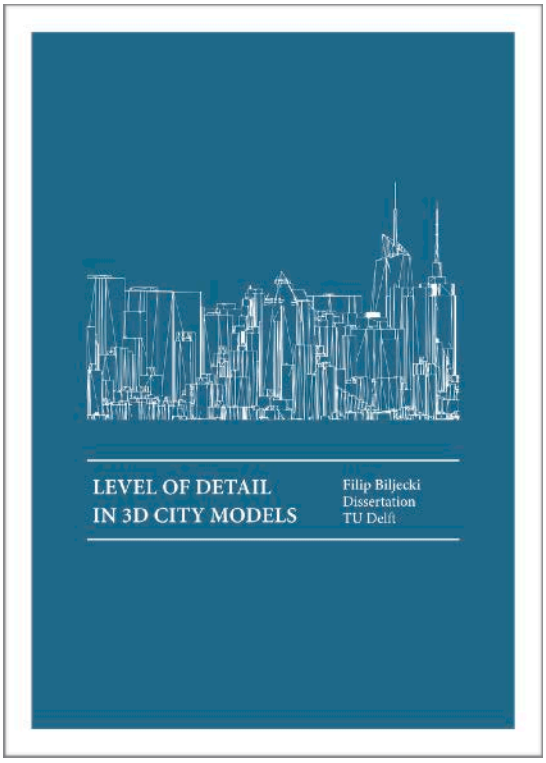
Please note that this is a compressed PDF of the dissertation

Biljecki F (2017): Level of detail in 3D city models. PhD thesis, TU Delft, 353 pp.
doi:10.4233/uuid:f12931b7-5113-47ef-bfd4-688aae3be248

To download the authoritative version, which contains high-quality figures and other material, please visit the TU Delft repository at http://doi.org/b463 (~99MB)



TU Delft
3Dgeoinfo

LEVEL OF DETAIL IN 3D CITY MODELS

Filip Biljecki
Dissertation
TU Delft

LEVEL OF DETAIL IN 3D CITY MODELS

Filip Biljecki

DELFT UNIVERSITY OF TECHNOLOGY

Doctoral dissertation

2017

LEVEL OF DETAIL IN 3D CITY MODELS

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof. ir. K.C.A.M. Luyben,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op
maandag 1 mei 2017 om 10:00 uur

door

Filip BILJECKI

Master of Science in Geomatics,
Technische Universiteit Delft,

geboren te Zagreb, Kroatië.

This dissertation has been approved by the

Promotor:	Prof. dr. J.E. Stoter
Copromotor:	Dr. H. Ledoux

Composition of the doctoral committee:

Rector Magnificus	Chairman
Prof. dr. J.E. Stoter	Delft University of Technology, promotor
Dr. H. Ledoux	Delft University of Technology, copromotor

Independent members:

Prof. dr. R. Billen	University of Liège, Belgium
Prof. R.J. Dijkstra	Delft University of Technology
Prof. dr. T.H. Kolbe	Technical University of Munich, Germany
Prof. dr. A. van Timmeren	Delft University of Technology
Dr. C. Ellul	University College London, United Kingdom

The research described in this thesis was carried out at the 3D Geoinformation Research Group at the Delft University of Technology (52° 0′ 17″ N, 4° 22′ 13″ E), and is a part of the project '5D Data Modelling: Full Integration of 2D/3D Space, Time and Scale Dimensions'. This work is part of the research programme Innovational Research Incentives Scheme with project number 11300, which is financed by the Netherlands Organisation for Scientific Research (NWO).

Typeset by the author in Minion Pro using XeTeX.

Cover design by Linda van Zijp (StudioLIN, Rotterdam). The image shows the geometry of an excerpt from the 3D city model of New York City (open data by the NYC Department of Information Technology and Telecommunications).

Printed in the Netherlands by Ridderprint.

ISBN 978-94-6186-800-8

An electronic version of this dissertation is available at http://repository.tudelft.nl

CONTENTS

ACKNOWLEDGEMENTS

A major perk of carrying out PhD research is collaborating with many enthusiastic and smart people. I was quite fortunate to have met many of them, and they have had a significant role in contributing to the work behind this thesis.

First and foremost, I could not be happier to have had Jantien Stoter and Hugo Ledoux as my supervisors. I have the utmost gratitude for their great mentorship that has played a monumental role during my doctorate. They have guided my research in the right direction, while simultaneously according me a great degree of freedom, thus enabling me to explore topics of interest for me and to fulfil my scientific curiosity. Jantien, you have kindly taken me on board for your Vidi project, and your leadership of the 3D Geoinformation Group has created an incredible working and social environment with a broad range of knowledge. I thank you for everything. Hugo, I learned a great deal about GIS and research from you, and the rigorous level of detail (pun intended) in your impeccably pertinent comments on my drafts is greatly appreciated. Above all, you have made our collaboration very enjoyable, cheers for that.

During the project I have closely collaborated with Ken Arroyo Ohori, who has also completed a PhD in the same group and project. His help and inexhaustible enthusiasm are acknowledged with great pleasure.

I gratefully acknowledge the Dutch taxpayers and the Netherlands Organisation for Scientific Research (NWO) for financing this project and my employment. I extend my gratitude to the partners in the project's users committee for their participation and the discussions: 1Spatial, Bentley, City of Amsterdam, City of Den Haag, City of Rotterdam, Esri Nederland, Kadaster, and Rijkswaterstaat.

Junqiao Zhao (Tongji University) contributed to the theoretical framework published in the paper forming the basis of Chapter 3. I thank him for the ideas he has brought to the project and for the fruitful discussions. Considerable appreciation goes to Gerard Heuvelink (Wageningen University) for mentoring me in the world of error propagation, leading to the publication of two nice journal papers (Chapters 11 and 12). George Vosselman (ITC, University of Twente) is gratefully acknowledged for sharing his expertise in 3D acquisition, and this collaboration resulted in a journal paper (Chapter 5). I owe gratitude also to Arzu Çöltekin (Uni-

versity of Zurich) for her contribution to our journal paper on applications of 3D city models (Section 2.3).

The members of the doctoral committee may have joined this project at its very end, but nevertheless with an important task. Great appreciation is paid to Roland Billen (University of Liège), Rients Dijkstra (TU Delft), Claire Ellul (University College London), Thomas H. Kolbe (TU Munich), and Arjan van Timmeren (TU Delft) for taking the time to read this thesis and for participating in the examination. I am honoured to have such an esteemed committee.

I experienced a marvellous amount of enjoyment in carrying out this work in the 3D Geoinformation Group. Thanks a million to my colleagues Abdoulaye Diakité, Anna Labetski, Jinjin Yan, Kavisha, Liu Liu, Ravi Peters, Sangmin Kim, Sisi Zlatanova, Stelios Vitalis, Tom Commandeur, and Zhiyong Wang. Furthermore, I am grateful that I could exploit Anna—our only native English speaker in the group—for spicing up a substantial portion of my thesis. Here I cannot go without also mentioning Veronica Alfano and Maaike Beliën (Faculty of Technology, Policy and Management) for their excellent training in writing and for checking my manuscripts. Needless to say, I am grateful to many other colleagues from the department, from the faculty, and from other faculties that make TU Delft such a great place to work. I am glad to emphasise the administrative staff for their dedicated work: Margo van der Helm, Daniëlle Karakuza, Linda van Keeken, and Ismail Yetim.

Thanks to Elfriede Fendel, Tjeu Lemmens, Martijn Meijers, Amin Mobasheri, Peter van Oosterom, Wilko Quak, Radan Šuba, Theo Tijssen, Edward Verbree, and Marian de Vries of GIS Technology for their contribution in the initial phase of the project.

In early 2016 I had the pleasure of conducting a research visit at the Singapore Land Authority. Many thanks to Kean Huat Soon, Victor Khoo, Sandy Teo Shen Ni, Richard Loo, Tong Wee Aw, Hock Kee Teo, David Tay, and Kheng Peng Soh. Beyond enabling a pleasant working environment, they have been quite welcoming particularly in offering a 101 on the local cuisine, culture, and lingo. I also warmly thank Susan van Boxtel from the Netherlands Embassy in Singapore for her interest in my stay.

Similarly, I appreciate my short yet productive stay at the HFT Stuttgart. Thanks to Romain Nouvel, Ursula Eicker, and Volker Coors for hosting me at their university, and the Erasmus+ Staff Mobility Programme for funding the trip.

The research described in this thesis has been awarded the Young Researcher in GIScience Award by the Austrian Academy of Sciences. I am honoured to receive this prize, and I thank the Academy and professors Thomas Blaschke and Josef Strobl for this recognition.

I am indebted to the creators of the open content used in this thesis: including the developers and contributors of open source software and authors of open data, such

as volunteered geoinformation, local authorities, and national mapping agencies. Here I should especially highlight the mind-boggling datasets provided by the Dutch Kadaster, Statistics Netherlands, AHN, the authorities of the Hague and Rotterdam, and the NYC dataset that is featured on the cover of this book. A large portion of the work in this thesis, especially Chapter 7 and Chapter 9, would not be possible without this contribution. Companies behind commercial software and other data are also gratefully acknowledged for providing free access to their products.

This thesis is based on my papers (pp. 350–353) that have been published in journals and conferences over the course of the past four and half years. Editors and anonymous reviewers (including Reviewer #3) are gratefully acknowledged for their insightful comments, which as a result have improved the content of this thesis.

The members of the OGC CityGML Standard Working Group, EuroSDR 3D Special Interest Group, Eurographics, and groups within the ISPRS are kindly acknowledged for the numerous challenging and insightful discussions. I especially appreciate the collaboration with Gerhard Gröger (University of Bonn & CPA Software), Karl-Heinz Häfele and Joachim Benner (Karlsruhe Institute of Technology), Marc-Oliver Löwner (TU Braunschweig), and Claus Nagel (virtualcitySYSTEMS) resulting in the proposal of a new LOD concept in CityGML 3.0, among other things. Furthermore, thanks are indebted to Jonas Teuwen (Radboudumc) for the discussions about mathematics and machine learning, and to Giuseppe Peronato (Swiss Federal Institute of Technology in Lausanne), Mickaël Brasebin (IGN France) and Kian Wee Chen (Singapore–MIT Alliance for Research and Technology) for exchanges related to environmental modelling and software.

I had the pleasure to be in frequent contact with many other people and institutions in the Netherlands and abroad who have impacted my work in a multitude of ways both large and small, my appreciation is also extended to those who are not listed above.

On a personal level, support and continuous interest from my friends, mostly in the Netherlands and Croatia but also elsewhere, is very much appreciated. I hope that this book will help them understand what I have been going on and on about for most of the time. While I cannot list them all I trust that everyone will feel recognised for their valuable contribution.

Undoubtedly the biggest thanks go to my closest companions as their support was essential in the completion of this doctorate. Tuğçe, thank you for your unconditional patience and for sharing this journey with me. My heartfelt gratitude goes to my parents Jasminka and Zvonko, brother Stefano, and the rest of my family for their continuous support. This thesis is also their accomplishment, hence I dedicate it to them.

Delft, March 2017 *Filip Biljecki*

CHAPTER 1

Introduction

1.1 CONTEXT [1]

Cities are increasingly adopting 3D city models. Providing further value and additional utility over 2D geo-datasets, 3D city models are becoming ubiquitous for making decisions and for improving the efficiency of governance. Local governments use 3D city models for urban planning and environmental simulations such as estimating the shadows cast by buildings, investigating how the noise from traffic propagates through a neighbourhood, and predicting how much solar irradiation a roof of a building receives in order to assess whether it is economically feasible to install a solar panel.

Similarly as traditional 2D geo-datasets, 3D city models are an approximation of the real world: features are modelled at a particular grade and certain elements are simplified or omitted. The quantity and mixture of content is driven by the intended use of the 3D city model, provenance of the base data, acquisition technique, invested funds, and spatial scale (Figure 1.1). The amount of detail that is captured in a 3D model, both in terms of geometry and attributes, is collectively referred to as the level of detail (LOD), indicating how thoroughly a spatial extent has been modelled. As a result, the LOD is an essential concept in geographical information science (GIS) and 3D city modelling.

In fact, the LOD concept is important in all steps of a typical 'life cycle' of a 3D city model, even prior to any acquisition having taken place. It is a significant factor in *contracting*—the LOD is frequently found in tenders to describe the characteristics of the desired 3D city model, e.g. whether roads or roof shapes of buildings shall be modelled. Hence, it can be considered as an element of public procurement—datasets are differentiated by the LOD and the amount of data that ought to be collected sets their value. Having the LOD in mind when *planning* the acquisition of data is essential for proper budgeting of resources, and the LOD determines the acquisition technologies that ought to be employed as different LODs are a result of different data acquisition approaches, e.g. it drives the minimum point cloud density when using airborne laser scanning, and determines whether a particular acquisition technique is sufficient or requires additional acquisition means. In the *acquisition* of the data, the LOD further serves as the principal instruction on how thoroughly to acquire the data. In fact, the term LOD can already be found in the earliest papers related to the acquisition of 3D city models [2].

Furthermore, not all 3D datasets represent a model of the current real-world; some of them are a result of *design*, e.g. in architecture to plan future scenarios. The granularity of planning is described by the LOD, meaning that the concept is appli-

[1] Some paragraphs in this section are partly based on my paper [1] Biljecki F, Ledoux H, Stoter J (2014): Redefining the Level of Detail for 3D models. *GIM International*, 28(11): 21–23.

Figure 1.1: Examples of 3D city models from around the world in different levels of detail. Acquisition techniques and many other factors such as the intended purpose of the data drive the nature of the dataset. The concept of LOD does not encompass only the appearance and visual detail, but also several other underlying factors that will be revealed in this thesis. Locations (clockwise from top left): Ettenheim, Germany; Punggol, Singapore; Chongqing, China; Adelaide, Australia; National Chiao Tung University, Hsinchu, Taiwan; New York City, United States.[2]

cable to fictional data as well. Since spatial analyses may be carried out for 3D city models that represent a setting that is yet to be constructed [3], they are important in GIS and therefore they should also be considered in the context of LOD.

Processing the data is influenced by the LOD as well. As I will demonstrate in

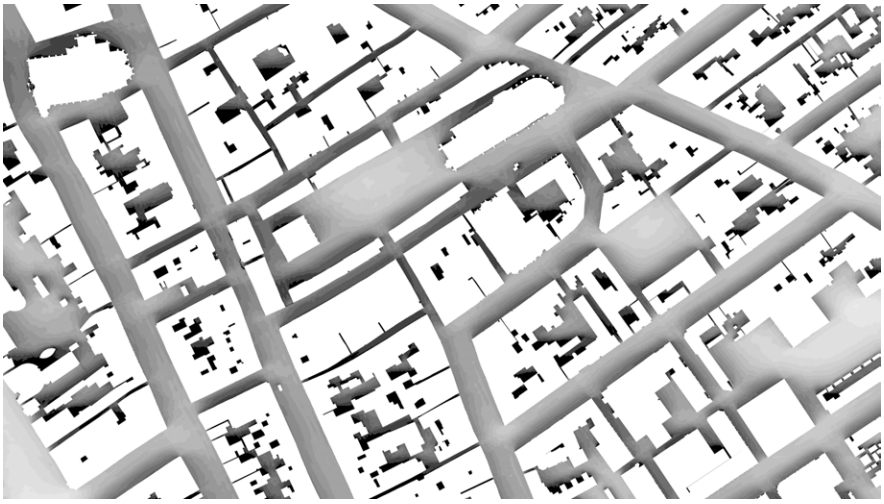


Figure 1.2: 3D city models are nowadays used for dozens of purposes not possible with 2D data. This example illustrates the predicted shadow levels on the ground as shaded from buildings (ortho view from above)—darker areas represent the surface that is shaded more over the course of a year. Spatial analyses such as this one serve several application domains, e.g. analysing thermal comfort.[3]

this thesis, the LOD does not refer only to the amount of geometric data, but also to the semantic richness. For example, data is subject to *conversion*, which usually depends on its LOD. After the dataset is acquired, the LOD influences the *storage* aspect, substantially influencing the storage footprint, and necessitating compression and integration techniques. *Quality control* is another aspect in the life cycle of the 3D city model where the LOD is consulted in order to ensure that all bits of data have been presented according to the specified LOD. Once the data is ready for *dissemination*, the LOD drives aspects such as the exchange of data, materialisation (3D printing), streaming, and delivery; topics all relevant for *interoperability*.

[2]Data courtesy of: Institute for Applied Computer Science, Karlsruhe Institute of Technology [4, 5]; CENSAM, Singapore-MIT Alliance for Research and Technology and OpenStreetMap contributors [6]; Chongqing Survey Institute; Adelaide City Council; Department of Civil Engineering, National Chiao Tung University [7]; Department of Information Technology & Telecommunications (DoITT) of The City of New York.

[3]Credit: The 3D city model of Delft is the result of the work of our group: generated with our software 3dfier by combining the BGT (large-scale topographic dataset of the Netherlands) and the AHN (the Actual Height Model of the Netherlands) datasets. The analysis was performed with Blender and VI-Suite (http://arts.brighton.ac.uk/projects/vi-suite).

Eventually, the data is delivered to users, who employ it for an *application* where the LOD may affect the performance and reliability of a spatial analysis (Figure 1.2). Data in higher (finer) LODs is believed to bring more accurate result in spatial analyses, but at the expense of a higher cost of acquisition. Hence it is important to benchmark their value and performance when used in an application. Because 3D city models are widely used in several spatial analyses, it is critical to give insight into their performance when employed for a particular purpose.

Since most workflows involve portraying the data, *visualisation* is also an important aspect where the LOD plays a prominent role in balancing the cognitive and performative aspects. Finally, after these steps have been completed, 3D city models are subject to *maintenance and update* in which the data is updated to reflect the latest real-world situation. In this process it has to be ensured that the refreshed data is of the same LOD as the original data.

1.2 PROBLEM STATEMENT

Despite the importance and omnipresence of the LOD concept in 3D city modelling, this topic has not been investigated extensively, and there has been no *holistic* research that encompasses the complete pipeline as described in the previous section. Most importantly, the term LOD has often been used interchangeably with scale, accuracy, and quality, and it has been largely used colloquially to generally point out the *richness* of a geographical dataset, without standardisation and formalisation. Moreover, it has not been determined what constitutes the LOD of a dataset, and the view of practitioners on it differ. Unlike 3D computer graphics where the topic has been thoroughly investigated in the past decades and where 3D models are differentiated by the amount of polygons, LOD in GIS encompasses semantics, attributes, and other aspects. Furthermore, there are challenges in quantifying the LOD. Unlike the explicit notions of scale in cartography, triangles in computer graphics, and resolution in rasters, determining the level of detail in 3D city modelling is a subjective task [8, 9].

There have been attempts to standardise the LODs to differentiate datasets. For example, the Open Geospatial Consortium (OGC) standard CityGML 2.0 [10], a prominent 3D GIS standard that will be described in Chapter 3, describes five LODs (Figure 1.3) to structure the geometric and semantic characteristics of 3D data. However, this thesis will demonstrate that these LODs are limited and ambiguous. At best, the understanding of the LOD concept stops at describing the grade and design characteristics of the 3D dataset, and there is not a lot of research on its influence on the post-acquisition phase of 3D data: maintenance, update, and application. For example, it is widely believed that a dataset with more detail is more valuable for

Figure 1.3: The five LODs of the OGC CityGML 2.0. The standard attempts to assign standardised classes to generally differentiate grades of 3D data. The geometric detail and the semantic complexity increase with each level. This LOD categorisation is well known in the 3D GIS community, however, it is not without shortcomings, and it is therefore thoroughly discussed in this thesis.

spatial analyses, but this presumption has not been well investigated.

As a consequence, all steps in the life cycle described in the previous section have been burdened by a lack of research in the domain of the LOD. For example, tendering documents usually specify the LOD with a short and ambiguous meaning. Hence the acquisition of datasets is prone to misunderstandings due to the absence of formalisation and a clear specification of the data that is planned to be contracted and acquired. Once acquired, processing the data without the understanding of the LOD may introduce errors, and quality control cannot be performed rigorously if an LOD standard is not present. The lack of understanding of 3D spatial analyses and how different LODs influence them may also result in errors, which may lead to wrong decisions in application domains. For example, a 3D city model produced with too coarse a LOD may mislead a user to believe that a rooftop is sufficiently insolated[4] to buy a solar panel, while in reality it is not. Another example is the shadow analysis shown in Figure 1.2. This analysis has been performed with a 3D dataset of coarse detail (LOD1 block models according to CityGML 2.0; see Figure 1.3). It is not clear whether modelling buildings in more detail (i.e. obtaining an LOD2) would bring an improvement in this spatial analysis. Such fitness for use of each LOD may depend on an application, and this topic has not been investigated to a great extent.

Lack of research in this topic may have further repercussions: it may lead to the acquisition of an inadequate dataset, not only insufficiently detailed, but also overly detailed, leading to unnecessary costs that do not bring a tangible benefit. As finer LODs trade the ease of acquisition and automation for realism and veracity, the re-

[4]Insolation, the exposure to the sun's rays, should not be confused with *insulation*, the action or state of keeping something insulated from unwanted loss of heat or from the intrusion of sound by interposing adequate material.

lationship is not always clear and may be unfavourable to the user. Finally, the lack of a well founded framework on the LOD concept inhibits maintaining and updating a dataset. For example, keeping the LOD consistent during maintenance is of utmost importance for change detection. Otherwise, an arising detail in the new version of a dataset may present an old feature that was not acquired in the previous acquisition campaign, rather than a change in the real world.

1.3 RESEARCH QUESTIONS

In this thesis I explore various topics related to the concept of level of detail in 3D city modelling. The main research question that I seek to answer in this thesis is:

> How should we consider, integrate, and improve the concept of level of detail in 3D city modelling?

The research question is subdivided into multiple research questions. The theoretical part of the thesis investigates:

1. What is already known about the concept of LOD in 3D city modelling?
2. How can we formalise the concept of LOD in 3D city modelling?
3. How can we design a consistent and unambiguous LOD specification?
4. Are there multiple valid variants of the specified LODs?

Implementation-wise I enquire into the following aspects:

5. How can we realise the developed LOD specification?
6. How can we increase the LOD of existing data?
7. How can we integrate multiple LODs of the same feature?

On the experimental side, the research covers the following questions:

8. How does the LOD influence the quality of spatial analyses?
9. How do acquisition errors influence LODs?
10. How can we distinguish errors induced by LOD specification and errors induced by the acquisition?

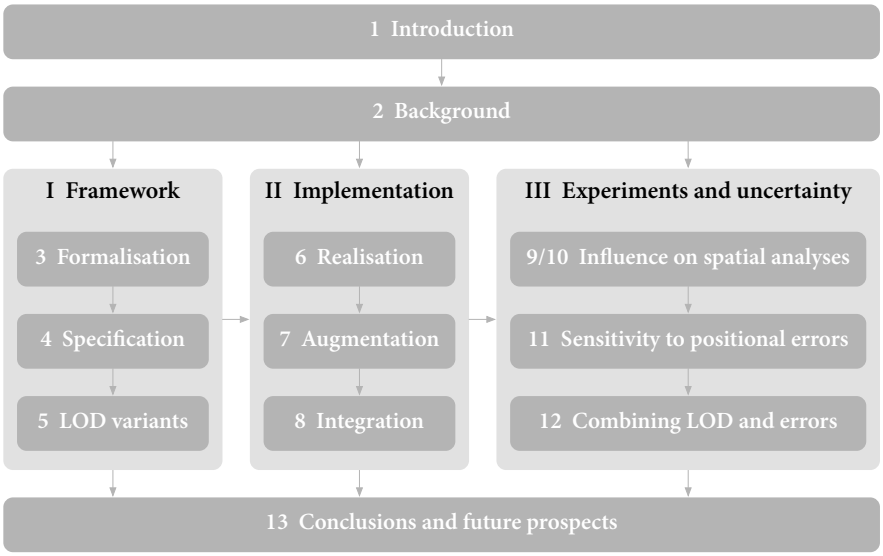


Figure 1.4: Organisation of the thesis. The report is divided into three parts.

1.4 OUTLINE OF THE THESIS AND SCOPE

This thesis is based on papers I have published during the course of my PhD research (listed on pp. 350–353). Some parts have been updated since their original publication. For example, the literature review has been extended to reflect the latest research findings. Furthermore, some papers have been modified, e.g. distributed across multiple chapters to enhance the flow of the thesis.

The thesis is organised into 3 parts and 13 chapters (Figure 1.4), as follows:

Chapter 2 discusses the fundamentals that are crucial to understand prior to delving deeper into the topic: the type of 3D geoinformation I focus on, how the data is acquired, and what are the applications of 3D geoinformation. This thesis focuses on spatial analyses and concrete examples of 3D geoinformation put into operation. Hence an emphasis is put into cataloguing uses of 3D city models.

The subsequent chapters are organised into 3 parts.

1.4.1 Part I: Framework and specification

Chapter 3 formalises the LOD concept in 3D city modelling. First, it investigates the current status of the LOD concept, present practices of stakeholders in the 3D GIS

community, and identifies the main shortcomings of the current LOD approaches. Next, the chapter discusses the relationship of the LOD concept to its progenitor in computer graphics, which has well developed mechanisms and theories about this subject, but these differ significantly from GIS. The chapter provides a framework that enables the quantification, description, and translation of presently available LOD standards, providing a more expressive approach to convey the LOD. Using this framework, communication of the fineness of data may be facilitated, and it may serve as a base for defining a standard for measuring the LOD of 3D city models. While the chapter provides a strict and unambiguous framework, its (academic) rigidity may not be practical because it may hinder its acceptance in practice. That leads us to the next chapter.

Chapter 4 is one of the core contributions of this research: it presents an improved LOD specification for buildings intended for stakeholders in the GIS community. It is based upon the LOD concept in CityGML 2.0 (Figure 1.3), solving its shortcomings, which I have identified after extensive discussions with researchers and practitioners, and examining documentation such as research papers and open data specification. The two main deficiencies of the standard are that each CityGML 2.0 LOD lacks a precise specification, and that the five LODs are too generic and therefore insufficient for differentiating distinct variants of the same LOD. The LODs defined in CityGML have been re(de)fined into a larger number of LODs that are supported by a firmer definition. Amongst others, a benefit of such a specification is helping the industry to better define their products, and customers (e.g. mapping agencies) to express their needs and specifications.

After establishing the LOD specification, *Chapter 5* investigates whether there are multiple valid variants of the same LOD, depending on the acquisition technique that is employed to generate a 3D city model. The chapter exposes several different variants of data that are of the same LOD, naming them geometric references. While these are distinct from the LOD, they play a hand in the specification; hence the chapter provides an addendum to the specification, and contributes with raising awareness of this topic.

1.4.2 Part II: Generation and management

The second part of the thesis focuses on realising the concepts presented in the first part. *Chapter 6* presents an approach to realising 3D city models in the newly defined LOD specification with procedural modelling using a software prototype that I have developed during the research. The chapter doubles as a source of multi-LOD data in the experiments carried out in the final chapters of the thesis.

In GIS, generalisation is used to reduce the LOD of a dataset: from a finer to a coarser LOD, thus removing complexity while preserving usability. The inverse

path, *augmenting* the LOD, has not been investigated thoroughly. *Chapter 7* introduces a novel approach to augment LOD using machine learning techniques. It results in generating 3D city models from inferior data (2D datasets without elevation measurements), being useful in regions rich with building footprints but missing elevation data.

Chapter 8 concludes this part of the thesis by investigating two approaches to ensure the consistency of multi-LOD datasets. The first approach focuses on linking the geometry across multiple LODs and preserving the topology. The implementation indicates that the method doubles as a compression method reducing storage footprint of multi-LOD data. The fundamental idea of the second approach is similar, with the difference being that the linking is established through modelling LOD as a spatial dimension perpendicular to the three standard spatial dimensions, resulting in a four-dimensional construction.

1.4.3 Part III: Experiments and uncertainty

Chapter 9 introduces the experimental portion of the work by answering the research question 'how does the LOD influence the quality of spatial analyses?'. In this chapter, geoinformation in different LODs is used for the same spatial analysis: estimating the population in the Netherlands at a fine scale for refining demographic maps; the results using different LODs are then compared. The chapter reveals the differences between two or more LODs when employed in a spatial analysis, and it contributes to the topic of population estimation using GIS and remote sensing by carrying out a large-scale study involving different datasets. While the experiments give good insight into the impact that different LODs have on a spatial analysis, the chapter has also the function of acting as a prelude to the succeeding chapters: it exposes several different problems with such analyses when using real-world data.

Chapter 10 continues the same line of thinking as the previous chapter, however, it uses procedural data generated in Chapter 6 avoiding difficulties induced by using real-world data. Three spatial analyses have been tested investigating the influence of different LODs on the result of a spatial analyses. It is demonstrated that procedurally modelled data is more suitable for such experiments than using real-world data. Furthermore, the chapter contests the often quoted assumption that finer LODs bring better results, and it compares the results of data modelled in different geometric references.

Real-world datasets always contain errors. *Chapter 11* investigates the influence of positional errors induced during the acquisition process. It introduces a novel approach for estimating the propagation of uncertainty in 3D GIS using intentionally degraded procedural models in repeated Monte Carlo simulations.

Chapter 12 revisits the previous two chapters, and investigates the relationship be-

tween the two types of errors—those induced by different LODs and those induced by acquisition techniques. The LOD concept is closely related to errors; given that acquisition techniques that are capable of producing data at finer LOD are usually more accurate, this results in a strong association between the level of detail and level of accuracy. However, inverse situations (i.e. coarse detail acquired with high accuracy, and vice-versa) are frequently becoming the case with the advent of volunteered geoinformation and new acquisition techniques. Hence, I raise questions about the relationship between the two and the relative contribution of each to the quality of a spatial analysis. The results of this chapter provide insight into whether it is worth acquiring a finer LOD if the acquisition method is not proportionally accurate, and vice-versa.

Chapter 13 concludes the thesis with the key takeaways, answers to the research questions, main contributions of the research, and proposes a roadmap for future work.

1.4.4 Scope

Each chapter defines its scope and possible extensions for future work. In general, from the thematic point of view, this research focuses on buildings as the most dominant feature of the urban environment. The research concentrates on outdoor 3D city models, a distinction that is defined in Chapter 2. From the implementation perspective, the work focuses on CityGML.

1.4.5 Personal pronouns

This thesis is based on my first-author publications in which I am the leading contributor but not the sole author. Hence I use 'we' as a courtesy to my co-authors when I refer to the experiments and results initially published in the papers. In contrast, I use 'I' in parts exclusive to this thesis, such as answering the research questions.

1.4.6 Open science

The main software that I have developed in this research is released as open-source. Likewise, the produced datasets have also been publicly released as open data. Furthermore, all the publications that form the basis of this thesis (and the thesis itself) are available as open access in the TU Delft repository.

1.5 PROJECT

The research described in this thesis is part of the project '5D Data Modelling: Full Integration of 2D/3D Space, Time and Scale Dimensions'. The project is funded by the Netherlands Organisation for Scientific Research (NWO).

The aim of this research project is to investigate the integration of the multi-dimensional characteristics of geographic data, i.e. 2D/3D geometry, time, and LOD at a fundamental level of data modelling. This thesis sheds light on the LOD aspect of the project.

The project was carried out in collaboration with several partners: 1Spatial, Bentley, City of Amsterdam, City of Den Haag, City of Rotterdam, Esri Nederland, Kadaster, and Rijkswaterstaat. Furthermore, the project was closely associated with the activities of the European Spatial Data Research (EuroSDR) and the Open Geospatial Consortium (OGC).

CHAPTER 2

Background

This chapter introduces fundamental concepts in 3D city modelling. First, it presents different types of 3D geographic information, and it aspires to define 3D city models. Second, it gives an overview of current mechanisms for producing 3D city models. Third, it covers their usability in different application domains. Finally, the chapter describes the concept of level of detail in building information modelling, a cognate subject of relevance in this thesis.

2.1 3D GEOINFORMATION

Three-dimensional geoinformation is data that describes geographic features in 3D space with a set of (x, y, z) coordinates. This general definition results in encompassing a broad notion of different forms of data, such as movement trajectories in 3D space, digital elevation models (DEMs), and 3D geological models.

One of the subsets of 3D geoinformation are 3D city models, which are usually defined as a representation of an urban environment with a three-dimensional geometry of common urban objects and structures, with buildings as the most prominent feature [11–14]. This idea leaves some ambiguity leading to different interpretations of what 3D city models are (Figure 2.1). For example, some researchers consider point clouds as 3D city models [15], while others include polygon meshes in that context [16]. Moreover, the definition is complicated by the fact that different types of data may be combined together [17].

In this thesis I define 3D city models as structured objects described by their boundary surfaces that may be semantically enriched (bottom right example in Figure 2.1), akin to the OGC CityGML standard.

However, in most other types of 3D geoinformation, LOD has an analogy. For example, point clouds are characterised by density and voxels by resolution. Both may be considered as equivalents to the LOD concept in 3D city modelling. Hence, while the thesis focuses on semantic 3D city models, most of the developed concepts and frameworks are, to an extent, also applicable to other types of 3D geoinformation.

2.2 ACQUISITION OF 3D CITY MODELS

2.2.1 Overview

3D city models may be produced with a variety of techniques and approaches. Understanding acquisition techniques is paramount when discussing the concept of LOD, as the LOD of a model is a direct result of the capabilities of an acquisition technique and other parameters in the procurement of data. In short, more detailed 3D city models require more labour and entail a reduced degree of automation [18]. However, due to different technologies and approaches involved, the delineation of acquisition techniques is not unambiguous and there have been efforts to derive taxonomies, e.g. analysing acquisition techniques for sourcing cadastre information [19, 20].

The approaches can loosely be categorised by platforms [21]: satellite (spaceborne) [22], airborne, unmanned aircraft systems (UAS; i.e. drones) [23, 24], mobile mapping [25–27], ground (static), handheld devices [28, 29], and crowdsensed [30]. Another viewpoint to distinguish acquisition approaches is by the type of the tech-

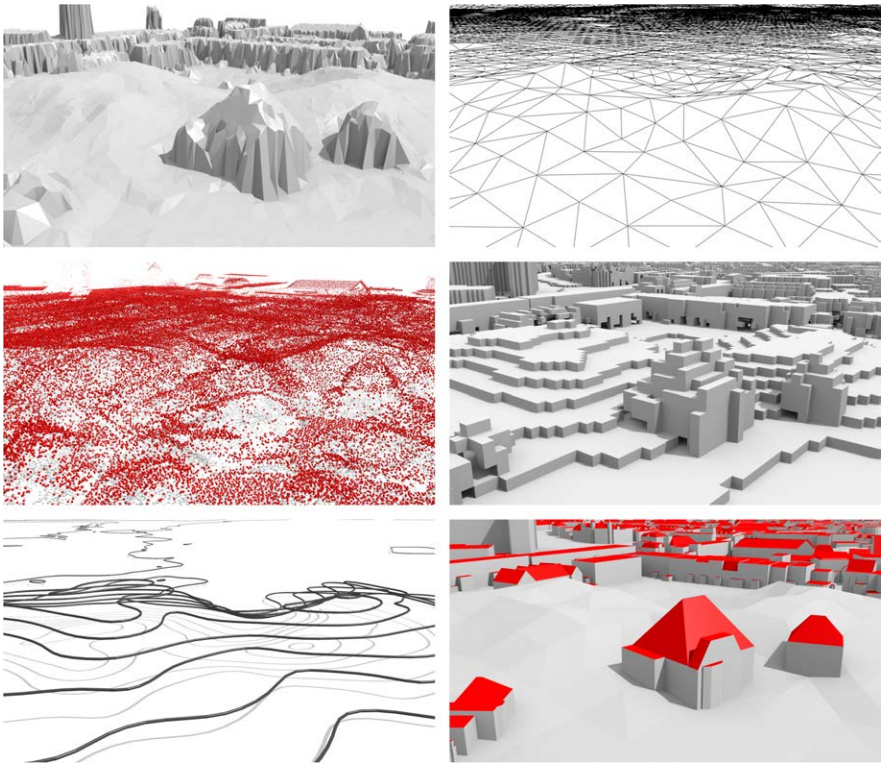


Figure 2.1: Different types of 3D geoinformation covering the same spatial extent. Clockwise from top left: a dense digital surface model, digital terrain model, voxels, a semantically structured 3D city model composed of boundary surfaces, traditional map contours visualised in 3D, and a point cloud derived from airborne laser scanning.[5]

nology (sensors). These are most prominently: lidar, radar, camera (photogrammetry), and total stations [31–33].

The permutations between platforms and technologies result in a number of approaches. For example, terrestrial acquisition may be performed with terrestrial laser scanning [34] or photogrammetry [35]; crowdsourced imagery from mobile phones and consumer cameras may be used to generate 3D geoinformation [36, 37]; and stereo satellite imagery may be used to generate 3D city models [38]. Moreover, each of these allows for further segmentation, e.g. imagery may be divided into or-

[5]Base data courtesy of the City of the Hague, and Actual Height Model of the Netherlands (AHN3).

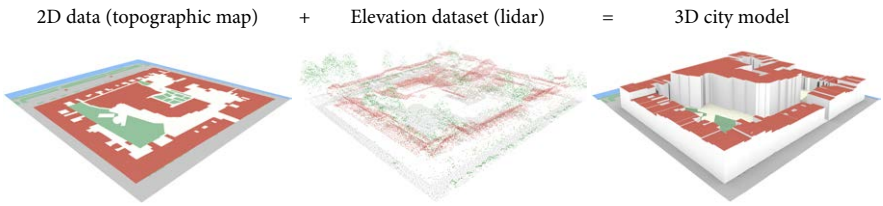


Figure 2.2: Extrusion of 2D datasets to derive 3D city models. This is a popular acquisition technique deriving simple block models (LOD1 according to CityGML), which—despite their crude detail—provide an increased advantage over 2D datasets, e.g. they can be used for visibility and shadow analyses (Figure 1.2). Notice that this method intrinsically results in flat roofs, in contrast to the 3D city model exhibited in Figure 2.1: the acquisition technique dictates how much detail is modelled. The point cloud (middle) is coloured according to different classes.[6]

tho and oblique; and both are used for the production of 3D city models, either alone or in combination with each other [39].

The combination of multiple techniques is also an important topic to consider when it comes to the LOD, primarily because certain workflows may result in datasets at a heterogeneous LOD, and because some LOD specifications can only be achieved with a combination of multiple techniques [40, 41]. For example, lidar and imagery may be used together [42, 43], and 2D GIS footprints are frequently used as a base when using airborne platforms [44]. In fact, the LOD1 model illustrated in Figure 1.3 (and the 3D dataset used for the analysis in Figure 1.2) is often obtained with a combination of a terrestrial survey of footprints of buildings, and airborne laser scanning (Figure 2.2). Additional acquisition techniques that cannot be placed in the observational context, as described above, include and are not limited to: obtaining 3D models from scanned 2D plans [45–47], from architectural CAD models [48], procedural modelling [49], by analysing shadows in imagery [50], and by analysing Global Navigation Satellite System (GNSS) signals [51]. Some of these methods, especially atypical methods such as generation from children's drawings [52] are not commonly used, but they are certainly interesting to mention to indicate the diversity and to give an impression of the challenges when delineating the approaches to acquiring 3D city models.

From the point of view of the LOD, I create a taxonomy of acquisition techniques that considers LOD as the focal point (Figure 2.3): direct acquisition, reduction, augmentation, and design.

[6]The datasets used to create the illustration are BGT and AHN.

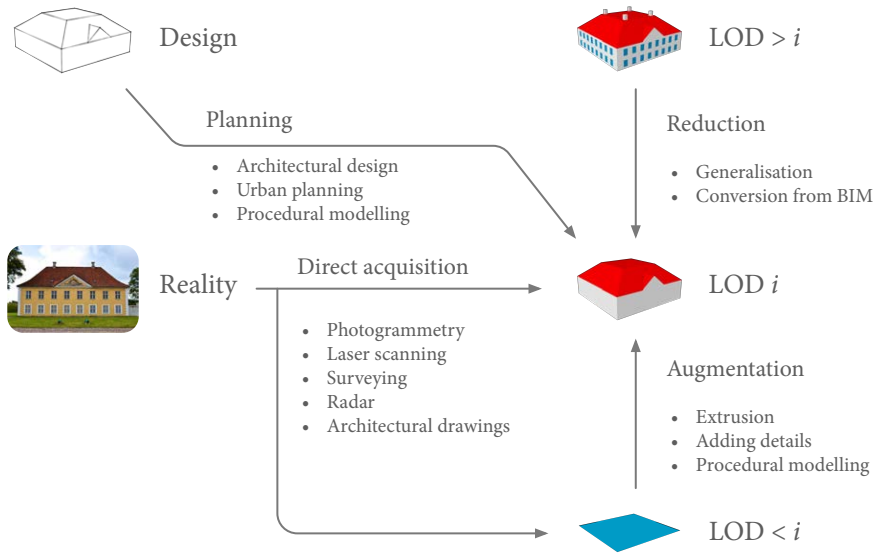


Figure 2.3: Production workflows of a 3D city model. Different acquisition approaches from the perspective of the level of detail.

2.2.2 Direct acquisition

This is the usual approach of deriving 3D city models: a subset of the real-world is modelled according to a predefined LOD, e.g. a photogrammetric survey is carried out to produce data at a specific detail. This category encompasses several techniques, such as airborne laser scanning and spaceborne radar.

2.2.3 Reduction

A dataset at an LODi is obtained by converting it from an architectural model or by generalisation from an existing 3D city model of a finer LOD ($> i$). This approach is employed for a multitude of reasons. Most notably: (1) if a 3D model is too complex for the intended application (spatial analysis or visualisation) it is generalised to facilitate the process [53]; (2) to reduce the storage footprint in a database [54]; and (3) to expedite streaming by reducing the amount of data that needs to be transferred [55]. Furthermore, there are other situations where it is important to reduce the LOD: (4) copyright aspects [56–58]; (5) automatic healing of 3D data models—if the repair fails, it is better to use a coarser (simpler) LOD that is valid than a

higher one that it is not [59, 60]; and (6) for privacy reasons, concealing sensitive details [61].

In the context of LOD it is relevant to note that the process of generalisation essentially results in multiple representations (the original one and its simplified counterpart).

2.2.4 Augmentation

Existing spatial data may be used as a base to generate data of a higher (finer) LOD. This process can be optionally aided by additional data sources. For example, if a footprint of a building is available, its height from a cadastral source can be used to generate a block model (extrusion; Figure 2.2). In comparison to the first two groups, this category has not been researched to a great extent, hence here I give an overview with a brief literature review.

Augmentation of the LOD involves its increase, and it may be considered akin to interpolation (the 'scale' of the data is increased during interpolation). Hence it is important to note that in such a case the quality may be degraded, despite the fact that the LOD is increased [62], an aspect which I will revisit multiple times during the course of the thesis.

The main reason to augment the LOD is because the current LOD is insufficient for a particular purpose. However, enhancing the dataset intrinsically requires the integration of additional data sources, which may be a hindrance. A literature review in relation to acquisition prompts grouping augmentation methods into three categories: (1) extrusion, (2) adding detail, and (3) procedural modelling.

Extrusion Extrusion is the most used augmentation technique. The concept of extrusion is that polygons in 2D datasets (LOD0 according to CityGML; see Figure 1.3) are extruded to a uniform height obtaining 3D volumetric models (LOD1) [63, 64]. In fact, it is one of the leading techniques in upgrading 2D to 3D [65]. Extrusion from 2D data has an advantage that the resulting 3D city model can be linked to the existing data [66].

The key information in extrusion is the height of a feature. There are many different ways to derive the vertical extent of a building. For example, taking a statistical measure such as the median or maximum elevation of lidar points that fall within a footprint. Other, albeit less common techniques, involve obtaining the heights with handheld rangefinding devices [67] and radar measurements [68].

Extrusion is the topic of Chapter 7 in which I present a new method, aided by machine learning, to extrude 2D datasets without elevation data.

Adding detail In this approach, an existing 3D city model may be taken as a base and enhanced [69]. For example, an LOD1 model may serve as a source in the photogrammetric mapping of roofs to produce LOD2.

A somewhat related method is to replace features, e.g. substitute certain parts in existing 3D city models with user-generated data that may be of a finer LOD [70].

Procedural modelling Procedural modelling is frequently employed to augment the LOD of existing data [71, 72]. For example, in some cases LOD1 models are produced by extrusion from LOD0, and are later embellished with roof shapes resulting in *faux* LOD2 models [73]. Furthermore, procedural modelling is often used in combination with other techniques, indicating that the taxonomy presented in Figure 2.3 inevitably contains blurry lines. For example, Edelsbrunner et al. [74] present a method to augment the LOD during the acquisition process where procedural modelling is used on the fly to aid the 3D generation.

Procedural modelling is the topic of Chapter 6 in which it will be described further, and it will be utilised as a technique in order to implement the theory presented in the first part of the thesis.

2.2.5 Design

A fraction of 3D city models are design models that do not represent an actual real-world setting, e.g. procedurally generated 3D models used in movies, architectural models of planned buildings, and models designed by urban planners to disseminate urban planning concepts [75]. However, such models also have their LOD, and they have not been used extensively in GIS. In this thesis (Chapter 10), I will demonstrate that their usability has been underestimated as they can be used in various GIS experiments where having real-world data is not essential. In fact, I argue that only synthetic datasets are viable for such experiments.

2.3 APPLICATIONS OF 3D CITY MODELS

2.3.1 Introduction

In the last decades, 3D city models have been predominantly used for visualisation. However, today they are being increasingly employed in a considerable number of domains for a large range of tasks beyond visualisation. As an example, Figure 2.4 illustrates four uses of 3D city models, utilising the dataset featured in Figure 2.1. Amongst other applications that will be overviewed in this section, these four analyses are a convincing case indicating the benefit of using 3D GIS, as some of them

are not possible with 2D GIS, or they provide an added value over 2D GIS (e.g. more accurate results and more insight).

Application of 3D city models is an important subject in this research, especially in Part III, which deals with the influence of LOD on the quality of spatial analyses. This section presents a condensed version of my journal review paper [76][7], which documented the state of the art of the utilisation of 3D city models across multiple domains based on a comprehensive literature study including hundreds of research papers, technical reports, and online resources. Section 2.3.2 describes the methodology and terminology used in the review, while Section 2.3.3 lists the identified uses of 3D city models.

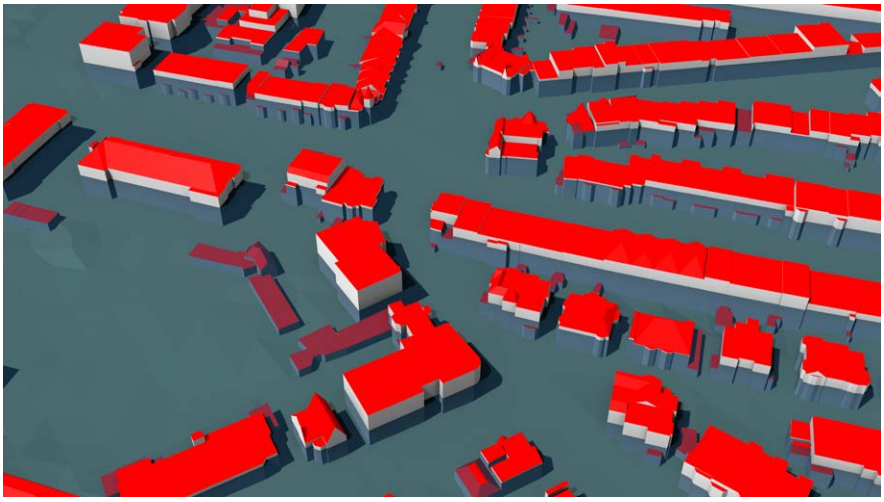
2.3.2 Methodology

The main methodology in this overview is a literature review and a synthesis. We have screened scientific literature, project reports as well as online resources on 3D geoinformation science with a focus on the utilisation of 3D city models in a comprehensive and systematic manner. The main challenge, encountered during the course of the review, is that terms such as 'use case' and 'application' are somewhat ambiguous in GIS-related resources, and that it was not immediately clear how to delineate different uses of 3D. A fuzzy terminology prevents an unambiguous organisation of the purposes of 3D city models.

The term *use case* has been defined in software engineering as a sequence of actions that provide an objective or subjective value to a user [77]. We second this definition: applied to GIS and 3D city modelling, a use case can be considered as a meaningful set of spatial operations that accomplish a goal a user wants to achieve with a spatial dataset. In this overview as in the rest of the thesis, use cases are viewed from the perspective that the user can *technically* arrive at her or his goal (that is, perceptual and cognitive aspects are not considered here). The use case is tied to a specific discipline (an industry or sphere of activity and knowledge) to which it may provide a substantial benefit. We infer the benefit based on the documented cases where people actually use the 3D city models in practice; this indicates to us that they at least see a subjective benefit in using the 3D models.

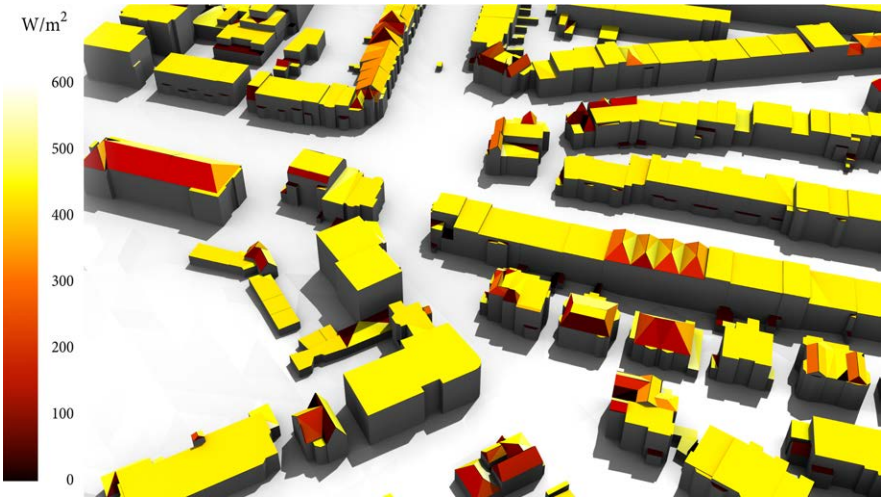
Following this definition, when a use case is employed in the context of a specific domain (e.g. archaeology) to solve an application problem we define an *application*. For example, the computation of the volume of a building is a spatial operation that is, among other operations, used in at least two use cases: estimating the energy

[7] Biljecki F, Stoter J, Ledoux H, Zlatanova S, Çöltekin A (2015): Applications of 3D City Models: State of the Art Review. *ISPRS International Journal of Geo-Information*, 4(4): 2842–2889. doi: 10.3390/ijgi4042842

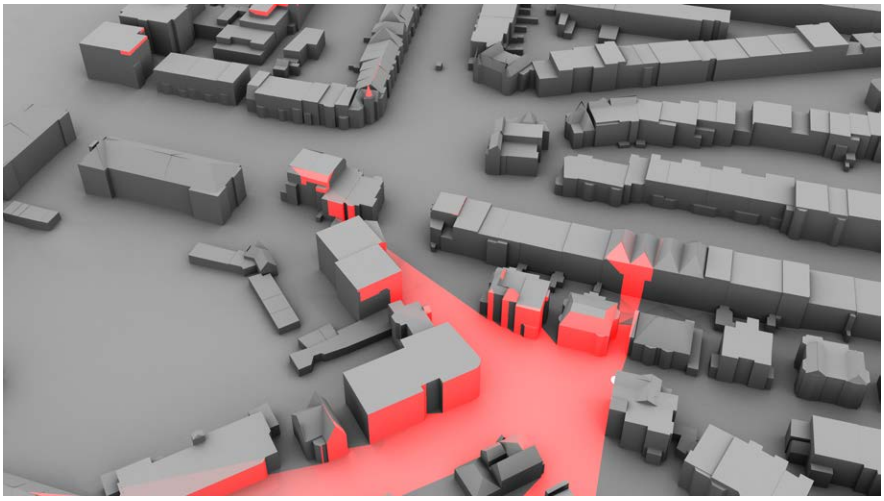


(a) Flood simulation. Carrying out flood simulations with semantic 3D city models enables deeper insight, such as understanding which buildings and which of their storeys will be flooded in a certain disaster scenario, which is in turn useful for applications such as insurance and disaster management.

(b) Estimating noise pollution. This is an illustration of noise from a planned tram track (red line) in an urban environment. Owing to analyses such as this one it is possible to predict how many citizens will be affected by noise if the tram line is built. This is a 3D analysis because noise propagates in 3D, and the analysis enables us to determine not only how many buildings are affected by excess noise pollution, but also which floors will be affected most.



(c) Estimation of solar energy of rooftops. Several application domains benefit from this spatial analysis in which 3D city models are indispensable. For example, it enables finding optimal locations to place solar panels, and those that are susceptible to overheating.



(d) Visibility analysis. This example showcases the use of 3D city models to predict the view from an apartment, benefiting applications such as security and mass valuation of real estate.

Figure 2.4: Application of 3D city models plays an important role in this research. These examples illustrate the use of the same 3D city model for different spatial analyses, benefiting multiple application domains. Understanding the data requirements, especially the level of detail, to suit different spatial analyses is paramount.[8]

demand (larger constructions require more energy to be heated) [78], and estimating the number of inhabitants of a building (larger buildings generally host more people) [79]. The latter use case is valuable in at least two application domains: for emergency response (estimating the number of people that have to be evacuated), and in environmental modelling (estimating the number of people affected by noise). Another example is the estimation of shadows cast by buildings on the surrounding area. This use case may be employed in adjusting the estimations of the solar potential of rooftops, but also in urban planning to assess whether a planned building, if built, would threaten a neighbour's access to sunlight.

In the subsequent overview, the focus is on listing use cases, but importance is also given to applications for a better understanding of use cases. Furthermore, as it will become evident from the list, some use cases have fuzzy boundaries, so a degree of subjectivity and personal choice has to be accepted by the reader.

2.3.3 List and description of use cases of 3D city models

Based on the methodology presented in Section 2.3.2, we have identified 29 distinct use cases used in several application domains. We describe and list the identified use cases in no particular order.

2.3.3.1 Estimation of the solar exposure

The estimation of the insolation of buildings is one of the most prominent use case in 3D city modelling (Figure 2.4c). It is a mature topic in GIS initially conducted on digital surface models [80]. However, the evolution of acquisition techniques and data models has enabled us to model buildings and their parts (e.g. roof), which has opened the door for multiple application domains requiring such a granularity.

3D city models are used to estimate how much a building is exposed to the sun in order to assess the suitability of installing solar (photovoltaic) panels on roofs [81–93]. 3D city models enable calculating geometric characteristics such as the tilt, orientation and area of the roof, which are used as the main input for the solar empirical models. Recent work has focused on extending the task for vertical façades, and taking into account the material of the receiving surface [94, 95].

This application may benefit from attributes such as the address and type of building for additional analyses [96], and it is being supported by an increasing number of software implementations [97, 98]. Furthermore, some researchers use digital surface models or dense point clouds rather than semantic 3D city models [99–103], indicating that the same application may be performed using different types of 3D

[8] Data courtesy of the City of the Hague. Analysis software: VI-Suite, Geomilieu–DGMR, and Blender.

geoinformation. For a comprehensive overview of research on solar potential applications see the recent review of Freitas et al. [104].

The estimation of the insolation of buildings is also vital to estimate the thermal comfort, i.e. the detection of buildings that are exposed to too much sunlight, potentially resulting in overheating during the summer [105, 106]. This use case also allows us to design an urban layout to maximise the insolation of a neighbourhood [107], and to estimate the capacities of decentralised energy sources for crisis management applications [108]. A further application is in the large-scale estimation of house prices. The information about the insolation can be used as one of the factors for estimating the property prices, under the assumption that solar radiation is capitalised in the value of a property [109]. Finally, 3D city models containing windows can be used to predict the indoor illumination and daylight autonomy [110, 111].

This use case will be a subject of experiments in Chapter 11, where it will be described in further detail.

2.3.3.2 Energy demand estimation

A use case that demonstrates the value of semantic information in 3D city models is the estimation of the energy demand of households.

Recent years have seen the advent of this application, where researchers, predominantly in Germany, have used 3D city models to combine data containing the volume of buildings, number of floors, type of the building, and other characteristics to predict the energy demand for heating and/or cooling [78, 112–123].

For example, estimating the energy demand is important for assessing the benefit of energy-efficient retrofitting. Previtali et al. [124] note the use of 3D city models in assessing the cost of retrofitting of a building. In combination with other data, 3D city models may be used for thermal assessment and to determine thermal bridges and heat losses from the building envelope. In a related retrofit planning analysis, Tabrizi and Sanguinetti [125] have used information related to the materials, weather data, and renewable energy sources.

2.3.3.3 Aiding positioning

Löwner et al. [126] and Cappelle et al. [127] present methods using 3D city models to improve positioning in urban environments. The rationale is that it is possible to derive a position from photographs if it is possible to match the same perspective from a 3D city model. This is useful for urban canyons where satellite positioning may be less reliable. Coors et al. [128] have developed a related method aimed at applications in tourism.

2.3.3.4 Determination of the floorspace

3D city models may be used for estimating the internal size of a building in the legal framework (i.e. net area, floorspace) [129–131]. The floorspace can be inferred from the exterior of a building, without indoor data. This use case has potential applicability in the taxation and valuation of buildings.

2.3.3.5 Classifying building types

Henn et al. [132] present a method for detecting the type of a building from its 3D geometry (e.g. apartment buildings and detached houses). The knowledge of the building type has an application in various domains. For example, the distribution and share of a building type in a neighbourhood is of interest in marketing and real estate management.

2.3.3.6 Geo-visualisation and visualisation enhancement

Visualisation is one of the fundamental purposes of 3D city models: it permits shape cognition and evaluation of complex spatial circumstances [133]. It is suggested that 3D city models generally provide an enhancement over 2D (map) data [134]. This use case is general and open-ended, since most of documented uses consist of visualising 3D data: e.g. for real estate [135], panoramic views [136], web visualisation [137–141], profiling [142], crime mapping [143, 144], serious gaming [145, 146], and augmented reality [147–157]. It is not our intention to further delineate each of these, as it entails more ambiguity with respect to the taxonomy. Therefore, only an overview of some applications follows.

3D city models are frequently used to enhance the presentation of results of analyses that are not necessarily related to GIS and 3D city models [158]. For example, economic activities [159], tsunami analysis [160, 161], and wind farms [162]. Such visualisations are meant to aid scientists analysing large amounts of data. Other analyses where researchers and practitioners have used 3D visualisation include human activity [163], wind fields [164], and air quality data [164–166]. 3D city models are also used in traffic and flight simulators [167], and for background and flythroughs in movies, documentaries, and news programs. Data used for these purposes is frequently procedurally generated [49, 168, 169].

2.3.3.7 Visualisation for navigation

3D city models, or sometimes 3D objects such as buildings on otherwise 2D visualisations, are used for facilitating the user's orientation in space for navigation purposes. Navigating urban spaces using 3D city models can help with orientation

as it offers familiar landmarks; and it has often been proclaimed that their 'more intuitive' nature in contrast to 2D maps provides more natural and realistic navigation cues [170–175]. At this point it is important to note that 2D aerial views (top views) are very important in navigation tasks as they provide overview information without occlusion as opposed to 3D views; and as mentioned earlier, they have a more consistent scale, and are thus better for distance estimation tasks. In a choice experiment, it has been recently demonstrated that people use 3D visualisations roughly 30% of the time for navigation tasks [176]. The more realistic representations appear to be helpful for rapid shape cognition, presenting the possibility that a mix of 2D and 3D views (multiple-linked views) are helpful in this case [177].

3D city models with semantic information provide added value in this use case as the visualisation can be enhanced to improve its function [178, 179]. For example, a landmark offers more navigational cues than a block of grey residential buildings, and can therefore be emphasised in the visualisation.

2.3.3.8 Visualisation for communication of urban information to citizenry

A regular application of 3D city models is to present the existing city and to disseminate urban information to citizens [180–183], and proposed developments and enhancements in a 3D virtual environment [184, 185]. For example, the model of the City of Adelaide in Australia provides a public consultation tool for assisting in visualising transport, urban design and planning [186]. Because most members of the general public are not urban planning professionals, visualisation should be carefully designed [180], and here it is noted as a distinct use case.

3D models are used also to investigate local dynamics and best fitting urban indicators for development [187], and find their use also in tourism for virtual tours [188].

The application of communicating urban information to citizens, impact of proposed projects, and to present the development of a city often results in the materialisation of the 3D city models as physical models [181, 189, 190]. Furthermore, within this use case, 3D city models may be used as a form of communication of cultural heritage [191].

2.3.3.9 Visibility analysis

3D city models are indispensable for many visibility analyses, such as determining the line of sight (LoS) between two points in an urban environment and for estimating the volume of sight [192–195] (Figure 2.4d). For example, they are used in estimating the visibility of a landmark [196, 197], assessing façade visibility for city marketing [198, 199], determining the optimal location for surveillance cameras [200–202], sensor coverage assessment [203], improving road safety [204], as-

sessing sniper hazards [205], and in real estate mass valuation in urban areas, based on the assumption that the view from an apartment is one of the factors driving its price [206–208]. Further applications involve predicting the visibility of GNSS satellites in the built environment and mitigating the multipath effect [209–219]. Such methods are valuable for enhancing map matching for navigation in urban canyons [220].

Visibility analyses with 3D city models are also used in studies on human perception of space [221, 222], and more advanced analysis resulting in distinguishing the view of water bodies, green spaces, factories, and roads [223, 224]. The information on the visibility from an apartment can also be used for taxation purposes [225].

Finally, 3D city models may be used for the estimation of the sky view factor (SVF)—the degree to which the sky is obscured by surrounding buildings [226]. Several researchers have demonstrated the use of 3D city models in estimating the SVF of their study areas for various purposes [227–234], e.g. for urban climate studies, and thermal comfort analyses.

2.3.3.10 Estimation of shadows cast by urban features

Estimating shadows cast by buildings is frequently used in urban planning [235], e.g. for assessing the impact of a planned building on its surrounding. Such analyses are legally required by some municipalities, such as The Hague in the Netherlands [236] and Mississauga in Canada [237].

This use case is also essential in the estimation of the solar potential of buildings, which may negatively affect the photovoltaic yield of a solar panel [90, 100, 238–243]. In this context, this use case is closely related to the previously mentioned one of estimating the insolation of buildings, and they are often used together.

Further applications include estimating the thermal comfort of buildings [230, 244], and the determination of solar envelopes [245, 246]. In the energy domain, Lange and Hehl-Lange [247] study shadow casting from a proposed wind turbine towards the existing surrounding residential buildings.

Finally, this use case has also an application in agriculture, such as estimating the predominantly shaded area of soil for calculating reduced growth in agricultural areas [248].

Using 3D city models to estimate shadows is one of the use cases that is employed in this thesis as an experiment. Hence this use case will be further described in Chapter 10.

2.3.3.11 Estimation of the propagation of noise in an urban environment

3D data is used to create models that answer how urban citizens are harmed by noise pollution [249–252] (Figure 2.4b), and how to mitigate it, i.e. where to place noise barriers [253, 254]. In Europe, the utilisation of 3D city models for this application surged after the Environmental noise Directive 2002/49/EC [255], which requires EU countries to produce strategic noise maps in order to inform the public about noise exposure and its effects [256, 257]. While 2D GIS is frequently used for this purpose [249], 3D city models provide an advantage over it, as due to refraction, sound levels may vary considerably at different elevations of the same planar coordinates [258]. Hence, since noise is simulated in 3D, it requires 3D geo-data as input.

Stoter et al. [259] produced a 3D noise map using 3D city models for obstacles in noise propagation, and Law et al. [260] for the visual impression of the results of the simulation. In this use case the semantics are not required, but may be helpful. For example, the knowledge of the type of object and the material of the walls may improve the results of the simulation of the propagation of the noise, and provide additional insight such as the number of apartments affected by excess noise [261, 262].

2.3.3.12 3D cadastre

Some governments have recently been focusing on developing property registration in 3D to provide insight into complex property situations, such as vertical ownerships in buildings and subsurface constructions (e.g. cables and pipelines, parking garage). 3D city models have been used to store and manage data about the physical counterparts of the legal objects and similar techniques have been used to collect, store and disseminate data about 3D legal objects as for 3D city models [263–273]. In the visualisation context, Shojaei et al. [274] and Pouliot et al. [275] investigate portrayal aspects in 3D cadastre.

2.3.3.13 Urban planning

3D geoinformation is ubiquitous in urban planning for various tasks, especially the visualisation of the urban environment [276–285]. Urban planning is a use case with blurry boundaries and a large number of actors [286]. However, there have been many documented specific purposes. For example, 3D geoinformation may be employed to facilitate park design [287], investigate urban objects that would interfere with the planning of a new metro line [288], temporal analysis of changes in the landscape [289], analysing the urban skyline [290, 291], and for traffic simulation [292].

2.3.3.14 Reconstruction of sunlight direction

Liu et al. [293] use 3D city models to determine the direction of the sunlight in photographs, which is useful for augmented reality, image processing, and object recognition.

2.3.3.15 Understanding synthetic aperture radar images

Several researchers in remote sensing have taken advantage of 3D city models for interpreting high-resolution synthetic aperture radar (SAR) images and predicting the reflectivity of future SAR image acquisitions with a ray tracing analysis [294–298]. The methods involve simulating the acquisition with virtual sensors and analysing SAR scattering effects with buildings of different configurations.

2.3.3.16 Facility management

Geoinformation is omnipresent in facility management. Recently, 3D city models have been employed for this purpose, e.g. in managing ports [299], airports [300], and utility networks [301–303].

2.3.3.17 Automatic scaffold assembly

Løvset et al. [304] presented a specialised use of 3D models of buildings for automatically designing an optimal scaffold assembly for it. Their method also takes into account the terrain around the building, and it complies with governmental rules and safety regulations.

2.3.3.18 Emergency response

3D geo-data can be used in disaster management and emergency response because they may provide valuable information such as the location of building entry points [305, 306]. In this context, 3D city models can be used to determine the best position for the deployment of the ladder trucks before the arrival of firefighters at the scene [307].

2.3.3.19 Lighting simulations

A seldom mentioned, but certainly distinct use case that we have encountered is the use of 3D city models in planning the lighting of landmarks [248]. Different lighting scenarios can be assessed without physical implementation and without visiting the sight, thus reducing associated costs.

2.3.3.20 Radio-wave propagation

Estimating the propagation of radio-waves for network planning is not a simple line of sight problem, since it involves concepts such as reflections and diffractions, which are more advanced that a straight line analysis [308]. This is an early GIS use case where DEMs have been used to extract a terrain profile between a transmitter and receiver, and then to apply propagation models [309].

This use case later evolved into 3D city models, and applications date back as far as 1994 in a research by Yang et al. [310] who uses ray-tracing on 3D buildings. Subsequent estimations of the propagation of radio-waves in an urban environment regularly include 3D city models [311–314]. Lee [315] demonstrate how to use 3D city models to predict Wi-Fi coverage.

2.3.3.21 Computational fluid dynamics

Computational fluid dynamics (CFD) and related analyses frequently take advantage of 3D city models [316–319]. They have been used for a wide variety of applications related to microclimate analyses: for estimating the wind flow and evaluating the wind comfort [320, 321], for understanding the urban thermal environment by estimating several environmental variables with CFD [322], estimating the physical effects of detonations and to determine the risks for structures and people [323], predicting the ground surface temperature [324], investigating the influence of air conditioning heat rejection management systems of residential buildings [325], and for the prediction of air quality [326].

2.3.3.22 Estimating the population in an area

Some application domains may require the number of inhabitants in a specific area, e.g. for assessing the population and number of affected buildings affected by the noise of a wind farm [248]. Since the size of a building and its type provide a clue as to the number of residents, using 3D geoinformation to estimate the population has been a topic of several research papers [79, 327–339].

The outcome of this use case can be used in multiple application domains. For example, for optimising the coverage of mobile radio signal coverage (i.e. to optimise the network to cover more people) [340], and emergency response for aid delivery and evacuation [341] (e.g. by estimating the population affected by a flood [342]).

This use case is focused upon in this thesis. It will be thoroughly described in Chapter 9, which presents a large-scale study of population estimation of 12 thousand census neighbourhoods in the Netherlands using 3D data in different LODs.

2.3.3.23 Routing

Routing is a traditional 2D use case that is gaining more importance in 3D city models since they may be used for outdoor navigation [343]. Slingsby and Raper [344] investigate pedestrian navigation enhanced by data not available in 2D, such as ramps and steps. This use case is considered as separate from the use case of visualisation for navigation purposes, as here the focus is on deriving the optimal route, rather than route portrayal.

3D city models containing indoor information can be used for route finding and accessibility [345–352], with specific applications such as evacuation [307, 353–357], navigating large train stations [358], determining indoor routes for the disabled [359], and locating the shortest path to the nearest automated external defibrillator [360]. Recent research efforts include the integration of indoor and outdoor routing for indoor emergency response facilitation [306].

2.3.3.24 Forecasting seismic damage

Christodoulou et al. [361] and Kemec et al. [362] use 3D city models to forecast and visualise damage to buildings from earthquakes, based on a framework for evaluating the seismic vulnerability. This use case is relevant for insurance, mitigation of earthquakes, and emergency response.

2.3.3.25 Flood simulations

Estimating the extent of floods has been a traditional topic in GIS, mainly with digital terrain models [363, 364]. However, models of the propagation and impact of flooding by an overflow of water from water bodies or heavy precipitation can be improved by using 3D city models [365] (see Figure 2.4a). Varduhn et al. [366], Amirebrahimi et al. [367], and Liu et al. [368] use 3D models to assess the flood risk and the potential damage at a micro-scale. This use case is important for insurances (risk management), evacuation, and utility management.

2.3.3.26 Change detection

Sharkawi and Abdul-Rahman [369], Pedrinis et al. [370], and Qin [371] use 3D city models for change detection in improving the quality of a city inventory. For example, it is possible to detect if an extension to a home has been built [372].

2.3.3.27 Volumetric-based density studies

A volumetric study is a research of the built-environment density, the volume and intensity of activities it generates, and its influence over an urban space [373, 374]. 3D city models are useful for volumetric analyses and they provide a substantial advantage over 2D data as they contain the height of buildings [373, 375]. The information on the volumetric density may also be used for modelling the dispersion of urban pollutants [374, 376, 377].

2.3.3.28 Forest management

Roßmann et al. [378] develop a forest management system that uses tree data at a comparable LOD to building models. The system may be used for several purposes: forest navigation, developing a sustainable management strategy for harvesting, and predicting tree growth.

Remote sensing has been extensively employed in the forestry sector, such as in estimating the volume of timber [379]. However, those applications are excluded from this analysis as they cannot be considered as 3D city models.

2.3.3.29 Archaeology

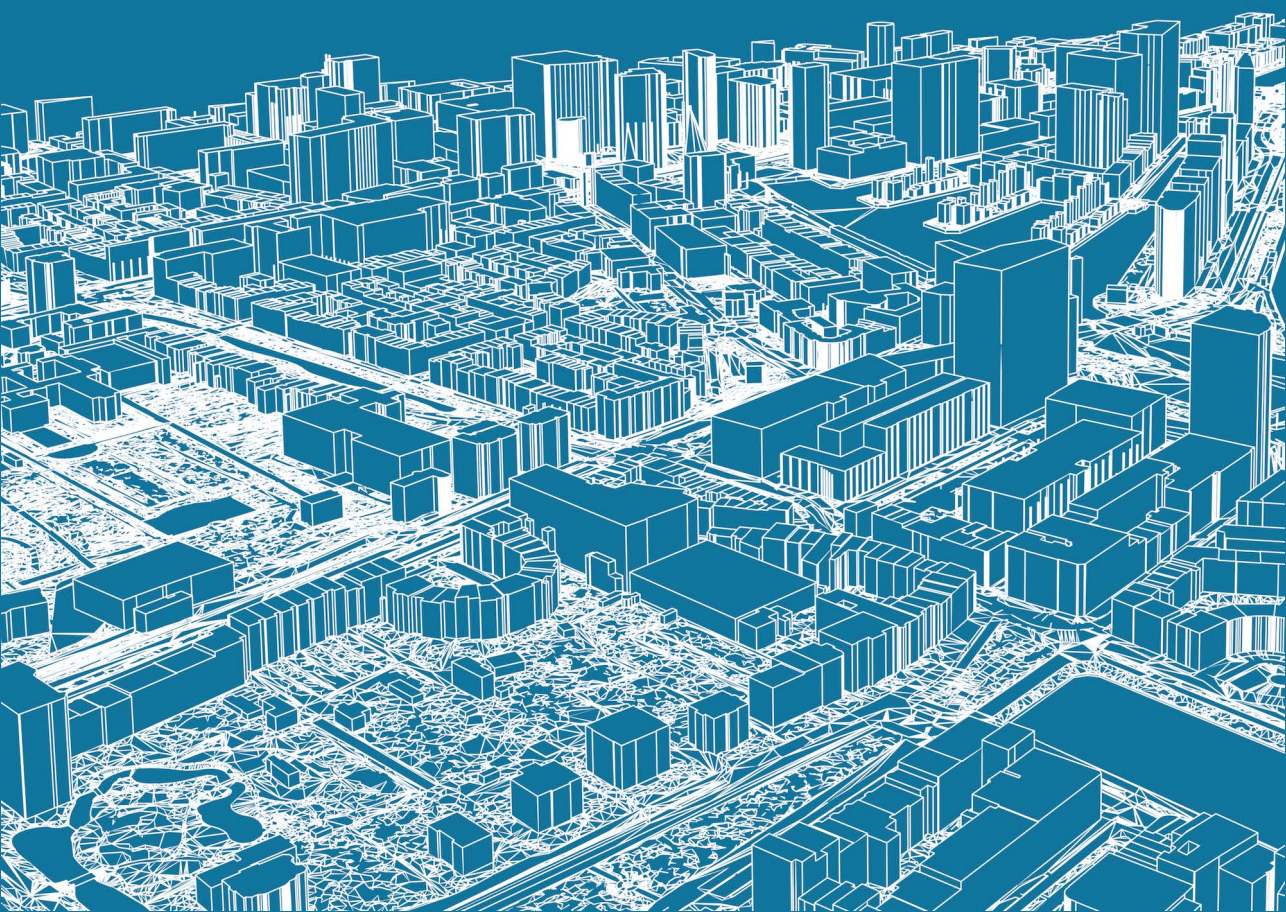
3D GIS is employed in archaeology, for example, for urban reconstruction of ancient cities, modelling of archeological 3D objects and their attributes, managing excavations, testing reconstruction hypotheses, and analysing development of sites over time [380–390].

2.4 LOD IN BUILDING INFORMATION MODELLING

Building Information Modelling (BIM) is a broad term involving processes, technologies, and policies to manage building design and construction [391].

BIM has its own LOD concept as well. However, it is of borderline relevance to LOD in GIS, as the 'D' in the LOD concept in BIM denotes *development*. The level of development provides the description of content and reliability of BIM models at various stages in the design and construction process [392].

There is also a certain degree of ambiguity when it comes to the LOD in BIM: some researchers state that BIM does not have an equivalent LOD as is in the GIS notion [393], while others borrow the concept from GIS [394]. Hence it can be concluded that LOD in BIM is subject to various interpretations [395]. For these reasons, the LOD concept in BIM is not considered thoroughly in this thesis. However, some of the concepts developed in this research may also be applied in BIM to some extent.



I

Framework and specification

Rotterdam, The Netherlands. Generated with 3dfier using BGT and AHN datasets.

CHAPTER 3

Formalisation of LOD

This chapter is based on my papers [396] and [397]:

> Biljecki F, Ledoux H, Stoter J, Zhao J (2014): Formalisation of the level of detail in 3D city modelling. *Computers, Environment and Urban Systems*, 48: 1–15. doi: 10.1016/j.compenvurbsys.2014.05.004

> Biljecki F, Zhao J, Stoter J, Ledoux H (2013): Revisiting the concept of level of detail in 3D city modelling. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, II-2/W1: 63–74. doi: 10.5194/isprsannals-II-2-W1-63-2013

Level of detail is an omnipresent concept in geographic information. It is used for the communication of how thoroughly real-world features have been acquired and modelled. The goal of this chapter is to present a fundamental overview of the concept of LOD, and to formalise it. The chapter starts by presenting mentions of the concept in literature highlighting its weaknesses and strengths, and it analyses existing concepts and arguments for a uniform LOD definition. The review of LOD concepts indicates that its definitions vary greatly between practitioners, standards and institutions. Furthermore, the LOD concept in 3D GIS is compared to the one in computer graphics. We define the LOD of a 3D city model as an indication of how thoroughly a 3D city model has been modelled, and the degree of its adherence to the corresponding subset of reality. We provide a formal and consistent framework to define discrete and continuous LODs, by determining six metrics that constitute it, and by discussing their quantification and their relations. The resulting LODs are discretisations of functions of metrics that can be specified in an acquisition-modelling specification form that we introduce. The advantages of this approach over existing paradigms are formalisation, consistency, continuity, and finer specification of LODs. One application of the framework is to provide a formal language to translate the existing LOD standards into a uniform specification. The chapter is concluded by reporting its mentions in literature since the publication of the original papers.

3.1 INTRODUCTION

The concept of level of detail (LOD) is essential in 3D city modelling. It is used to define a series of different representations of real world objects, and to suggest how thoroughly they have been acquired and modelled. Although the background and intention of the concept are intuitively recognised, in 3D city modelling the term LOD has been borrowed from 3D computer graphics and accepted without extensive discussion. In this chapter we argue that the term of LOD in 3D city modelling is currently incoherent, and that it is different from the one in computer graphics. It does not have a significant overlap other than the goal of the selection of a model sufficient for accomplishing a required task while balancing computational, economical, and cognitive limitations [398–400].

While the term is prevalent in several papers in the 3D GIS research community, it is influenced by computer graphics and its meaning often differs. For example, Meng and Forberg [401] define LOD as a uniform number of milestones along the scale space when taking the scale space as a linear continuum. For Glander and Döllner [402] it is a degree of generalisation. Forberg [403] expresses that it as a common way to enhance the performance of interactive visualisation of polyhedral data. According to Sester [404] and Goetz [405] LODs are multi-scale models for different applications, while for Granshaw [406] it is a hierarchical structure. Lemmens [407] equals it to the term of resolution and states that it is related to how much detail is present in the data and may refer to space, time, and semantics.

As explained in Section 3.2, the LOD in 3D city modelling in practice serves as a specification-related instruction for the acquisition, modelling, generalisation, and exchange of spatial data. This is in contrast with computer graphics where 3D models are simplified to their coarser counterparts in a dynamic process. Moreover, LODs of 3D city models do not differ only by the amount of geometric data, richness of *details*, and visual properties, but also they may differ by the semantics [408]. While researchers recognise that there are no universally agreed LODs for 3D buildings and other objects comparably to the 2D topographic maps that have official scale series [401], there is still not much work on the formalisation of LOD, i.e. a fundamental discussion that would standardise and unify the different approaches.

As introduced in Chapter 1, the CityGML 2.0 standard of the Open Geospatial Consortium [10] contains the *de facto* LOD concept of 3D city modelling (Figure 1.3). The specification of LOD for CityGML establishes quality classes for data acquisition, and the 3D model's LOD roughly reflects the model's complexity and accuracy [409, 410]. However, as it is the case with other standards, the LOD concept of CityGML has deficiencies. For example, the geometric content for each LOD is not strictly defined. For these reasons, discussions for its improvement are undergoing, in which I am taking part [411].

The goal of this work is to formalise the concept of LOD in 3D city modelling, and to provide a general and scientific framework for specifying LODs. Lacking a definition, specification, and a universal standard, the current LOD standards cannot be compared, translated, sorted, and evaluated. This leads to ambiguity in the communication of the acquisition-modelling properties of a 3D city model between users and producers.

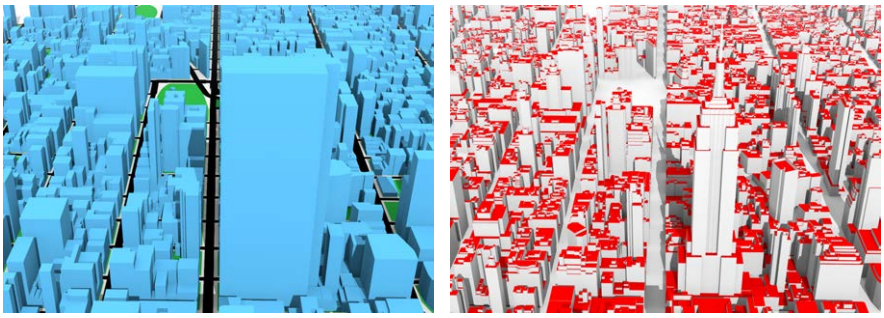
We define the LOD of a 3D city model as an indication of how thoroughly a 3D city model has been modelled and as the degree of its adherence to its corresponding subset of reality. In this chapter we decompose the LOD into six metrics that may be defined by continuous functions (Section 3.4), yielding a continuous LOD approach. In this view, the LODs are discretisations from a series of functions of such metrics (Section 3.5). We highlight that in such a case the traditional term LOD might be misleading. However, we do not propose linguistic modifications because we are aware that the current term is deeply ingrained in the GIS community, and that introducing a new term would be a futile task.

3.2 ANALYSIS OF EXISTING CONCEPTS AND THE NEED FOR AN LOD DEFINITION

We have evaluated different 3D city model representations and LOD concepts found in academia, standards, products, and guidelines (Section 3.2.1), and we have made an analysis and summarised their shortcomings (Section 3.2.2). We have found that these standards are essentially different not only by their specification, but also by their driving aspects (metrics), targeted usage, and arrangement of thematic classes and elements. In total, 26 level of detail paradigms comprising 79 mutually exclusive LODs have been evaluated.

3.2.1 Analysis of the concepts of LODs

The CityGML 2.0 standard of the OGC [10] defines five LODs. The concept is intended for several thematic classes of objects but it is primarily focused on buildings. The five described instances increase in their geometric and semantic complexity (Figure 1.3 in the Introduction of this thesis). LOD0 is a representation of footprints and optionally roof edge polygons marking the transition from 2D to 3D GIS. In LOD0 there are no volumetric representations. Subsequent LODs are improving in terms of the complexity of objects in the geometric and semantic sense. LOD1 is a coarse prismatic modelled usually obtained by extruding an LOD0 model. LOD2 is a model with a simplified roof shape, and where the object's parts can be modelled in multiple semantic classes (e.g. roof, wall). LOD3 is an architecturally detailed



(a) Block buildings with additional themes. **(b)** Fine buildings with no other themes.

Figure 3.1: LODs are not always ordinal. The dataset on the left has buildings modelled at a coarser LOD in comparison to the dataset on the right. However, it has additional thematic features (urban vegetation and roads). Which dataset has a higher LOD?[9]

model with windows and doors, being considerably more complex than its preceding counterpart. LOD4 completes an LOD3 by including indoor features [412], but otherwise it retains virtually the same properties as LOD3. This taxonomy has been developed in the German Special Interest Group 3D (SIG 3D) initiative [413–415], and has been further described in Gröger and Plümer [408]. The textures can be added to any LOD (i.e. the texture is not part of the LOD specification), and the standard includes other thematic classes, e.g. roads and vegetation which have their LOD description, but they are less prominent and their acquisition is not prescribed for each LOD (Figure 3.1).

The progress of the LODs is not consistent: the first LOD is 2.5D, while LOD1–3 improve the exterior geometry, and LOD4 adds one LOD of the interior that is indeterminate. Therefore, instead of five LODs, with respect the 3D city models and exterior geometry, there are three distinct LODs.

Since its inception, this LOD categorisation has became the primary LOD standard. In fact, CityGML partly owes its popularity to this simple and straightforward LOD concept. In the next chapter we argue that this concept has shortcomings and drawbacks, making it unsustainable as the number of producers, applications, and users grow with each having their own preferences.

National mapping agencies recently started adopting 3D city modelling standards. Examples include the Netherlands [417], whose standard is tied to CityGML, and China [418] developed from scratch not basing their model on any international

[9]Data courtesy of Chair of Geoinformatics, Technical University of Munchen [416]; and Department of Information Technology and Telecommunications (DoITT) of The City of New York.

standard. The Dutch standard extends CityGML classes and attributes, being more precise in the specifications. It also gives recommendations for textures. The Chinese standard contains four LODs, and defines which topographic objects should be modelled, and their thresholds (minimum size). The building LODs are defined by accuracy and basic description of the geometry. Also, different models have different requirements for the texture resolution.

There are also many other LOD concepts outside of national and local authorities. In academia, especially in the field of 3D generalisation, there are different specifications of the discrete LODs [401]. For example, Thiemann [419] defines three LODs for settlements and buildings: LOD1 contains aggregated settlement blocks with a uniform height, LOD2 blocks of the individual buildings without roof form, and LOD3 is LOD2 enhanced with a simplified roof form. Schilcher et al. [420] describes three LODs for individual buildings: LOD1 is a model popping up of the ground plan to a uniform height, LOD2 is LOD1 enhanced with a standard roof form, and LOD3 is an LOD2 enhanced with photorealistic textures and small surface features. These specification are clearly in conflict with the labels of CityGML, potentially resulting in misunderstandings and inconsistencies.

A few companies offer product portfolios of off-the-shelf 3D city models or for integration in a product (e.g. navigation software). Examples include Blom ASA [421], Vertex Modelling [422], NAVTEQ [423], CyberCity 3D [424], Sanborn [425], and TeleAtlas [426]. The companies offer 4 or 5 LODs, which are distinguished by the wealth of details and/or textures, and where landmarks have a special status. The semantics and the required accuracy are seldom specified. Furthermore, some companies offer additional adaptation and customisation of their models to fit the needs of their clients, making these LODs rather generic guidelines and frames of a final product later to be more precisely agreed by the two parties. However, most of the producers of 3D city models do not advertise their models in form of a series of LODs with a description and usage recommendation for each. Their internal standards rather serve as a general product frame, and may differ for each client or project. By direct inquiry, we have obtained the modelling specifications of a few other companies. They are essentially different but commonly contain a few LODs where the texture is not a part of an LOD specification.

The popular applications on smartphones for personal navigation, such as Google Maps and Apple Maps, recently started including 3D city models for their 3D visualisation mode. They contain up to two LODs distinguished by the complexity of the geometry and appearance.

We have studied a few tenders for the procurement of 3D city models, and publicly available models maintained by local authorities, such as the ones from the Glasgow City Council [427], Lusail in Qatar [428], and Australian cities: Wollongong [429], Perth [185], and Adelaide [186]. The tender specifications of 3D city models define

Table 3.1: Overview of the analysed LOD paradigms per group. All the series of LODs are driven by the granularity of the exterior geometry. Beyond this mutual aspect they differ in other drivers.

Group	No. of LODs	Main driving metric(s)
Standards	4–5	Ext. geometry, texture, semantics
Academia	3	Ext. geometry, texture
Off-the-shelf	4–5	Ext. geometry, texture, type of object
Mobile applications	1–2	Ext. geometry, texture
Internal guidelines	4	Ext. geometry
Municipalities	1–5	Ext. geometry, minimum accuracy
Special cases	1–5	Ext. geometry
All	1–5	At least ext. geometry

one LOD, and are often not detailed: they rather specify the minimum requirements for the deliverables, e.g. minimum accuracy, which features of a building should be included, and a set of library roofs to be used.

In this analysis, we have also encountered specific cases. These include the integration of the interior in a CityGML LOD2 model [131], mixing LOD for buildings of different types [158], and further, mixing CityGML LODs in the same object (different LODs for the wall and roofs) in an application for communicating future urban design with physical 3D models [189].

3.2.2 Analysis and critical overview of the current LODs

The overview of the described paradigms for each group is listed in Table 3.1.

From the paradigms briefly presented in the previous section, it is obvious that the main deficiency of the current LOD approaches is that it is not universally standardised what the LOD is (it is merely clear that it is a conception of the design quality of the 3D city model and the amount of data) and what it comprises. If a series of multiple discrete LODs is available the driving aspects differ between paradigms. For example, LODs in one standard are driven by the complexity of the geometry and the semantics (CityGML), while in the other by the characteristics of the texture (Blom3D). On the other hand, NAVTEQ's LODs are driven by the type of objects, i.e. the finest LOD only contains better representations of landmarks, while residential buildings remain the same as in the previous LOD. Thus, an LOD does not only define the wealth of the geometry, but a lot more: semantics, texture, interior, acqui-

sition techniques, and so on. The formalisation and quantification of such metrics is rarely discussed.

The paradigms all describe one or multiple discrete LODs, which are not linked and are not continuous. In case of multiple LODs, the improvement in the specifications of finer LODs is obvious. However, the functions of the progress are not specified, implying that the specifications are derived rather arbitrarily without connections between LODs, and the refinement of the range into intermediary LODs is not possible. For example, the quality of LOD2 is not 2 times that of LOD1, and it is not possible to derive an LOD1.7. In the paradigms defined by the producers, the LODs progress in the sense that are adapted according to the acquisition technique, e.g. features mandated in all but the finest LOD can be acquired with aerial photogrammetry. However, modelling features such as high-quality photorealistic textures, the geometry of awnings and openings rather requires a terrestrial acquisition technique.

Because there is no general definition for an LOD, LODs cannot be compared, translated, sorted, and evaluated. Furthermore, some paradigms define properties (semantics, texture, accuracy) that others do not mention. Since the LOD of a 3D city model is one of its most important properties, the industry and the research community suffers from such a deficiency, not being able to easily and efficiently communicate the acquisition requirements of a 3D city model in question [397].

Considering the thematic classes, the current standards are mostly focused on defining buildings, and with the exception of CityGML, they pay little attention to other classes of city objects. There are, however, independent studies about improving the LOD specification of other thematic classes, e.g. Chen [430] and Rafiee et al. [431] do it for trees, and Tamminga et al. [432] for roads.

From the semantic perspective, the paradigms do not offer a full semantic integration, and understate the importance of expressing the semantic requirements. For example, in CityGML a chimney is stored as a building installation and does not have its separate semantic class, which might be useful for some applications. In addition, if a chimney has its flue modelled, it is stored together as one building element in a class which is general and assigned to dissimilar types of building elements such as outer stairs. Moreover, while CityGML went furthest with storing semantics, there are no minimum requirements about semantics. For example, it is not required that windows in LOD3 are semantically labelled as windows, as they may remain unlabelled.

Another observation is that standards deal mostly with single objects without a specification for generalisation from a finer to a coarser LOD, and the modelling of aggregated multiple objects is not clear. Fan and Meng [433] notice that there is not a robust relationship between the elements in different LODs and reusability across multiple LODs are not possible.

Continuing the analysis, the current LODs in all paradigms do not appear to be specified exactly, and are rather considered as ranges since at least two different models of the same object can be of the same LOD. For example, Benner et al. [434] expose 12 different models that are LOD2 variants in CityGML. This causes ambiguity which is a twofold problem: due to incomplete LOD specifications it is not possible to precisely build a model of a specific LOD, and it is not possible to precisely evaluate existing models and determine their LOD. Beside stricter specifications, a solution would be to increase the number of discrete LODs to accommodate different variants, something which we accomplish in Chapter 4 presenting a refined LOD specification.

In brief, LOD, as one of the most important concepts in 3D city modelling, needs clarification, formalisation and improvement, and more research to address the detected issues.

Not much related work has been done on the formalisation of the LOD in the frame of 3D city modelling. Bandrova and Bonchev [435] suggest defining LOD by distinguishing it from scale and accuracy, and propose six discrete LODs for 3D maps. Löwner et al. [436] and Benner et al. [434] develop the separation of the LODs into semantic, exterior, interior, and appearance (sub)LODs for CityGML. Furthermore, they refine the LODs for the interior.

3.3 LOD IN COMPUTER GRAPHICS

3.3.1 Overview

The concept of level of detail in 3D city modelling has its roots in computer graphics (CG). The motivation is straightforward: it is inefficient to render a detailed 3D model that is far from the observer, because a highly detailed object will be shown only as a small object and the high detail will anyway not be noticed by the observer. Hence, in order to improve the efficiency of rendering, an object can be displayed at different LODs, depending on the context (Figure 3.2).

The LOD concept in computer graphics is important for GIS as it influenced the GIS and remote sensing community [438]. Several research groups operate on the intersection of the two fields, for example, to optimise the rendering of terrain and visualisation of 3D city models [9, 439–442].

Level of detail in computer graphics was pioneered by James Clark in his 1976 seminal paper [443] proposing using simpler counterparts of a geometry for 3D models that were less important, and since then it became one of the integral topics in computer graphics [400]. Its namesake in 3D city modelling is closely related, but with significant differences not making it appropriate to entirely consider it as its descendant. Some of the analogies from CG to 3D GIS have been considered in

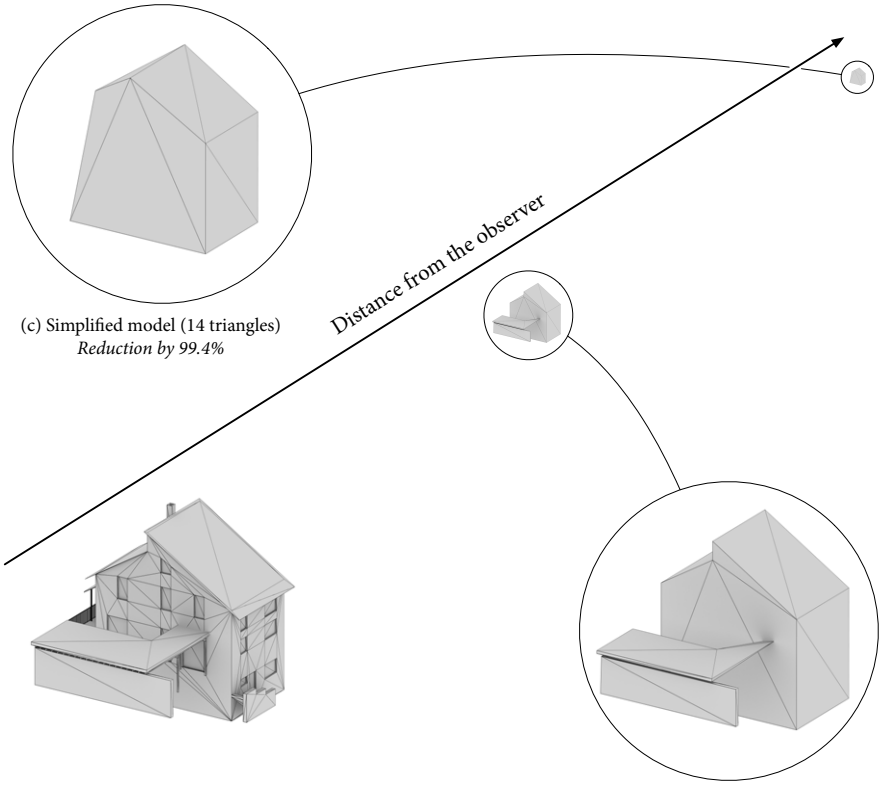


Figure 3.2: Level of detail in computer graphics. The original model (a) was simplified with quadric based simplification [437]. The face count was considerably reduced retaining a satisfying visual fidelity with respect to the context.[10]

this research, which will become relevant in the experimental chapters of the thesis (Part III).

The two principal LOD-related tasks in computer graphics are (1) *simplification*—reduction of detail of the displayed 3D object considering the limited computer's capabilities while not compromising the user experience [446, 447]; and (2) *appli-*

[10] The original model is courtesy of the Institute for Applied Computer Science, Karlsruhe Institute of Technology [444]. The models were simplified with Meshlab [445].

cation—determining the most appropriate representation of a 3D model to be visualised. Both are briefly overviewed in the continuation.

Simplification attempts to minimally change the shape and appearance of an object. Performing this task inherently entails errors and potentially the loss of visually pleasing detail that the original 3D model offers. Simplification seeks to find a compromise between the two, i.e. to efficiently reduce the complexity of the model while preserving *visual fidelity*. In the simplification, each LOD representation is quantitatively characterised by the metric of the number of polygons, and it presents an indication of *cost*. Since the number of polygons is a measure of the complexity of the model and it is directly related to the visualisation performance, it is a main metric for the complexity or the *detail* of the model. The fidelity could be considered as both a subjective and quantitative concept. The visual fidelity of the LOD representation may be assessed visually by humans or more commonly by *error metrics*. These are quantitative measures how similar one polygonal model is to another, which serve as a proxy for the visual fidelity. There is a variety of different error metrics, such as the Hausdorff distance[11] [449].

The main question of LOD application is to achieve another balance—that between fidelity and performance [450], selecting the most optimal LOD to be rendered. The LOD selection rationale relies on *factors* that form the *context* of the scene. These are numeric values that help in determining the most appropriate LOD. The most notable factors are culling, distance to the observer (Figure 3.2), importance of the object, and angle of view [451, 452]. Hence, the relationship in computer graphics can be reduced to a statement that LOD is defined by metrics such as the polygon count, which are the result of a function involving one or multiple factors: LOD = metrics = f(factor).

3.3.2 Comparison of LOD in 3D GIS to computer graphics

While the term LOD in 3D GIS originates from computer graphics and there is an overlap, there are some important differences that make the concept in GIS independent. A practical difference between the two fields is that CG often deals with too much detail that needs to be simplified, while researchers in GIS strive to maximise the acquired detail essentially working in the opposite direction.

Most importantly, the early purpose of the LOD in 3D GIS was to improve visualisation performance [453]. However, as demonstrated in the previous Section 2.3, visualisation is now only one of the applications of 3D city models. The key dissimilarities between the two fields are given in the following paragraphs.

[11] The Hausdorff distance [448] is described further in Section 10.3.4 where it is used as an error metric to find differences in different outputs of a spatial analysis.

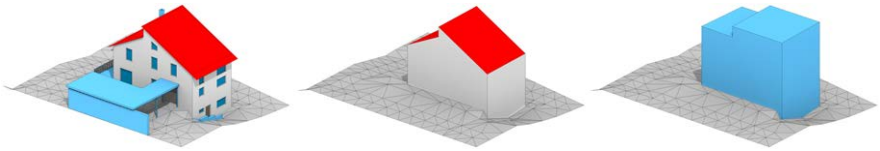


Figure 3.3: LODs in 3D city modelling as opposed to the ones in computer graphics. This example illustrates a few LODs from the same model as in Figure 3.2. Generalisation strives to preserve the structure, as opposed to the approach of simplification in computer graphics that is focused on conserving visual fidelity.[12]

3.3.2.1 Simplification

Geographical data may also be simplified, i.e. generalised. The primary motivation behind simplification in computer graphics is improving rendering performance, and uses outside rendering are seldom [454]. Hence, the simplification has intrinsically different aims—it aims at preserving the structure of features, e.g. rectangular footprints and walls of buildings [455, 456], employing cartographic approaches [457]. An example is given in Figure 3.3. The original model shown in the computer graphics example in Figure 3.2 is simplified according to the simplification nature in GIS, preserving the structure and topology of the building, and not generalising the terrain. Furthermore, the different LODs contain different semantic granularity. Computer graphics does not consider semantics as much as GIS, which is another difference that is relevant to note.

3.3.2.2 Measuring the detail

The main measure of complexity and differentiator between different LODs in computer graphics is the amount of geometry, e.g. the number of triangles. Expressing detail in GIS is not so straightforward. As a consequence, it is not straightforward to attach a numerical and rational value to each LOD. We clarify this difference in Figure 3.4 with two arguments. The illustration shows three 3D models of the same building: the left and centre model are both considered as LOD2[13] according to CityGML, while the one on the right is LOD1. First, this example exposes the fact that the two (different) models on the left have the same polygon count, the foremost metric in computer graphics to distinguish two representations. Second, besides the aforementioned LOD2 models, the model on the right is a geometry extruded from a fine footprint of the same building. This LOD1 model has a higher

[12] Data courtesy of the Institute for Applied Computer Science, Karlsruhe Institute of Technology [444].
[13] This ambiguity will be addressed in Chapter 4 in which such variants will be regarded in a refined and improved specification of CityGML LODs.

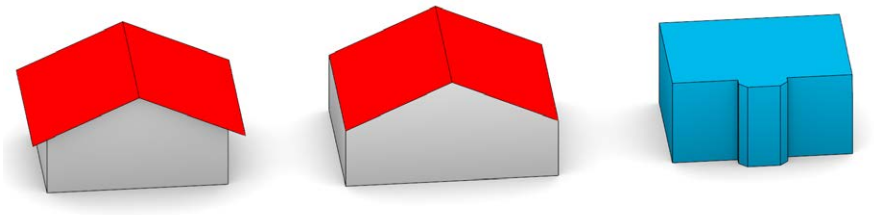


Figure 3.4: Two variations of LOD2 and an LOD1 model exposing the shortcomings of the CityGML LOD concept. The figure also demonstrates why the computer graphics principles cannot be fully applied to GIS and 3D city modelling.

face count than the LOD2 models. While the number of the primitives generally gives a good indication about the geometric complexity of a 3D city model, it cannot be considered as an unambiguous differentiator as it is the case in computer graphics (the only exception to this in GIS and in 3D city modelling is terrain because of its usually triangular representation: lower LOD means less triangles [458, 459]). Consequently, unlike in computer graphics, the LODs in 3D GIS cannot be ordered: the LOD1 model, intrinsically considered inferior to an LOD2, may be accounted as more valuable than an LOD2 for certain scenarios when a finer footprint is more useful than the acquired roof shape. An example of such a use case is the computation of the net internal area of a building, useful for energy estimations, real estate valuation, and population counts [78, 121, 129, 329]. Hence it does not strictly hold that LOD$(i+1)$ > LOD i, i.e. the LODs are rather nominal, instead their ordinality rather depends on the use case and other aspects [396] (consider again the case in Figure 3.1).

Figure 3.5 illustrates another example with real-world datasets. Some datasets at a coarse LOD contain more geometry per building than datasets at finer LOD. Admittedly, this may, to some extent, be caused by different architecture and modelling rules, where buildings with complex design require more geometry to be represented. However, the samples are sufficiently large to derive a conclusion that the number of surfaces cannot be used to define an LOD.

3.3.2.3 Dynamic nature

The LOD in GIS is determined before any acquisition has taken place: by designing the specifications for acquisition. LOD in CG is more dynamic: multiple LODs are generated on the fly depending on the context. As a consequence, in GIS practitioners use data in one LOD and multi-LOD datasets in GIS are currently not com-

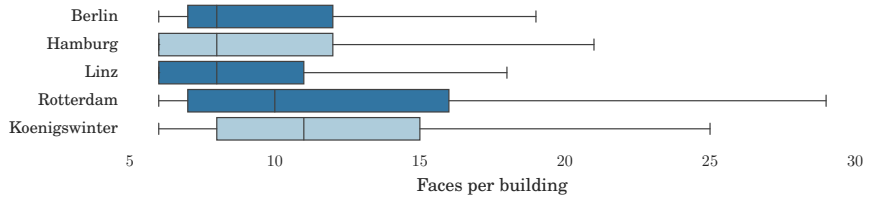


Figure 3.5: The count of faces cannot be used to define the LOD. These box plots indicate the average number of faces uses to model a building in different datasets. The colours represent different LODs: LOD1 (lighter) and LOD2 (darker hue). As opposed to computer graphics, there is a large variation in the number or faces within the same LOD, as an LOD1 can be more detailed than LOD2 in certain cases.[14]

mon, and LOD management in GIS is practically non-existing: there are no LODs to switch.

3.3.2.4 Context, cost, and factors

While the fundamental logic of the LOD concepts in GIS and CG are similar, discussing terms such as context, cost, and factors is difficult because data in GIS is used for different purposes other than visualisation. First, each spatial analysis may have different requirements, resulting in different factors. Second, cost in GIS has a different meaning than in CG. In CG, cost refers to the computational load required to visualise a dataset (e.g. time in milliseconds). Furthermore, such a value is not so much applicable in GIS because a spatial analysis is usually ran only once and the results are stored, i.e. the models are not repeatedly used (rendered) as in CG.

In GIS the main cost-related concern is the cost of acquisition—not necessarily only the price of acquisition but also technologies involved and the time required. This is hard to take into account, because prices change through time, they may vary from country to country, and they are not easy to obtain or to find a proxy for it. For these reasons, the cost of procuring a dataset is not taken into account in this thesis. However, a parallel to cost that can be drawn is the error of a spatial analysis. It is believed that coarser LOD result in larger errors in the result of a spatial analysis. This widely accepted assumption is challenged in Part III of this thesis, and the thesis discusses balancing the LOD and inevitable errors.

[14] Data courtesy of: Berlin Senate Department for Economics, Technology and Research; City of Hamburg; City of Linz; City of Rotterdam; and University of Bonn.

3.3.2.5 Heterogeneity

In GIS many 3D datasets have been produced from multiple sources, hence their LOD is not consistent. For example, a model derived with airborne laser scanning will likely have roofs mapped in more detail than walls. A similar case is a dataset derived from a mobile laser scanning unit, resulting only in data of features visible from a road. Furthermore, it is common to map more important features at finer LOD (e.g. see the landmark in Figure 1.1) or to provide context (e.g. compare in the same figure the building with indoor data with the neighbouring block models). For these reasons, the LOD in GIS may be considered heterogeneous in comparison to the one in CG (see Figure 3.1).

Additional reasons for having a mixed LOD in the dataset may be: privacy and security (some features may deliberately be mapped in lesser detail to conceal sensitive detail), failure of automatic reconstruction methods which may reconstruct an LOD1 when LOD2 fails, and the inability to repair a building (an LOD1 is generated to replace an invalid LOD2).

3.4 CONSIDERATIONS FOR THE IMPROVEMENT OF THE DEFINITION OF THE LOD CONCEPT

3.4.1 Composition of the design characteristics of a 3D city model (metrics)

By analysing the standards presented in Section 3.2, by discussing with other researchers and practitioners, and based on own experience, we propose a list of 6 metrics (quantifiable *ingredients*) that constitute an LOD of a 3D city model. These metrics may be applied separately to all data in a spatial extent (or dataset), to a class of city objects (thematic class), and to their elements (features which cannot be semantically further decomposed). The spatio-semantic design characteristics of any model could be specified with these six metrics.

1. Presence of city objects and elements The selection of objects and elements that have to be modelled is an obvious metric. However, this concept is hampered by non-standardised nomenclature and semantics, unclear hierarchies of objects and elements, their aggregation relations, and modelling rules (e.g. should a roof be modelled with overhangs or not). Here we propose how to define such a metric and we use common-sense nomenclature.

The presence of an object or element is a binary property (something should be acquired or not), and it can be easily expressed as a list of real-world features that have to be geometrically mapped. Figure 3.6 illustrates an example of three different LODs (A, B, and C) for two city objects: building and vegetation. The coarsest LOD

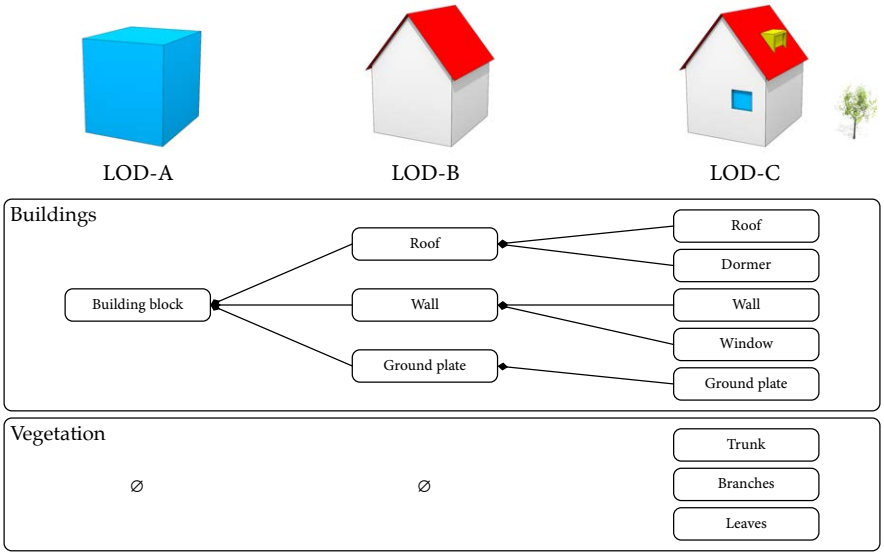


Figure 3.6: A simplified explanation of the concept of presence of city objects and their elements in three different LODs of a 3D city model. The aggregation relationship between elements is expressed in the diagram below the graphical representations. Notice that the element Building block is not found in finer LODs (it is replaced), contrary to the case of walls and roof structure.

(LOD-A) requires only buildings to be modelled and only as a block model. So here only one element of the building has to be acquired, in this case defined as a building block. Vegetation and other thematic classes of city objects are not required to be included in this LOD. The LOD-B improves the LOD-A by replacing the block model with the walls, ground plate, and roof structure. The finest LOD in this series, LOD-C, adds the thematic class vegetation, and to buildings it adds elements such as windows and significant roof structures such as dormers. This LOD also contains the walls and the roof structure as in the coarser counterpart.

This metric also covers the aggregation of objects and elements, and we add a structure defining the relations of objects and elements in multi-LOD cases. However, exact aggregation rules are out of scope of this work.

For simplicity, in this example the elements and objects contained in coarser LODs remain unchanged in finer LODs.

Figure 3.7: Two buildings modelled in different feature complexity. The wall surface (white) and top surface (red) at the finer LOD (right) are modelled more precisely due to the finer feature complexity requirement.

2. Feature complexity Besides listing objects and/or elements that have to be modelled, one important description of an LOD is the complexity or fineness of their geometry with respect to the real-world. This metric defines the geometrical correspondence of the model to the reality, which most of people perceive as resolution or (the level of) detail. A straightforward way to realise it is as the minimum length of linear elements in the real-world that will be taken into account in the model.

Continuing on top the previous metric presence, a city object and its elements may be present in two LODs. However, with a different feature complexity. Thus one model may geometrically be a better approximation of a feature of the real-world. Figure 3.7 illustrates an example of a house modelled in two different feature complexities.

This concept is linked to shape complexity, and it can be expressed in various ways [460]. For example, it can also be quantified with a fractal dimension [461]. In this chapter we choose a metric magnitude, i.e. modelling all the details of an element that are bigger than a certain size, such as in the Figure 3.7 where the 'dents' and 'irregularities' of the wall surface are in reality bigger than 30 cm.

Feature complexity is related, but should not be confused with the complexity of the model. For example, a simple garage with flat elements may be modelled with a fine feature complexity, and while the resulting model would closely correspond to the real-world, it may still look simple and coarse.

Furthermore, in addition to elements, this metric may be applied to city objects. In such a case it specifies the minimum size that an object has to have in order to be acquired, a metric currently included in CityGML.

For example, for coarse LODs in some applications such as noise pollution estimation, it is not relevant to acquire smaller buildings such as garages, or regarding object elements, small balconies.

3. Dimensionality While 3D city models imply the use of 3D volumetric geometries, primitives of lower dimensions can be used as well. Examples are 2D footprints of buildings which are used in the LOD0 of CityGML, or point representations [462]. Roads may be represented as (1D) lines, and trees as (0D) points, which may be satisfying for a number of applications (e.g. Figure 1.1 shows an example of a 3D dataset in which roads are modelled as 1D lines). We consider dimensionality of the geometry of an object different from the concept of feature complexity, and as a separate metric. The dimensionality can be defined for elements as well. For example, a 3D building may contain windows modelled as polygons, and chimneys as points on the roof surface or vertical lines.

4. Appearance (texture) Textures are important in visual applications, and particular use cases in which visual appearance is appreciated, e.g. archaeology [463]. Furthermore, the appearance may complement the presence of the elements which are not geometrically and semantically acquired. For example, while windows of a building may not be geometrically present in a model which is textured, they are available for visual inspection and approximate measurements.

5. Spatio-semantic coherence The semantic richness of a 3D city model has been an important concept of CityGML from the very beginning. Semantic information is critical for a number of applications [464].

The spatio-semantic coherence describes the granularity of the semantics in a model and its correspondence to the geometry, since it is possible to have different LODs of the semantics for the same geometry (and other metrics). For example, while a tree may have its canopy and branches modelled, in one case they may be assigned the same semantics (e.g. tree), while in another each element could have its distinct semantics (1:1 mapping).

Stadler and Kolbe [465] recognise six general cases of spatio-semantic coherence. In the finest case, the semantic model and the geometry are given as a complex aggregation where both components correlate on the same levels of the hierarchy. Following their work, it is possible to define different levels of spatio-semantic coherence (e.g. a level where a dormer and roof surface are distinct classes, and another case where they are attached the same class).

6. Attribute data Each component of a 3D model (a dataset, a city object, and their elements) can be assigned one or more attributes, e.g. year of construction, and use of the object, access to the roof, roof material, address, and type of a road [466]. However, the list cannot be compiled at a generic level because it is dependent on the application.

3.4.2 Points of discussion related to the metrics

Coverage and applicability of metrics A metric may be expressed from a discrete or continuous function. For example, the presence of elements is discrete (cf. Figure 3.6), while the resolution of the texture may be continuous (e.g. a value ranging from 50 to 10 cm/px). Table 3.2 outlines such a property for each metric, and the applicability of metrics per dataset or spatial extent (multiple objects covering all thematic classes), city object, and their elements. For example, presence is a metric that can be defined for both an object (e.g. bridge) and its elements (e.g. cables).

Table 3.2: Properties of the metrics (discrete and continuous). The overview includes the applicability of the definition of the values of metrics per dataset (or spatial extent), classes of city objects, and their elements.

	Metric	Property	Extent	City object	Element
1.	Presence	Discrete		•	•
2.	Feature complexity	Continuous	•	•	•
3.	Dimensionality	Discrete			•
4.	Appearance	Continuous	•	•	•
5.	Semantics	Discrete	•	•	•
6.	Attributes	Discrete	•	•	•

Quality vs LOD It is important to note that the two considered qualities of geographical data—LOD and accuracy (one aspect of quality)—are unfortunately perennially misapprehended as synonyms. One may often hear that a model is described as of a 'high quality model', which is colloquial and it has no underpinning according to international GIS quality standards. While there is an association between the two (representations at finer scales tend to be of higher quality [467]), these are two independent concepts [468]. That means that two datasets of the same area may have the same LOD but different spatial accuracy, e.g. an LOD1 may have a positional error of 0.3 m, while another one an error of 0.6 m. Furthermore, in some cases features may be considered having a fine LOD, however, their geometry and other data may considerably deviate from the real-world (i.e. it can have a low accuracy). Finally, some datasets are design (see the acquisition taxonomy in Figure 2.3), and are not representations of a real-world setting, but they still have an LOD. Hence, it is meaningless to discuss quality characteristics such as positional accuracy.

This distinction has been only briefly discussed in related work. For example,

Oude Elberink and Vosselman [469] highlight that specifying the LOD does not mean that the geometric accuracy of the model has been determined. In 2D, Daly [62] point out that having a finer LOD of GIS data does not necessarily increase its quality, in fact it may degrade it: interpolation from a coarser LOD to produce data at finer LOD does not warrant a more accurate dataset, on the contrary.

The separation between quality and LOD is in line with the framework presented in this chapter, i.e. LOD is seen as a product specification of a 3D model, and it is different from the data quality concepts such as positional accuracy or completeness. However, on top of the LOD specification, practitioners may define recommendations or requirements of spatial accuracy for an LOD.

The thesis revisits this topic in Chapter 12 where datasets of different combinations of LOD and accuracy levels are used in spatial analyses.

Scale vs LOD Similarly with quality, scale is sometimes used interchangeably with the LOD [470, 471]. For example, Luo et al. [472] state that they generate virtual 3D city models 'at a scale of 1:1000'.

To add to the confusion, scale has multiple meanings in GIS [8, 473]. In cartography, it refers to the relationship between distance on the map and the corresponding distance in reality. A certain map scale implies a particular LOD that ought to be mapped depending on the purpose of the map [474]. In addition, it refers to the extent of a mapped area, and the resolution of the data [475, 476]. The latter is closely associated with the LOD [477].

With the transition from paper maps the term scale is losing its meaning [478, 479]. Because of the digital environment in 3D city modelling[15], scale does not have such an explicit meaning. Therefore, we conclude that while there is an association between the terms scale and LOD, using scale in 3D city modelling should be in general avoided.

Partial specifications Notice that when defining an LOD it is not required to specify all the metrics. For example, in the production of 3D city models represented as digital surface models (DSM), it is not possible to specify the presence of objects and elements to be collected. Producers rather specify the feature complexity as a loose equivalent to resolution, and do not have control over the collection of different types of objects and elements (e.g. vehicles may also be present in the model).

[15] There are physical representations of 3D models to which the term scale, as in the cartographic context, may apply. For example, in some applications such as wind flow simulations [480] and sound propagation [481, 482] physical 3D models may be used. In such a case of a tangible representation, the term scale may be used to indicate the ratio of the size of the model.

Generic (library) features In order to improve the efficiency of modelling, storage, query and visualisation, it is not uncommon to just acquire a footprint or a point representing a feature, and then to include a finer model of the model from a library for visual representation. Coors [141, 483] distinguish this as a query data model (feature geometry) and a presentation model (view), each having its LOD, the latter possibly having multiple LODs depending on the visualisation requirements. This is common in tender documents where it can be specified to use features from a library (e.g. for common types of roofs, cf. Kada [484]) instead of modelling each feature. Another example is the acquisition of roads as 1D lines, but their portrayal as a 2D surface feature (i.e. extruding the line according to a width based on the attribute of the number of lanes).

The framework presented in this chapter takes into account such a concept. For each representation its LOD can be specified. However, the correspondence of the models and their real-world situation, is a matter of spatial data quality elements.

3.4.3 Interior of the model

CityGML 2.0 supports only one LOD for the interior which is not particularly defined and requires a fine exterior geometry and semantics (LOD3). Researchers identify this problem and propose multiple LODs for the interior [485, 486].

The interior is not listed as a metric because we consider that the interior geometry is a different concept that should be separated from the exterior geometry, and all the presented metrics could be applied to it. For example, depending on the application, we may consider a model with fine exterior geometry and texture, but with no interior, still as a high-LOD model.

With this view we first decompose a city object to the exterior and interior, and use the metrics separately to each. We consider the interior of the model a separate concept from the exterior for the following reasons: (1) the applications using only exterior geometry are by far outnumbering the indoor applications, which are fundamentally different and have different users; (2) the vast majority of 3D city models do not contain any interior [487]; (3) with the recent introduction of indoor standards the interior geometry developments appear to be dissociated from the 'orthodox' 3D city modelling field [488]; and (4) the acquisition and modelling techniques for the interior are different, e.g. see the work of Sternberg et al. [352].

3.5 PROPOSAL OF A FORMALISED LOD FRAMEWORK

Based on the analysis presented earlier in this chapter, it is obvious that there are different views on the LOD concept. Hence coming up with a definition is difficult and may be subjective. We define the LOD of a 3D city model as an indication of

the content of the data and a degree of its spatio-semantic adherence to its corresponding subset of reality. The values through which the adherence is expressed are the metrics introduced in the Section 3.4.1. When an LOD is decomposed into these quantifiable components, a straightforward comparison of two or more LODs is possible. Each such different combination of the values of metrics is a different LOD, and a small difference results in a change in the LOD, making this framework continuous.

Here we introduce two concepts: series of LODs and continuity. The series of LODs are a 'sensible' combination of metrics such that their progress is overall. Viewing this concept in a continuous world, LODs are discretisations of progressive and monotonic functions of the values of metrics, i.e. each discrete LOD i is a collection of the values of n metrics M, which are the result of the discretisation of functions of metrics:

$$\text{LOD } i = \begin{bmatrix} \mathrm{M}_1 & \mathrm{M}_2 & \dots & \mathrm{M}_n \end{bmatrix} = \begin{bmatrix} f_1(i) & f_2(i) & \dots & f_n(i) \end{bmatrix}, i \in \mathbb{R}, n \in \mathbb{Z}.$$

In the following two sections we demonstrate how to specify the functions of the metrics, how to define a series of LODs, and give an example of the realisation of the framework. We also introduce a specification format for discrete LODs.

3.5.1 Specifying the metrics

3.5.1.1 UML model of the specification

A UML diagram (Figure 3.8), based on the discussion in Sections 3.4.1 and 3.4.3 and on Table 3.2, is created to support the specification of the metrics for an LOD. Each 3D city model consists of a selection of one or multiple city object types (thematic classes) in a spatial extent, with general properties (`GeneralMetric`) that apply to the hierarchy, such as the feature complexity for all types of city objects and semantic requirements. Each city object that has to be modelled can be defined by metrics such as attributes and again feature complexity. As argued in Section 3.4.3, an object should be separated into exterior and interior, however, the latter is constrained with the exterior. The focus of the specification is on objects' elements (features that make up a city object), which acquisition–modelling specification is defined by the presented 6 metrics.

The city objects can be specified further than the typical classes in CityGML. For example, instead of defining a class for buildings, two classes such as residential buildings and landmarks may be defined, and different acquisition–modelling specifications can be applied to each. This is in line with 3D city models that are used for navigation purposes. Equally important, it is possible to define city objects based on

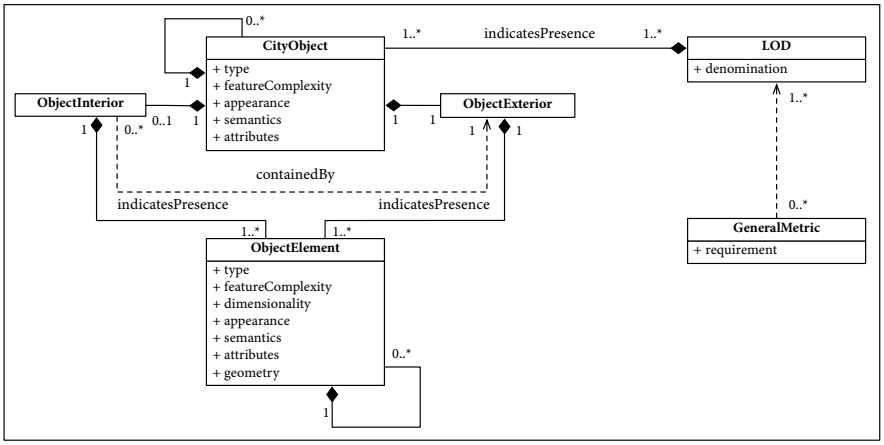


Figure 3.8: The UML diagram of our formalised LOD framework.

other criteria such as their size, e.g. trees taller than 2 meters can be defined as a city object class of tall trees that should be modelled with higher specification requirements than shorter trees.

Aggregation is supported within this framework as well. A `CityObject` may be aggregated to another `CityObject`. The same applies to elements. Furthermore, the framework permits us to mix LODs (e.g. roof and walls at different LODs).

3.5.1.2 Specifying the functions of the metrics

Some of the the presented metrics (see Table 3.2) are quantifiable and of continuous nature, hence their values may be expressed in functions.

In Figure 3.9 we give an example of the function of the metric feature complexity for a dataset (general metric), that applies to all objects and elements, unless specified otherwise. For all metrics, first an arbitrary interval range of LODs (e.g. from 0 to 10) is assigned. In the construction of the functions we must determine a range of values and their function, i.e. in this case the feature complexity ranges from 6 m to 0.1 m in an exponential function. In this way, it is possible to discretise the function and obtain the value of the feature complexity at one or more specific points in order to define a series of LODs, e.g. LOD 4.88 specifies the feature complexity at 0.81 m.

The discrete metrics, especially the presence of objects and elements require a different approach. Our goal is to define in a integrated approach: (1) how to denote the presence of objects and elements in a series of LODs; (2) how to specify the depth

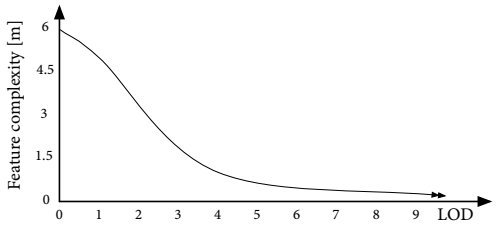


Figure 3.9: Example of the function of the metric feature complexity. This is one of the functions that define the series of LODs.

of spatio-semantic coherence; and (3) how to define the aggregation of objects and elements; and finally (4) how to specify the required attributes. Similarly to the continuous functions, these functions specify the values of the metrics at discrete steps.

In order to express the metric presence, first we have to define an inventory of city objects and their elements that are relevant for 3D city modelling. While CityGML is a good start, such an inventory does not exist. For this reason we have analysed the standard ISO 6707-1:2004 Building and civil engineering – Vocabulary – Part 1: General terms [489], which is extensive in the inventory of elements that compose a building. With this standard we have virtually all elements that are relevant for 3D city modelling in a well defined list, and beyond that since the standard is quite detailed. This example shows how to base the list and semantics of elements, in this case for a building. Afterwards, it is required to select the scope of the selection of the relevant elements since 3D city modelling has no clear boundaries. For example, in a coarse LOD buildings could be aggregated in building blocks (of a few adjacent buildings), and further in a neighbourhood block. For many stakeholders and applications this might be out of their interest. On the other side of the LOD axis, each element could be decomposed into very detailed elements. As an example, a wall could be decomposed into bricks and mortar, and each could be individually modelled. This may also be out of the scope of 3D city modelling.

After setting the scope, based on the described with a few additions we have made an inventory of elements of a building, and their aggregation hierarchy across a series of LODs from 0 to 9. An excerpt example of the elements per discrete LODs are illustrated in Figure 3.10. Since a building can be composed of more elements than it can be fit in a simple diagram, this example is limited. The designation in the brackets refers to the definition in ISO 6707-1. Such an inventory may also apply to the LODs typical in BIM, and small-scale objects, which also means that this LOD

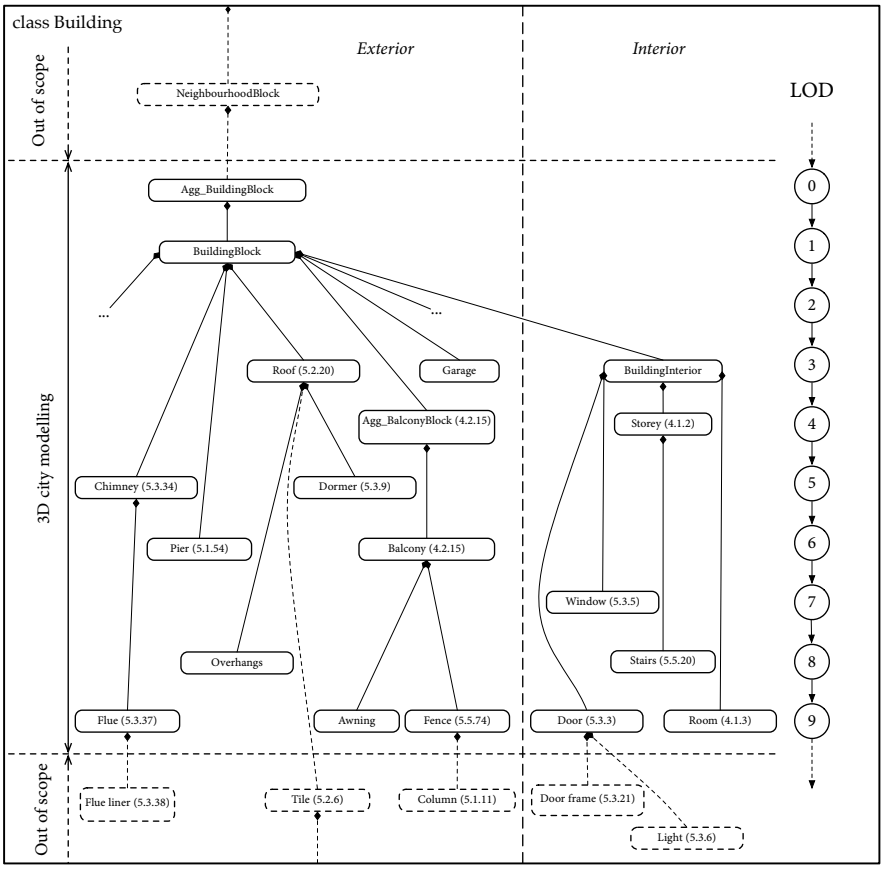


Figure 3.10: An example of the object building and its elements in the scope of 3D city modelling. The diagram is designed in a hierarchical form in order to support the metrics of presence (with aggregation), semantics, and attributes for a series of LODs. The right axis shows exemplary LOD values. Referenced names are from ISO 6707-1.

concept, if viewed broadly, can be extended to both areas and their ranges of LODs.

Second, the metric spatio-semantic coherence fits such a hierarchy, but is defined separately from the metric presence. The specification can define the metric equal or lower than the presence. For example, while awning and fence of a balcony may be modelled (in this particular example LOD9 with respect to the presence) the se-

mantics of both elements may be balcony. That is, the presence of elements has the value 9, but the semantics the value 6. For a full spatio-semantic coherence, their LODs should correspond.

Third, the aggregation is specified by introducing new city objects and elements such as aggregated balconies, i.e. balconies that are close enough to each other get aggregated in one object. This may be done by creating new objects and elements in the hierarchy. For example, the hierarchy specifies that the balconies are required to be modelled in LOD6 and finer, but LOD4 specifies that it enables the aggregation of balconies. Similar is with the buildings in LOD1 and aggregated buildings in LOD0. The aggregation rules can be attached to their relations (e.g. aggregate balconies if their distance is less than 3 meters), and are out of scope of this chapter.

Finally, the required attributes may be defined for each LOD. For example, one LOD may not require any attribute of buildings, but a finer LOD it may require an attribute such as year of construction. This can be realised by specifying such a requirement in the hierarchy at a respective discrete point. This is not given in the figure due to the limited space, but it is shown later in the example of a specification of a discrete LOD (Table 3.3).

3.5.2 Defining a consistent series of levels of detail

Sensible combinations of different and improving metrics values are what we consider a consistent series of LODs. Figure 3.11 gives a clarification where the left graph decomposes the CityGML's five LODs into the two leading metrics: exterior geometry and interior geometry. While in our framework we do not consider these two as separate metrics, in CityGML they can be seen as metrics (Section 3.2.1). The progress of the LODs is directed towards each metric at a time, meaning that the progress is not overall, but also that their relationship does not exist. Other standards are similar. The right graph illustrates our approach, simplified with two arbitrary metrics (e.g. texture and feature complexity). Any combination of the values of these two metrics can be a certain LOD, and any combination of the functions of metrics could represent a series of LODs. While this enables an infinite number of permutations and supports any combination of metrics, it would yield numerous senseless combinations (e.g. acquiring a high texture elements with a low feature complexity, and vice-versa), and a series with no clear logic (e.g. a finer LOD having a lower feature complexity than a lower LOD).

However, we consider that only a subset of such combinations may be a consistent series of LODs, where the progress of each metric is well defined in a monotonic and progressive function, and comparable (two examples are given here). Any combination lying on such a line is considered as a discrete LOD of a particular (continuous) series, and in this figure two examples of functions are given.

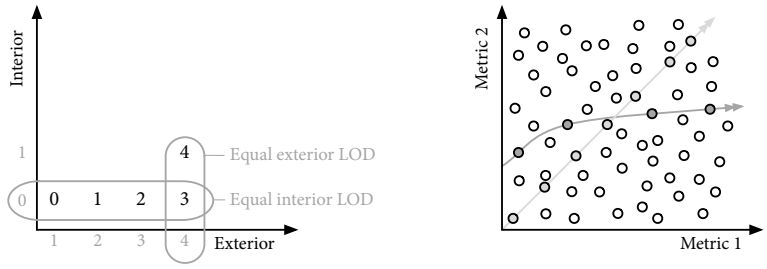


Figure 3.11: Simplified example of the construction of an LOD function from metrics. The left graph shows the theory on CityGML, an approach applicable to most of current standards, while the right graph suggests our view of the consistent series of LODs.

3.5.3 Specifying a discrete LOD of a 3D city model

In order to express the specification of a discrete LOD, a straightforward specification format (for acquisition and modelling guidelines) should be available. However, we have realised that in 3D city modelling such a specification does not exist, therefore we propose our own, in tabular form. An example of the specification of a discrete LOD (e.g. named LOD i) is given in Table 3.3.

First of all, the general metrics that apply to all city objects and elements (unless specified otherwise for each) are given in the beginning: the feature complexity is 0.4 m. The semantics should be fully specified (each object and element that cannot be semantically further decomposed should have its semantic class).

The second part contains the list of city objects and their elements that should be acquired (the names of the classes are according to the previous section), with the realisation of the rest of the metrics in vertical columns. If a value is left out it means it is inherited from the general metric or the metric of the object. This list indicates presence by itself—if an object or element is not there, it should not be modelled.

In this example, buildings should be acquired with attributes of the occupancy and energy rating. The elements of the building that should be acquired are the walls, roofs with dormers, chimneys, balconies, piers and openings. In this example the roof, dormer, and chimney are required to be modelled with a finer feature complexity (0.2 m), overriding the general guidelines for the dataset.

Other types of city objects that should be modelled are roads, and street lights. Their specifications are self-explanatory following the logic for the specifications of the buildings. The street lights do not require any appearance, while the roads are coloured with a single colour. If a street light is shorter than one metre, then it should

Table 3.3: Example of the specification of a discrete LOD. This LOD is derived as a discretisation from a series of functions of metrics, with three city object types and their elements. For simplicity, the number of objects and elements is limited.

	LOD specification				LOD i			
General metrics			Feature complexity		0.4 m			
			Appearance resolution		0.3 m/px			
			Semantics		Yes, full spatio-semantic coherence			
City objects and elements	**Object**	**Feat. c.**	**Attributes**	**Elements**	**Feat. c.**	**Dim.**	**Appearance**	**Attributes**
	Buildings		+ Occupancy	Wall		2		+ Material
				Roof	0.2 m	3		
				Roof.Dormer	0.2 m	3		
				Chimney	0.2 m	3		
				Balcony		3		
				Pier		3		
				Opening		2		
	Roads		+ RoadUse	Traffic area- cars		2	Black	+ SpeedLimit
				Traffic area- bicycles		2	Red	
	Street lights	1 m	+ PowerConsumption	Pole		3		

not be modelled.

This table answers two relevant questions when modelling: which objects and elements to acquire and how, and we believe that it represents a concise way to represent an acquisition-modelling specification of a discrete LOD of a 3D city model.

3.5.4 Terminology

The classic and well-established term 'level of detail' implies a degree of generalisation or amount of geometric features of a 3D city model. Hence it is not entirely correct with respect to the definition given in this chapter, but also current concepts. LOD is often described in other incorrect terms, such as *level of accuracy* [490], *level of completeness* [491] or *level of quality* [492]. The term *detail* is loosely equivalent to the presence and complexity of the objects and their elements, which makes it incomplete. While more appropriate terms that would accompany this definition might be terms such as *level of abstraction* [158], *level of representation* [493], *levels*

of realism [159, 494], or *levels of complexity* [495] the conventional term of LOD is so well established in the GIS community that we do not propose any change here.

3.6 CONCLUSIONS

This chapter has given a comprehensive LOD analysis with a list of most shortcomings of current LOD paradigms both in computer graphics and 3D city modelling. The chapter observes that the outlook on LOD may be subjective, and that it is not easy to come up with a definition. This is also a challenge in cartography [496].

Furthermore, the chapter points out some misconceptions in the GIS community. For example, it reveals that many of the LOD concepts are nominal rather than ordinal, and that their value depends on different aspects (see Figure 3.1 for an example). It also reveals that the LOD concept should be treated separately from spatial quality, and that it is distinct from its cousin in computer graphics. These circumstances hamper the understanding of the concept and its description.

We have defined and formalised the LOD concept, and we have established a harmonised LOD framework that is applicable to any format and standard in 3D city modelling. We see the LOD as the degree of design correspondence between the model and the real-world object and indication of how thoroughly a dataset has been acquired, being driven by the geometry, appearance, semantics, and other related metrics that can be quantified. Such an approach enables a consistent specification of the data, and facilitates the translation between the standards and the comparison of their LODs by decomposing them into quantified metrics, since LODs are discretisations of continuous functions of the metrics.

It is also important to mention that the presented framework, while strict, does not solve all problems and does not purport to answer all questions. An example are curves, where the concept of feature complexity falls short. However, these are also problematic to specify in 2D GIS [497].

Our concept has several advantages over other concepts: (1) a thorough 3D city model base specification (with UML) used for expressing properties of 3D city models that can be used to create continuous LOD series and their discretisations enabling a higher number of LODs which are also consistent; (2) we recognise the importance of semantics and give an example how to extend and refine the semantic classes; (3) we give the possibility to decompose the city objects into more city objects based on their properties; (4) we give emphasis to the aggregation of objects and elements; and (5) allow mixing conventional LODs in the same city object by defining the requirements per each element rather than the city object.

The presented concepts are extensible and adaptive for different thematic classes, city objects and their elements. The LODs may also be defined for smaller, and on

the other side, large scales, enabling the applicability in BIM or very detailed virtual reality or architectural models, meaning that this framework is not restricted to 3D city modelling. The framework also enables finer distributions of LODs than presently available series of contemporary LOD concepts usually found in industry, academia, and government institutions. The progress of LODs can regard all the metrics, and can be more consistent compared to the present solutions.

After the publication of the papers on which this chapter is based, the LOD framework was found of interest to specify data requirements, e.g. in the frame of 3D cadastre and BIM for spatial planning tasks [280]. Furthermore, the discussion on feature complexity was adopted by Wong and Ellul [498] for defining geometric characteristics (metadata) of 3D city models. Finally, the framework was reviewed by Löwner and Gröger [499], concluding that it is the most comprehensive framework currently available. A point of criticism is that the concept has a low freedom of interpretation due to its formalism, precision, and scientific aspect. This aspect is addressed in the Chapter 4 where the insights obtained in this chapter are used to define a new LOD specification in a simplified manner, leading to straightforward understanding and leaving freedom of interpretation to suit specific use cases.

CHAPTER 4

Designing an LOD specification for buildings

This chapter is based on my paper [500]:

> Biljecki F, Ledoux H, Stoter J (2016): An improved LOD specification for 3D building models. *Computers, Environment and Urban Systems*, 59: 25–37. doi: 10.1016/j.compenvurbsys.2016.04.005

The previous chapter discussed and formalised the concept of LOD in 3D GIS, which is intended to differentiate multi-scale representations of semantic 3D city models. The aim of this chapter is to realise the framework obtaining a consistent LOD specification. In this chapter we focus on the crucially important geometric aspect of the LOD concept, and further analyse the prominent LOD concept found in the CityGML standard. Despite the popularity and the general acceptance of this LOD categorisation, we argue that from a geometric point of view the five CityGML LODs are insufficient and that their specification is ambiguous. We solve these shortcomings with a better definition of LODs and their refinement. Hereby we present a refined set of 16 LODs focused on the grade of the exterior geometry of buildings providing a stricter specification. The specification is designed using the insights presented in Chapter 3, by focusing on geometry related metrics of presence, feature complexity, and dimensionality. Our specification also includes two hybrid models that reflect common acquisition practices. The new LODs are in line with the LODs of CityGML 2.0, and are intended to supplement, rather than replace the geometric part of the current specification. While in our work we focus on the geometric aspect of the models, our specification is compatible with different levels of semantic granularity. Furthermore, the improved LODs can be considered format-agnostic. Among other benefits, the refined specification could be useful for companies for a better definition of their product portfolios, and for researchers to specify data requirements when presenting use cases of 3D city models.

4.1 INTRODUCTION

The OGC CityGML 2.0 standard contains one of the most prominent LOD taxonomies (see Section 3.2 and Figure 1.3). The five CityGML LODs have become widely adopted by the stakeholders in the 3D GIS industry and they now also describe the grade and the design quality of a 3D city model, especially its geometric aspect (i.e. 'how much detail should be acquired?'). They have gained importance also in the computer graphics [501, 502] and BIM [503] communities when dealing with 3D building models.

The LOD concept of CityGML is primarily intended to differentiate the grade of data resulting from different production workflows, and they are driven by semantics as much as geometry. In the industry and research community they were accepted from the outlook on geometric richness, which was partly caused by the lack of applications that require semantics. For example, we have observed that while the LOD2 from the point of view of CityGML developers represents a model with differentiated semantic surfaces, practitioners primarily refer to models with roof shapes, even when not dealing with data that is semantically structured. For this reason, from now on this thesis is mostly focused on the geometric aspect of the LOD.

While the five LODs generally provide a categorisation of the overall level of abstraction, content, value, and usability of 3D city models, this classification has several drawbacks and shortcomings as we further elaborate in Section 4.2. Since the specification is crucial among practitioners and researchers for conveying the grade of a 3D city model and its adherence to the real-world, in this chapter we present a refined specification to solve such problems. It should be noticed that the topic of refining and improving the current specification of the LODs is currently under consideration in the CityGML community for version 3.0 [499], and the work presented in this chapter is intended to aid the discussions. However, our work is independent of any particular 3D format, and applicable to any format that can be used to store 3D building models, including ones such as COLLADA and OBJ.

In Figure 3.4 in the previous chapter we have argued the difference between the LOD concept in 3D GIS and in computer graphics. In the context of this chapter, here we use the same figure giving an example of the shortcomings of the current CityGML LOD concept, from the point of view of the geometric detail. The image illustrates two LOD2 models: the model on the left has been acquired with two acquisition techniques, the walls are at their actual location and the roof overhangs are explicitly present. The representation in the middle has been acquired with one technique (aerial photogrammetry) where the walls are derived as projections from the roof outline. This example illustrates how the CityGML LOD concept is ambiguous and that it falls short in defining the complexity of the models: the two models are of

the same LOD (LOD2) according to CityGML while the first one is more laborious to acquire and it may bring better results in a spatial analysis (e.g. more accurate volume). Hence, practitioners would not consider them to be of equal value and usability. For these reasons we argue in this chapter that they should be considered as different LODs, and our specification differentiates such cases.

This ambiguity is most evident in the production of 3D models. For example, in 3D generalisation where researchers produce multiple geometric variants of LODs and discuss the ambiguity, among others see the papers of Guercke et al. [54], Fan and Meng [433], Stoter et al. [504], Noskov and Doytsher [505], and Deng et al. [506].

Solving the ambiguity is also important considering: (1) the increasing number of acquisition techniques (see the overview in Section 2.2); (2) the number of data producers and national mapping agencies requesting 3D data is increasing [507, 508], and without a finer specification data producers and users may resort to creating their own specifications (e.g. see Section 3.2), which might increase the ambiguity; (3) the increase in quantity of datasets with non-homogeneous LODs [497, 509]; and (4) use cases have different requirements when it comes to the complexity and quality of the data. Furthermore, the number of 3D use cases is rapidly increasing (see Section 2.3). Each of these use cases may have different requirements when it comes to the LOD of the models.

In this chapter we improve the geometric aspect of the LOD specification of 3D building models. We provide an extended and more informative series of 16 LODs that are compatible with the existing CityGML LODs. The refined taxonomy is a result of a research into currently available 3D city models and an investigation of the acquisition workflows. We review related work on this topic (Section 4.3), and for each LOD we give requirements and show an example (Section 4.4).

In this chapter we focus on the exterior of buildings (i.e. their exterior shell in LOD0–3). The refinement of the indoor and semantics aspect of the specification can be considered as perpendicular topics to this one. These topics are being tackled by other researchers who decompose it into different LODs and integrate them into expanded LOD1, LOD2 and LOD3 models (for examples see the work of Boeters et al. [129] and Löwner et al. [436]). While the semantic LOD and indoor LOD are out of scope of our work, present work on these topics is compatible with our work because such a specification can be supplemented to ours. For example, each of the newly refined LODs can be assigned a semantic LOD depending on the achieved spatio-semantic coherence.

4.2 SHORTCOMINGS OF THE CURRENT CITYGML LOD CONCEPT

Focusing on the CityGML 2.0 LOD concept, it might come to a surprise that this ubiquitous standard does not provide a strict specification for the five LODs. It gives short narrative descriptions, with a table (see Table 3 in the standard [10]) that is considered as a recommendation, and not a requirement. The description actually specifies the *upper limit* of each LOD, and not the minimal restriction for each, i.e. it restricts what can be a part of each representation. For example, LOD2 cannot contain openings, but it is not stated that LOD3 must contain openings.

Hence, besides an insufficient number of LODs and their condensed grouping, the main drawback of the concept in the current version of CityGML is that it does not mandate what features and how detailed they should be acquired, and therefore it leaves ambiguity and freedom for the implementation. For example, it is not stated whether dormers and other larger roof details should be acquired in LOD2. This may lead to misunderstandings between stakeholders, and errors in the utilisation of the models. For example, in solar potential estimations, which are most frequently carried out with LOD2 models, it is important to have roof superstructures since they cast shadows and they may reduce the area available for the installation of photovoltaic panels. Hence there may be substantial differences between analyses carried out with LOD2 models with and without roof superstructures. For this reason there is a need to differentiate between variants of LODs, and it is therefore important to provide a more expressive specification that diminishes errors caused by an ambiguous LOD specification.

CityGML 2.0 provides several conformance rules to test the validity of CityGML data, and there are other efforts such as the ones of Gröger and Coors [510], Wagner et al. [511], and Coors and Wagner [512] to provide extended modelling guides and rules. However, these do not cover the geometric detail of the models. This drawback results in many valid variants to be considered of the same LOD.

The size (e.g. length or footprint) of real-world features and their parts (e.g. a balcony of a building) that have to be acquired is designated as one of the main differentiators of the LODs, as argued in Chapter 3. However, this cannot be used as the general guideline to further specify LODs. For example, if an LOD2 requires that certain building parts bigger than a threshold should be acquired, this cannot be applied to windows, because they are not intended to be acquired in LOD2, irrespective of their size. A second example are roof overhangs (such as in Figure 3.4). They may be required by a stakeholder. However, in size they are smaller than other features that may not be required at all (e.g. dormers), hence each group of related features should be treated separately. Finally, nowadays a significant amount of models are constructed with a combination of different data sources. The LOD concept does not consider the LOD of combined data, where some parts of buildings may be ac-

quired in a finer or coarser detail than other parts.

These shortcomings could be solved together by providing a general list of features that should be acquired and the minimum size for each. However, CityGML does not provide such. We provide these in our specification described in Section 4.4.

4.3 RELATED WORK

The general LOD notion was examined in our earlier work in Chapter 3 where the concept is decomposed into six metrics: list of features, their geometric complexity, dimensionality, appearance, spatio-semantic coherence, and attributes. We take into account the first three metrics when defining the geometric aspect of the LOD.

Stoter et al. [504] recognise that CityGML lacks precise LOD definitions and allows ambiguity, and in a later research, Stoter et al. [513] argue that the specification should be further defined by practitioners, depending on the intended application of the 3D city model to be acquired. We agree with this reasoning, and think that our approach may help practitioners to do so in a standardised and justified way, while still leaving a significant degree of freedom to accommodate specific requirements of use cases.

Due to the ambiguity and the differences of models that CityGML considers to be of the same LOD, He et al. [514, 515] refer to the CityGML LODs as LOD groups, and further define inter-level LODs within the LOD1 group that vary in their geometric complexity. Besuievsky et al. [231] have a similar approach for LOD3 buildings where they create three variants of LOD3 buildings that are distinguished by the size and type of features to be acquired. We have considered their granular LODs when designing our refined specification.

Borrmann et al. [516] and Breunig et al. [517] provide an extended LOD specification for tunnels defining five LODs to create consistent multi-scale models and to use them for synchronous engineering collaboration. The LODs are discerned primarily by the list of railway elements that are acquired. Chen [430] does a similar work for trees defining four LODs i.e. 'Level Of Tree-detail'. The feasibility of the acquisition of these representations has been conducted with different airborne laser scanning scenarios.

In the BIM community, van Berlo and Bomhof [518] have worked on the refinement of the BIM LODs after analysing industrial practices and conducting a series of geometric tests. This is similar to our approach. Related to the BIM domain, Tolmer et al. [519] propose additional LODs to allow for a more transversal decomposition of data and objects organisation, and apply them to an urban motorway project.

Wate et al. [520] emphasise the importance of the relationship of the acquisition technique to an LOD, and give acquisition technique guidelines for each CityGML

LOD. Vosselman and Dijkman [521] notice that the capabilities (resolution) of acquisition techniques have a direct impact on the LOD of the reconstructed 3D city model. In our work we have analysed acquisition workflows, and we have taken them into account when designing the specification.

Döllner [492] expresses that in the current LOD approach it is difficult to integrate buildings from different sources and of varying LOD. Furthermore, they discuss the models that can be considered of an LOD between LOD2 and LOD3. Their observations are important for our work because we introduce two LODs that are designed to be acquired with a combination of different sources.

Benner et al. [434] and Löwner et al. [436] propose the orthogonal decoupling of the exterior and indoor geometry, and a refinement into multiple semantic LODs. The number of permutations, excluded by some prohibited variants, is large enhancing the specification, since they still fit within the present CityGML LODs. For example, a building with a coarse exterior with no semantic structuring may include a fine interior, and such has a unique designation. The refinement of the indoor LODs is a current research topic, which is also in focus of Hagedorn et al. [485], Kemec et al. [486], Billen et al. [522], Kang and Lee [523], Kim et al. [350], and Boeters et al. [129].

4.4 REFINED LEVELS OF DETAIL FOR BUILDINGS

We provide a series that contains 16 LODs (4 refined LODs for each of the LOD0–3), which are shaped after a literature review and inventory of presently available models by finding their main relevant similarities and mutual aspects. A visual example of the refined LODs is illustrated in Figure 4.1. We believe that these LODs allow for less ambiguity, and they aid practitioners to standardise their data with an improved definition of the complexity of the models. As mentioned before, this work is not intended to extend CityGML 2.0, it rather provides a supplementary specification that reflects the current practices and that conforms to the current concept, and at the same time solving the ambiguities elaborated in Section 4.2.

4.4.1 Methodology

Besides investigating workflows for producing 3D city models (e.g. [40, 524, 525]), we have examined several categories of sources of data. For instance, national standards and guidelines [418, 513, 526–528], examples of series and specifications of 3D city models not related to CityGML [43, 133, 182, 419, 420, 486], usually in the field of 3D generalisation [529, 530], visualisation [135, 531], and 3D reconstruction [532, 533].

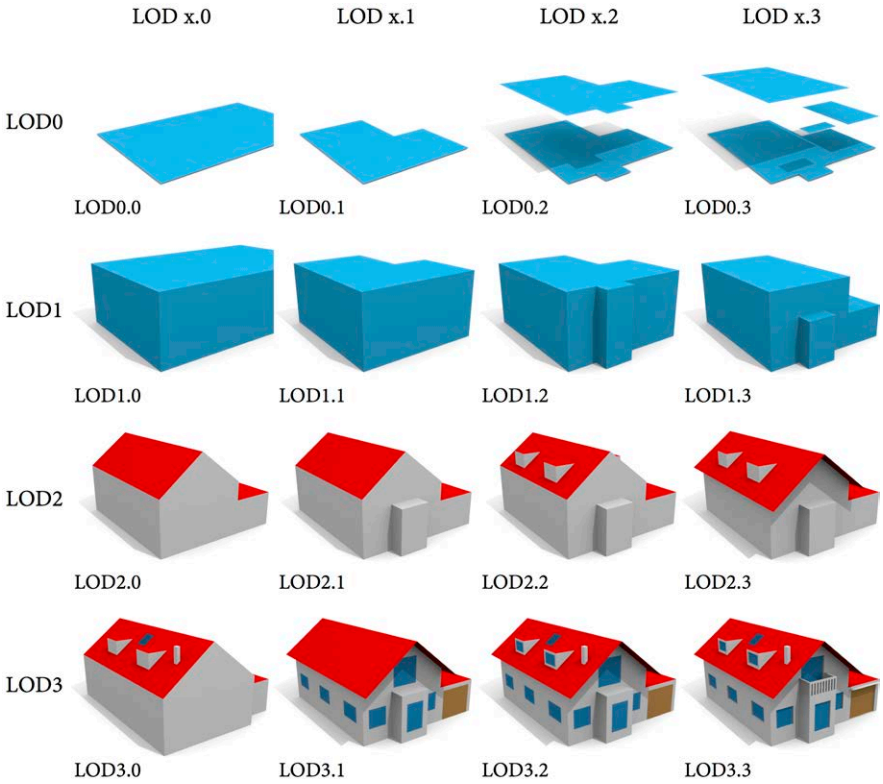


Figure 4.1: Visual example of the refined LODs for a residential building.

Furthermore, we have examined examples of models that refer to the CityGML LODs but do not appear to be stored in CityGML or have any other relation to the standard [22, 24, 289, 371, 534–540]. Finally, a number of publicly available datasets and specifications from companies, tenders, and local governments have been examined as well [181, 189, 421–427, 541].

We have analysed the available models from different angles: usability (their intended use cases), acquisition techniques (implying their cost and availability), and specification if available. The dozens of variants of models have been grouped into original CityGML LODs to which they correspond, and which we name *LOD families*. We have found a few general aspects, and various spatio-semantic ambiguities that surface in each LOD family, resulting in four groups of models within each family. The LOD groups are concentrations of models and partially imply their cor-

respondence to the real-world, acquisition technique, accuracy, domain of applications, etc. For each of these LOD groups a common set of requirements has been established, resulting in a specification and refined set of LODs.

Note that in our specification we do not focus on the semantic aspect and the texture. We address the amount of geometric detail that has to be acquired, by focusing on the list of elements of a building, and their granularity. Non-geometric requirements can be supplemented to our specification if required.

A few observations in the survey motivated specific choices. The first matter that we have noticed in our survey is the large number of unique specifications and combinations of various aspects. It is not possible to regard each of the aspects while retaining a reasonable number of LODs. Hence, in this process we have balanced the scope of the specification and the number of the refined LODs, and we have taken care that the models unambiguously correspond to a refined LOD. The goal is to provide a finer specification, but flexible enough to still allow some freedom of modelling and not to result in a too large number of levels. This is beneficial for use cases, since a large number of models have been acquired bearing in mind a specific use case, and a strict specification would not be favourable towards such practices.

Another observation that led to a specific choice was a small number of *outliers*—specifications that are not in line with the common practices in 3D city modelling. They are rather designed for a specific application and as such cannot be accommodated in a uniform specification. For example, Kemec et al. [486], Ioannidis et al. [542], and Frommholz et al. [543] define an LOD2 model with generalised footprint and fenestration; and the specification in Kartverket [544], which defines LOD1 models with non-flat top surfaces. Such models are not considered, since their inclusion would compromise the simplicity of our concept, but it would also not be in line with the standard CityGML LODs. However, the rationale of our specification (and framework presented in Chapter 3) can still be used to define such *customised* LODs in addition to our series.

4.4.2 Selection criteria for objects to be acquired

The selection of objects (e.g. building elements such as dormers) to be mapped is an important part of the LOD specification. This is also analogous to 2D maps [545], where the selection criteria is mostly based on their significance and minimum size, and it depends on the object's class [546]. This reasoning can be followed in 3D as well. However, it should be stated that the significance and size are both fuzzy terms that also depend on the use case, and cannot be strictly defined, as other concepts related to scale and LOD.

The minimum size can be expressed as the minimum length and/or width of an object, and/or the minimum footprint area. For example, a requirement may state

that dormers that are wider than 1 m, and/or their footprint projected onto the roof is larger than 1 m^2 should be acquired. This can be applicable to both the size of a feature and its granularity, e.g. minimum size of a land cover area or its *spikes*.

In expressing the thresholds, it is important to define both the 1D and 2D requirements. For example, a chimney may be longer than a dormer (1.5 m > 1.0 m), but much smaller when considering its footprint on the roof (0.15 m × 0.15 m ≈ 0.02 m^2). A 3D requirement (volume) will not be used because it is not applicable to all types of features (e.g. windows).

4.4.3 General rules and principles of the specification

The main principles of our specification are:

1. A model must adhere to all of the requirements to be considered of a specific LOD.

2. With a few exceptions, the specification is designed in such a way that each LOD$x.i$ contains more detail over LOD$x.(i-1)$, i.e. all requirements of an LOD$x.i$ should also satisfy the requirements of an LOD$x.(i-1)$. However, observe that an LOD$x.i$ is not necessarily more *detailed* than an LOD$(x-1).(i+1)$. This is comparable to the discussion related to Figure 3.4. Furthermore, note that LOD$(x-1).i$ is not necessarily derivable from LOD$x.i$.

3. The list of building elements to be acquired is composed from the most common occurrences of such features, such as windows and dormers. This list applies also to other elements of comparable function and size.

4. The selection criteria to model an object or not have been determined based on the minimum size of the building's element to be acquired. The minimum size is expressed as a distance, which can be applied to the length, width, or height of a feature, and as a projected footprint area. The footprint projection is not necessarily on the ground. For example, for windows the projection onto the walls is considered, and roof superstructures on the (inclined) roof. These requirements are specified separately for each feature's type (e.g. chimney). If no footprint selection criteria is stated, only the length criteria applies.

   Such minimum size selection criteria may cause disregarding some features that in certain settings could be important for a particular application. If that is the case, users may still opt for modelling smaller features beyond the threshold.

5. The specification provides general requirements that leave space for an extra number of variants. An LODx model that is modelled finer than it is required in LOD$x.i$, but below the specification of LOD$x.(i+1)$, is considered as LOD$x.i$. For example, if an LOD$x.i$ and LOD$x.(i+1)$ require all buildings parts larger than 4 m and 2 m to be acquired, respectively, and a model contains a part 3 m long, it may be considered as LOD$x.i$.

6. The specification defines that sizeable building parts, extensions and annexes such as garages and alcoves, may be acquired and treated distinctly. This should be distinguished from the cadastral point of view. Such objects are still part of the building, but their geometry is acquired in a way that these features are perceptually distinguishable.

4.4.4 Refined LOD specification for the geometry

The specification is given in Table 4.1, and a visual representation of the models is provided in Figure 4.1. In this section each LOD family is refined with four LODs that are described.

Because LOD1 is in essence an extrusion of the LOD0 model (or its generalised product), both are examined and redefined in conjunction.

LOD0 and LOD1 families The coarsest volumetric representation that the standard contains is the LOD1 model, a generalised model only described as 'the well-known blocks model comprising prismatic buildings with flat roof structures' [10]. Block models have also been described in 'patent language' in the US Patent Application by Guskov and Brewington [547], as a set of extruded polygons (right prisms) that comprise a volume defined by a base height from which extrusion begins, and an extrusion distance.

LOD0 is briefly described as a representation by 2.5D horizontal polygons with footprint level height and optionally roof level height [408].

LOD1 models are usually derived with extrusion to a uniform height [63, 64, 548–552], and generalisation from finer LODs [401, 553], such as a bounding box of an LOD2 [554, 555]. As a consequence, there are only horizontal and vertical surfaces, and no projection to the *xy* plane of two horizontal surfaces can overlap (not counting the *ground surface*).

Although LOD1 models are the coarsest volumetric representation, they can be derived from very accurate and detailed footprints [416, 556] (see also Figure 3.4).

LOD1 models provide a relatively high information content and usability comparing to their geometric detail [96, 132]. For example, they may be used for shadowing

Table 4.1: Specification of the refined levels of detail fitting the current CityGML 2.0 LODs.

Requirements	Refined levels of detail															
	0.0	0.1	0.2	0.3	1.0	1.1	1.2	1.3	2.0	2.1	2.2	2.3	3.0	3.1	3.2	3.3
Individual buildings		•	•	•		•	•	•	•	•	•	•	•	•	•	•
Large building parts (>4 m, 10 m^2)		•	•	•		•	•	•	•	•	•	•	•	•	•	•
Small building parts, recesses and extensions (>2 m, 2 m^2)			•	•			•	•		•	•	•	•	•	•	•
Top surface[(0)]			S	M	S	S	S	M								
Explicit roof overhangs (if >0.2 m)												•		•	•	•
Roof superstructures[(1)] (larger than 2 m, 2 m^2)											•	•	•		•	•
Other roof details (e.g. chimneys >1 m)													•		•	•
Openings[(2)] (>1 m, 1 m^2)													R	W	•	•
Balconies (>1 m)														•	•	•
Embrasures, other façade and roof details, and smaller windows (>0.2 m)																•

[(0)] Applicable only to LOD0.y and LOD1.y: S—Single top surface; M—Multiple top surfaces if the difference in height of the extruded building elements is significant (larger than 2 m).
[(1)] It includes dormers and features of comparable size and importance (e.g. very large chimneys).
[(2)] R—only openings on roofs; W—only openings on walls. In R, openings on dormers are not required.

simulations [90, 92, 239], estimation of noise pollution [259], energy demand estimation [114, 115], simulating floods [366], urban air flow analyses [320, 557, 558], visualisation [559], solar potential estimations [560, 561], determining the sky view factor [562], and satellite visibility predictions [209].

While the LOD1 model is the most basic volumetric 3D city model, it may be modelled in multiple ways. Götzelmann et al. [563], Glander and Döllner [158], Meng and Forberg [401] and Mayer [564] all generalise LOD1 models creating a coarser LOD1 model. Agugiaro [565] generates two variants of block models from footprints: one from a cadastral source, and one from a topographic map. Therefore,

we have identified the following relevant aspects in the LOD0 and LOD1 families:

- The models may represent individual or aggregated buildings (buildings that in reality are not adjacent and between which there is a gap, but are close enough to be modelled as one entity, at a smaller scale as in cartography). Some specifications enforce this, e.g. the 3D standard of the Netherlands requires individual buildings prohibiting their aggregation [513, 528]. However, it should also be noted that the notion of a building varies depending on the jurisdiction [566, 567].
- Since 2D footprints are extruded to a uniform height, and the resolution of the footprint directly implies the LOD of the 3D model, they are the focus of LOD0 and LOD1. However, their complexity may considerably deviate [568], from coarse to fine footprints as seen in Prandi et al. [569] and Ellul and Altenbuchner [570]. This is especially the case in generalisation from finer LODs [403, 571].
- Besides the footprint, LOD0 models may contain a roof-edge surface.
- Multiplicity of top surfaces: LOD1 models are not necessarily extruded to a uniform height, which is a common misconception about the production of the LOD1 model, since the number of top surfaces are not restricted by the standard. We have encountered several instances of enhanced LOD1 block models with *differentiated* roof tops [183, 425, 444, 515, 570, 572–574] that include multiple flat surfaces instead of a single surface for the top, to differentiate terraced houses, large roof constructions, etc. (see also Figure 3.3 in which LOD1 is taken from the CityGML 2.0 standard [10]). Related to extrusion, sizeable parts of buildings (e.g. veranda, carport, garage, and alcove) may be modelled separately with a different value of the height, resulting in multiple top surfaces, even if they *belong* to the same footprint. This may be to a degree incompatible with the traditional LOD1 notion, however, their occurrences warrant a separate LOD, and not all LOD1 models are derived with extrusion.

We define four LODs in each LOD family with the following minimum requirements: LOD0.0 and LOD1.0 are the coarsest models: they require all buildings larger than 6 m to be acquired, and buildings may be aggregated. These are the only two instances in our specification in which neighbouring buildings may be aggregated in a single geometric entity. In LOD0.1 and LOD1.1 buildings must be individually modelled and all large building parts shall be acquired. LOD0.2 and LOD1.2 have the requirement that smaller building parts and extensions should be acquired

(e.g. alcoves), and are extruded to a single height. This addition may result in more accurate spatial analyses, such as line of sight [201]. In addition, LOD0.2 requires the roof-edge polygon to be acquired as well. We have found that the LOD1.2 is the most common LOD1 model in practice. A shortcoming of most LOD1 models is that they encapsulate a single height for buildings, which might not be appropriate in cases of buildings with a complex architecture [538, 575]. Hence we introduce LOD0.3 and LOD1.3, which also require the same features to be acquired, but it allows multiple top surfaces if their differences are higher than a threshold (e.g. 2 m). For example, a large recess in a wall might have its height separately modelled and may be individually extruded. This approach can benefit use cases such as estimating the internal area. Finally, LOD1.3 cannot contain multiple horizontal surfaces at the same planar coordinate.

LOD2 family LOD2 is a more detailed model than LOD1, in which individual buildings are mandated, and are usually modelled as simple structures containing standard and simplified roof structures. They are commonly derived from point clouds or photogrammetry, and their combination with building footprints [576–578]. LOD2 models can also be derived with generalisation from an LOD3 [579].

The models provide a relatively favourable relationship between the costs of acquisition and usability. Acquisition-wise, they can be automatically derived from point clouds [44, 580, 581]. Usability-wise, they can be used in a wider range of applications than LOD1, such as the estimation of the solar potential of rooftops [81]), or as an improvement in accuracy over LOD1 (e.g. in energy demand estimation [78]).

Stoter et al. [513] recognise the ambiguity of roof overhangs in LOD2, i.e. whether they should be explicitly modelled or not, as overhangs may add value to certain use cases. The CityGML standard allows overhangs in LOD2 if known, but it does not require them (see again Figure 3.4). This results in various LOD2 models with and without overhangs (for examples see [284, 285, 582–585]; and [65, 433, 581, 586–590], respectively). Most of the models from national mapping agencies do not have explicitly modelled roof overhangs [67, 227]. LOD2 models with differentiated roof overhangs cost more to acquire since they generally require a combination of two acquisition techniques (airborne and terrestrial). When roof overhangs are not available, the walls are usually obtained as projections from the roof edges to the ground, inherently increasing the volume of the building.

A second ambiguity of LOD2 buildings are building installations such as dormers, and chimneys, which are allowed in LOD2, but they are not always found in practice (Figure 4.2). LOD2 models that contain building installations only include large features which considerably protrude the wall or roof structure [510, 521, 591]. The

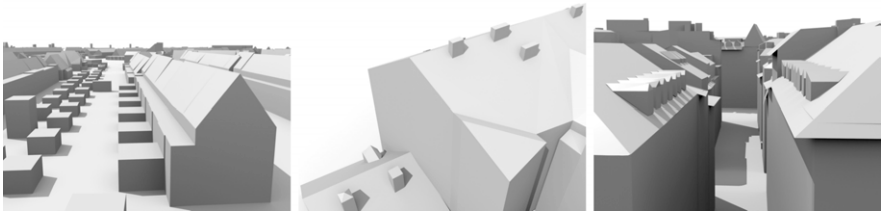


Figure 4.2: Roofs in LOD2 models mapped in different detail. The roofs in these datasets are modelled according to different levels of geometric detail: *clean* generalised roofs with no detail, roofs containing superstructures such as dormers and large chimneys, and roofs with explicitly modelled overhangs. The current CityGML LOD specification does not acknowledge these differences.[16]

German Federal 3D building modelling guideline [527] explicitly states that dormers and other objects of similar size should not be acquired. On the other hand, we have encountered LOD2 models which have dormers and other roof superstructures modelled, in academia [548], national mapping agencies [507, 544], and in municipal datasets [183].

We have grouped the occurrences of LOD2 models into four LOD2 variants based on the aforementioned aspects:

- LOD2.0 is a coarse model with standard roof structures, and potentially including large building parts such as garages (larger than 4 m and 10 m^2).
- LOD2.1 is similar to LOD2.0 with the difference that it requires smaller building parts and extensions such as alcoves, large wall indentations, and external flues (larger than 2 m and 2 m^2) to be acquired as well. In comparison to the coarser counterpart, modelling such features in this LOD could benefit use cases such as estimation of the energy demand because the wall area is mapped more accurately.
- LOD2.2 follows the requirements of LOD2.0 and LOD2.1, with the addition of roof superstructures (larger than 2 m and 2 m^2) to be acquired. This applies mostly to dormers, but also to other significantly sized roof structures such as very large chimney constructions. Because the roof is mapped in more detail, this LOD can be advantageous for the estimation of the insolation of roofs.
- LOD2.3 requires explicitly modelled roof overhangs if they are longer than 0.2 m, therefore the roof edge and the footprints are always at their actual loca-

[16] Data: Rotterdam, Netherlands (courtesy of the City of Rotterdam); Zagreb, Croatia (courtesy of Geofoto Ltd); and Geneva, Switzerland (courtesy of Geneva's Territory Information System).

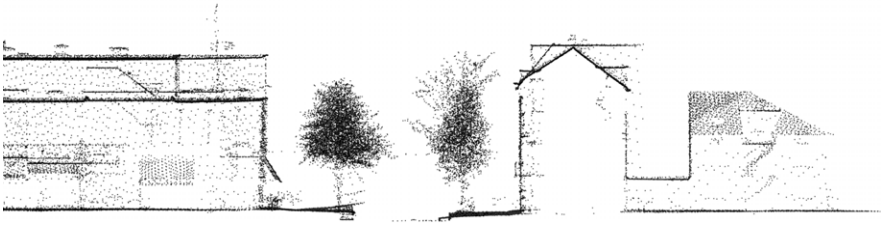


Figure 4.3: Dense airborne lidar scanning provides sufficient data for LOD2.3. This cross section view of the recent national lidar campaign in the Netherlands (AHN3) suggests that overhangs and roof superstructures are discernible in dense point cloud acquired only from an airborne platform. This may lead to LOD2.3 generation in future only from airborne lidar, without the need to integrate other sources of data.[17]

tion, which has advantages for use cases that require the volume of the building. The production of such models usually involves the integration of data from multiple sources. However, with the advancement of airborne sensors, it is foreseen that in future the acquisition of LOD2.3 will be more common (Figure 4.3).

LOD3 family LOD3 adds openings (i.e. windows, doors), balconies, more detailed roof structures (e.g. chimneys and antennas), and mandatory roof overhangs. This enhancement benefits some applications, e.g. openings are important for estimating heat losses [319], luminance mapping and glare analysis [110], planning energy-efficient retrofits [124], and for accounting the area available on vertical walls for solar panel installation [94]. LOD3 models are also appreciated in visualisation [592].

The acquisition of LOD3 models is a laborious process [184], hence they are in practice of limited availability, and are usually restricted to smaller areas. They are normally derived from terrestrial laser scanning [593, 594], very dense airborne laser scanning [595], their combination [40], from the conversion and generalisation from architecturally detailed models such as BIM [5, 596–598] and CAD [46, 48], from architectural plans [599], ground imagery [600], and with procedural modelling [169, 405, 601]. However, recent research in remote sensing focuses on automatic detection of windows and other façade details [535, 602, 603], leading to an automatic acquisition of LOD3 models [604].

LOD3 models are significantly more detailed than LOD2 models, and less ambiguity is present from the 3D GIS point of view (while it would be possible to nitpick

[17] Point cloud data courtesy of the Actual Height Model of the Netherlands (AHN).

among different variants of LOD3, it is safe to assume that they will make no difference in spatial analyses in the GIS domain). The only differentiation between LOD3 models that we have detected is the minimum size of features that are acquired, especially whether the embrasures of openings are modelled or not. For example, Besuievsky et al. [231] create multiple variants of LOD3 models, one with flat windows, one with the embrasures, and a third with minor façade details. The finer two have the roughly the same size of linear features rendering their difference negligible from the GIS standpoint, hence, we merge them.

We define: LOD3.2, an architecturally detailed model that contains features of size larger than 1.0 m, and LOD3.3—one that contains features of size larger than 0.2 m, including embrasures of windows (i.e. making windows '3D'), awnings, and similar features of comparable size. The latter may be beneficial over the former for high-quality visualisation and virtual reality applications [147], and it is usually a product of the conversion from BIM and architectural models. While it is less likely that LOD3.3 will make a significant difference in most GIS analyses with respect to LOD3.2, its delineation is necessary in order to recognise datasets produced by conversion from BIM and other highly detailed sources.

However, we have encountered a number of models that cannot be fully accommodated in the traditional LODs, but are common in the acquisition workflows and which technically belong to LOD3. For example, Franić et al. [181] and Novaković [189] create a model from an aerial survey with roof details finer than in LOD2.3, but with other features of lesser detail comparable to LOD2. We denote such models of mixed LODs (e.g. different LOD for *aerial* and *terrestrial* features) as *hybrid* LODs. Two such LODs that reflect the acquisition workflows have been defined, and to attempt to accommodate models of specific configurations. Because they both contain openings, we add them to the LOD3 group.

LOD3.0 is a model where roof structures are mapped in finer detail than LOD2.2, but other features such as walls are acquired as LOD2.2. Roofs may include roof windows. Windows of dormers do not have to be acquired.

LOD3.1 is its terrestrial counterpart, defined for terrestrial acquisition techniques, such as mobile mapping systems [26]. Since these techniques operate from the ground, roof features may be out of reach and it is mapped in coarse detail [605]. Therefore this instance requires all features below roof to be acquired at LOD3.2 grade, and the roof as LOD2.3, since overhangs may be explicitly modelled. This LOD is advantageous for use cases that require only the wall surface to be modelled in finer detail, such as estimating the solar potential of façades or for pedestrian navigation.

4.5 CONCLUSIONS

The current LOD categorisation of CityGML has two shortcomings:

1. lack of a precise specification of each LOD; and that
2. the current five LODs are too generic and therefore they are not always capable to separate one LOD from the other (i.e. two significantly different levels of abstraction may still be considered as the same LOD as per the current specification).

We introduce a refined specification solving these shortcomings. Our refined LODs present multiple geometric realisations for the standard CityGML 2.0 LOD groups. The refined LODs that we have introduced are a result of a literature review, analysing acquisition workflows, and examining publicly available specifications. The specification is compliant with the existing LOD concept in the international standard CityGML. Hence, while improving the shortcomings of the current concept, the refined specification does not damage the commonly used five standard LODs—it is completely compatible with it. With this specification it is possible to determine the LOD of an existing dataset, and to store in the documentation or metadata of the model (since CityGML 2.0 does not support storing such information). However, we foresee that our specification may be integrated as a user defined profile in the upcoming CityGML 3.0, where a generic LOD framework is planned [411, 499].

We have covered the vast majority of CityGML LOD cases we have found in practice, and while further differentiations are possible, we believe that it is not beneficial to define more than 16 LODs when only considering the geometry. The extended specification is not complicated and it is intended to be of special interest to the data producers, addressing their critic of ambiguity that the current LODs present. Most importantly, not a lot of modelling freedom is allowed anymore, diminishing potential misunderstandings between 3D stakeholders and potential errors in the usability of the models.

As much as the current LODs of CityGML 2.0 are used outside the CityGML format, these improved LODs may also be considered as independent of CityGML, and applicable to any other 3D format or context. For example, they may become an important factor in contracting the data acquisition, and as a more precise paradigm to express and benchmark the capabilities of a 3D city model reconstruction technique. Furthermore, practitioners can find them useful to better define their product portfolios, and national mapping agencies to implement them in their specifications. The classification has already been adopted by the EuroSDR 3D Special Interest Group

(SIG) as a base for a forthcoming European standard for national mapping in 3D. Finally, they can be useful to express the LOD of models they are analysing, primarily in work on generalisation and 3D use cases.

Even though we did not use to the full extent of the strict framework presented in the previous chapter, the refined specification is considerably more expressive and precise than the one in the current standard. A balance between a formal and practical specification is achieved to ensure its adoption in the community. An advantage of permitting a degree of flexibility is allowing additional requirements driven by use cases, following the reasoning of Stoter et al. [513].

This improved LOD specification is the core contribution of this thesis. Most of the following chapters will use it as a base for answering the research questions. For example, in Chapter 10 we conduct detailed experiments of the performance of each LOD in a particular spatial analysis, aiding the practitioners to choose the optimal LOD for a use case while balancing the costs of the acquisition.

For future work, we plan working on the semantic aspect and on other thematic modules. From this chapter, the thesis focuses mostly on the geometric component of the LOD concept. However, semantics will not be completely ignored, and it will be subject of a few of the forthcoming chapters (e.g. in Chapter 6 the presented specification will be realised with a sample dataset with multiple levels of semantic detail for each of the refined 16 LODs; and in Chapter 9 a spatial analysis will be carried out with multiple datasets of different semantic richness).

Furthermore, we plan to investigate how to automatically validate whether a dataset is modelled according to a certain LOD.

CHAPTER 5

Variants of LODs

This chapter is based on my papers [606] and [607]:

> Biljecki F, Ledoux H, Stoter J, Vosselman G (2016): The variants of an LOD of a 3D building model and their influence on spatial analyses. *ISPRS Journal of Photogrammetry and Remote Sensing*, 116: 42–54. doi: 10.1016/j.isprsjprs.2016.03.003

> Biljecki F, Ledoux H, Stoter J (2014): Height references of CityGML LOD1 buildings and their influence on applications. *Proceedings of the 9th 3DGeoInfo Conference 2014*, Dubai, UAE. doi: 10.4233/uuid:09d030b5-67d3-467b-babb-5e5ec10f1b38

The implementation portions of these two papers are presented in the Chapter 10, dealing with experimental research.

The previous chapter introduced a refined LOD specification. This specification is stricter than the one found in the CityGML standard, seeking to diminish misunderstandings between stakeholders, among other advantages. However, specific acquisition approaches may produce data of the same LOD that geometrically differ from each other. For example, the top surface of an LOD1 block model may be modelled at the eaves of a building or at its ridge height. Such variants, which we term geometric references, are often overlooked and are usually not documented in the metadata. In response to this research gap, we investigate a variety of LOD1 and LOD2 geometric references that are commonly employed. The chapter catalogues geometric references found in practice, and presents a specification that serves as a supplement to the one introduced in Chapter 4. A notable result is that there are 21 valid variants of LOD1 meaning that there are many different ways to realise a relatively simple model such as LOD1. The enhanced specification developed in this chapter further reduces ambiguity in the production of 3D city models, concluding the theoretical part of the thesis.

5.1 INTRODUCTION

So far this thesis has indicated that in practice the LOD concept is used to specify the fineness of the geometry of the models to be acquired. However, specification-wise the LOD is only one of the aspects to consider when acquiring 3D city models because from a geometric standpoint, there exist multiple variants of models within each LOD. For example, an LOD1 block model of a building may be modelled in a multitude of possibilities (Figure 5.1): among other options, the top surface might represent the height at the eaves of a building or the height at the top of the construction. If we ignore the elevation, the footprint may be modelled at the position of the walls, or it may represent a projection of the roof edge polygon to the ground. This example already results in four variants of an LOD1 model, a fraction of all the possibilities, as we point out in Section 5.3.

It is our experience that these modelling choices, which we describe as *geometric references*, are often overlooked by practitioners when acquiring, processing and utilising 3D city models. Furthermore, they are rarely documented in the metadata of the dataset, usually because it is not possible to store such information, as in the case with CityGML. The awareness of the geometric reference is important because, as we suggest later in the thesis, different geometric references within the same LOD may lead to considerable differences in the results of a spatial analysis. As a conse-

Figure 5.1: Four variants (geometric references) of an LOD1 block model. The elevation of the top surface of a block model may be modelled at, among other options, at the eaves and at the top of the construction. Similarly, the footprint (and therefore the *walls*) may be modelled at the footprint or at the outline of the roof edges. Such combinations result in a multitude of modelling possibilities within the same LOD, which can cause errors in a spatial analysis if not documented properly.

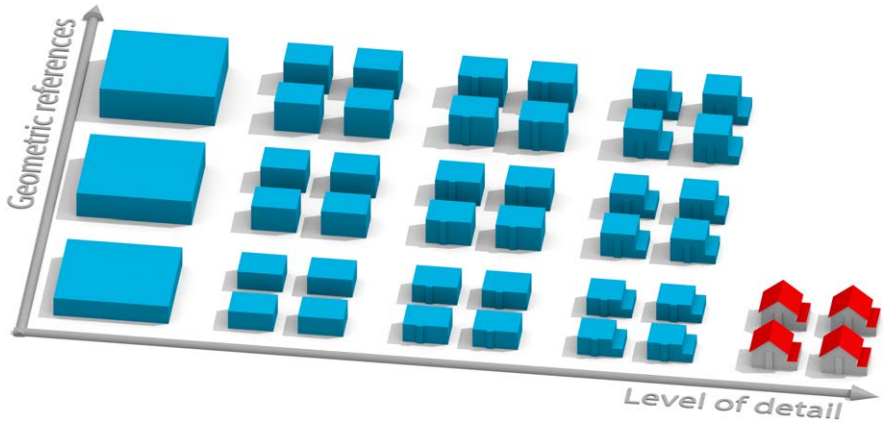


Figure 5.2: The orthogonal relationship between the LOD and geometric reference concepts. The image contains five LODs (four LOD1 block models: LOD1.0, LOD1.1, LOD1.2, LOD1.3; and LOD2.3 as a more detailed model for reference). Within each of these LODs there exist multiple variants of geometric references. This example illustrates the different geometric references for the height of the LOD1 block models (heights at the eaves, half of the height of the roof structure, and top height of the building) The figure is limited since it is only a subset of the possible LODs and GRs.

quence of the ambiguous specifications, this may lead to errors in the utilisation of the models.

In this chapter we provide an insight into this topic by covering the following aspects: we (1) derive an inventory of the most frequent geometric references, based on a survey of current practices of acquisition and modelling (Section 5.3); and we (2) propose a number of recommendations, such as an extension of the INSPIRE Building model standard (Section 5.4).

While we focus on CityGML, our work is applicable to any other 3D standard and LOD taxonomy.

5.2 BACKGROUND AND RELATED WORK

We define a geometric reference (GR) as the modelling choice of the boundaries of the captured feature. This concept is orthogonal to the concept of LOD, since the LOD refers to the spatio-semantic richness of the representation (Chapter 3). The relationship between the LOD and geometric reference concepts is outlined in Figure 5.2.

In this research, we focus on the two geometric references that, in our opinion, account for the majority of the ambiguities found in practice:

- Vertical reference of the top surface of block models: what does the elevation of the top surface of the block model represent. This is applicable to LOD1 models only, and those LOD0 models that contain the roof edge polygon (however, note that in this chapter we will not focus on LOD0).
- Horizontal reference of the footprint: what does the footprint (and the generated walls) represent. This is applicable to LOD0, LOD1, and LOD2 models, and to some extent to LOD3 models that are derived by supplementing LOD2 models with detailed façade geometry [602] (i.e. our LOD3.0 model), and mixed-LOD models [181, 189].

5.2.1 Acquisition of LOD1 and LOD2 models and their ambiguities

The reason for the existence of the multiple geometric references lies in the lineage: the different workflows and approaches for acquiring 3D city models. This is especially the case for LOD0, LOD1 and LOD2, which are largely derived automatically or semi-automatically with a number of different techniques (see Haala and Kada [578], Tomljenovic et al. [608], Musialski et al. [502], and Verdie et al. [501] for overviews).

Figure 5.3 clarifies this diversity by illustrating some of the general scenarios to derive LOD1 and LOD2 models: (left) the acquisition with airborne techniques, (right) with terrestrial observations coupled with the information about the height, and (centre) the scenario of the combination of the airborne and terrestrial measurements.

Airborne techniques (airborne laser scanning—ALS, and photogrammetry) are frequently employed for deriving LOD1 and LOD2 models (e.g. see the work of Suveg and Vosselman [44], Xiong et al. [525], Rottensteiner [609], Sirmacek et al. [610], and Demir and Baltsavias [43]). These techniques generally involve the acquisition of the roof surface, and subsequently the projection of its edges to the ground to derive the walls and the footprint. This inherently causes buildings to be wider than they are in reality. In such a scenario, LOD1 models are usually derived by constructing a horizontal plane at an elevation such as roof edges or roof ridges, and LOD2 models do not contain differentiated roof overhangs—they are part of the (combined) geometry of the roof (with the exception of our LOD2.3).

When airborne techniques are supplemented with terrestrial measurements, such as a geodetic survey (centre example), the footprint and walls are at their *actual* position. Their location also serves as a *constraint* for the reconstruction of the models. For instance, in case of ALS, only points with planar coordinates within the footprint

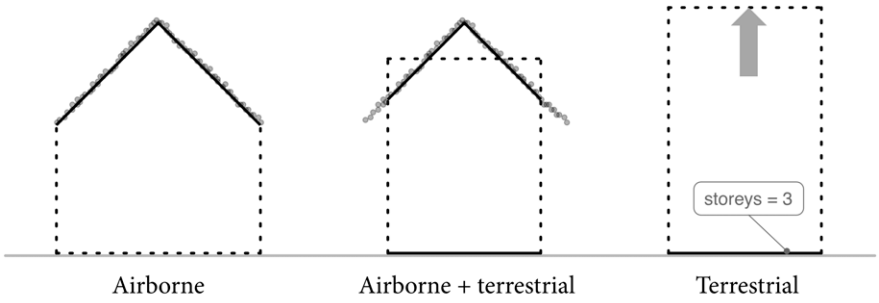


Figure 5.3: The acquisition approach dictates the geometric reference. This example illustrates some of the common acquisition workflows accounting for the vast majority of sources of LOD1 and LOD2 models, resulting in different geometric references.

are considered in the reconstruction of a building. This means that in the reconstruction of the LOD2 model the roof surface may be smaller with the roof edges reduced, since the area representing the roof overhangs is usually not included, as it is constrained by the footprint polygon. When producing LOD1 models, the elevation of the top surface is commonly derived from a statistical analysis of the lidar data, such as the median or average of the heights of all points within the footprint. Both the LOD1 and the LOD2 models are shown in the centre example.

Finally, the example on the right indicates the extrusion from 2D footprints in combination with various forms of attribute data, such as the number of floors [141, 405, 549], or a building height derived in cadastral measurements, e.g. height at the eaves [67, 514]. We revisit this topic in Chapter 7 where we present a novel method to extrude building footprints with non-elevation data.

A prominent technique of deriving LOD1 and LOD2 models, that is outside this observational context, is generalisation from existing models at finer LODs [403, 529, 530, 554, 555, 579]. In the same way, most papers on generalisation do not specify the horizontal and vertical reference of the generalised models.

5.2.2 Related work

The topic of geometric references has been explored in 2D GIS only to some extent, especially in point-based representations. For example, on a coarse scale a city may be represented as a point. The placement of the point (e.g. centroid versus a point placed at the most populous area of the city) may affect the calculation of the distance between two cities. Such examinations have been the subject of many research papers [611–614].

In 3D research on this topic is limited. The INSPIRE Building model is a relevant source on this subject, as it provides metadata to express the references in 3D city models, and we cover the document in Section 5.3.1.

Brasebin et al. [227] partially investigated this problem. They term the different horizontal references as *modelling choices* and estimate the influence of two references on the estimation of the sky view factor—the degree to which the sky is obscured by surrounding buildings [226].

Sargent et al. [552] point out that different users prefer different height values of buildings in 2D topographic databases, and recommend that national mapping agencies should provide more than one building height value. The values are termed as *building height characteristics*.

Pedrinis and Gesquière [615] acknowledge multiple forms of footprints as *inconsistencies*, and deem them inconvenient when matching datasets from multiple databases. In their work they present a remedy for correcting the offset between two data sources caused by different geometric references, to allow their merging.

Oude Elberink [616] recognised the problem of the ambiguity of the uncertain reference of the footprint when reconstructing the LOD2 models from point clouds in conjunction with 2D data. The research does not further investigate this topic, but it is important to mention as one of the first sources we have found that indicates the implications of unknown geometric references, e.g. in the combination of multiple geo-data sources to produce models of unknown lineage.

Another work that is related to this chapter is the 2016 paper of Ellul et al. [209], investigating the impact of geometric references in LOD1 models on the prediction of GNSS visibility. This recent study involved generating multiple LOD1 models with different heights based on lidar settings (e.g. median and maximum elevation). The experiments indicate that there is an impact of different GRs in the results of that particular spatial analysis.

5.2.3 Refinement of LODs

In Chapter 4 we have presented a redefined specification of LODs. This chapter continues the work and the presented concepts are developed on top of the LOD specification. The geometric references that we investigate are applicable to the refined LODs (in Figure 5.4 we indicate the relationship between the two concepts). Each of these may be modelled according to several different geometric references (e.g. $\text{LOD1.2}^{\text{GR1}}$, $\text{LOD1.2}^{\text{GR2}}$, etc.), resulting in a large number of different representations. There are some exceptions. For example, the refined specification differentiates a finer version of the LOD2 model, specified as LOD2.3, that is acquired as a combination of terrestrial and airborne techniques, where the roof overhangs are explicitly modelled. In such cases there are some restrictions, e.g. the walls cannot

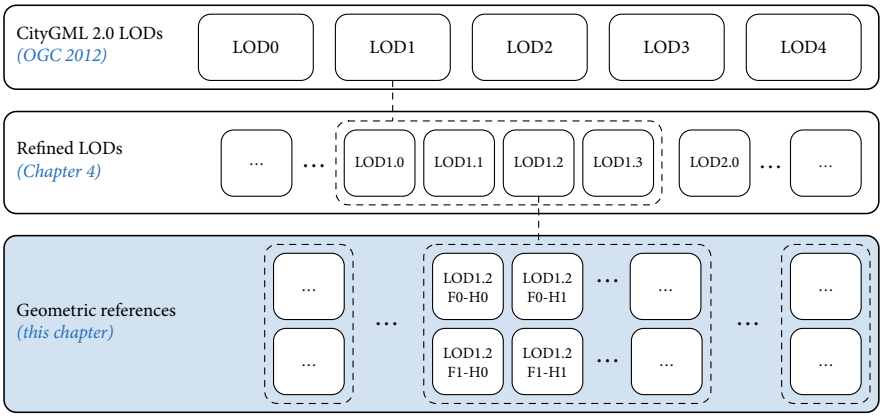


Figure 5.4: The relationship of this chapter with the previous work and other efforts. Our work extends the refined LOD specification by describing multiple variants of each LOD, resulting in dozens of combinations. The shorthands will be explained thoroughout the text.

be modelled as projections from roof edges because that is not compatible with the LOD specification.

5.3 INVENTORY OF THE REFERENCES IN LOD1 AND LOD2

5.3.1 The INSPIRE Building Model

The INSPIRE Data Specification for the spatial data theme Building [617, 618] covers 2D and 3D representations, and focuses on the footprint and elevation in the context of buildings, representing a solid foundation for our work.

The specifications mandate that the horizontal and vertical levels of a building that was chosen to represent its footprint and elevation have to be documented. It presents a code list and definition of a number of references of buildings in LOD1 and LOD2, that are intended both as descriptors of the geometric representation and attributes (e.g. in 2D data). These references are represented by a value type `horizontalGeometryReference` or `ElevationReferenceValue`, i.e. a list of self-explanatory elements considered to capture a horizontal or vertical geometry.

For the vertical reference (elevation of the top surface of the LOD1 block model), this is realised through the attribute `verticalGeometryReference3DTop`, and a value from the corresponding code list `ElevationReferenceValue`, such as `topOf-Construction`. The list is extensive as it takes into account uncommon cases, such as the instance where the height of the eaves is not equal (e.g. the eaves on one side

of the building are higher than the eaves on the other side). At the same time, the standard contains the value `generalRoof`, which is ambiguous considering that it may refer to any point on the roof.

While the standard recommends a number of references for the top surface in LOD1, as we outline in Section 5.3.2, this list is not complete. For instance, another relevant value may be the highest point of a building (not to be confused with `topOfConstruction`), a height level that includes above-roof elements such as chimney and antennas, and that is frequent in generalisation, where the coarse models are sometimes constructed as bounding boxes capturing the extent of a building. This level is a possible value in the code list `ElevationReferenceValue` as `highestPoint`, but for some reason it is not listed as a recommended value for `verticalGeometryReference3DTop`.

The second relevant concept is the footprint, which is also covered by the INSPIRE Building model. The reference for the geometry of the footprint is expressed through the `HorizontalGeometryReferenceValue`, with possible values such as `footPrint` and `roofEdge`. This reference is applicable to both LOD1 and LOD2.

5.3.2 A survey of geometric references in practice

We have made a survey of current practices of horizontal and vertical references in LOD1 and LOD2 through an extensive review of research papers that deal with the acquisition of 3D city models, by contacting 3D acquisition companies, and by investigating the specifications of national mapping agencies (NMAs).

An overview of research papers [54, 515, 554, 573, 574, 579, 588, 619–624] yielded an extensive overview of the references, but also strengthened our impression that the majority of research papers that describe methods to acquire buildings do not explicitly elaborate on the employed geometric reference.

The specifications of local governments and national mapping agencies contained some information about the geometric references. We have obtained them through publicly published specifications of data or tenders (e.g. Netherlands [513], UK [552, 625], Germany [67, 526, 527], and Switzerland [551]), and through our involvement in the EuroSDR 3D Special Interest Group [507].

Most of the references that we have found are standardised by the INSPIRE Building model (Section 5.3.1). However, while INSPIRE provides a sizeable list of vertical and horizontal geometric references, we have discovered that there are additional values occurring in practice, rendering the standard incomplete. We list these below.

5.3.3 Vertical geometric references (top of the LOD1 block model)

Our survey has suggested that the height of the top surface of the LOD1 block model may be modelled at a multitude of different elevations. We list and describe them in details, and group them into three categories. For each height reference we assign an internal shorthand for easier referencing in the continuation of the thesis, and in Table 5.1 we give an overview with a relation to the INSPIRE Building model where possible. Table 5.1 also highlights that the INSPIRE Building model does not cover all references, hence this inventory can be seen as possible extension the standard (this is elaborated in Section 5.4).

A. References related to the roof structure As indicated in Section 5.2.1, the vertical geometric references are mostly related to the roof structure, and this category accounts for most of the specifications observed in the survey. For example, in photogrammetry the height of LOD1 models is usually taken from the roof edges or at the ridge of the roof.

H0 Height at the roof edges. Because of the roof overhangs, roof edges may have an elevation that is lower than the one of the highest point of the walls, hence this is the lowest point of the roof structure, and therefore the lowest possible reference of the top surface.

H1 Height at the roof eaves (the intersection of the roof and the wall plane; see Figure 5.1). This value is common in terrestrial measurements, and it is usually not visible for airborne techniques. It corresponds to the reference H0 in the case when there are no roof overhangs.

H2 Height at one third of the height of the roof structure (with H0 as the lowest point of the roof structure).

H3 Height at half of the height of the roof structure.

H4 Height at two thirds of the height of the roof structure.

H5 Height at the top of the roof (top ridge). This is a value typical in the generalisation from models with a finer LOD. It can also be derived from point clouds.

H6 Height at the top of the construction of the building. This value encompasses the whole construction (similar to a bounding box), and it is usually used with generalisation from LOD3 where antennas and chimneys are available. In case there are no such structures that extend beyond the top of the roof, the value corresponds to H5.

B. References based on the statistics of a point cloud Many LOD1 models are derived from airborne lidar data by extruding the footprint to a certain height derived from the points whose projection is within the footprint of the building (see the central scenario in Figure 5.3). Some of these references overlap with certain references in the previous category. However, the usual approach is to take the average or the median height of all points within a footprint [513, 539, 626].

This approach is ambiguous and it depends on the characteristics of the ALS survey, and the reflection properties of the various roof surfaces of a building. For example, due to the relative position of the aircraft, in one occasion the point cloud of a building may contain only points that represent the roof. However, in another survey the point cloud of the same building may contain lots of points representing walls (if not filtered), essentially resulting in a different elevation of the median of the heights and other similar statistically derived heights from a point cloud.

HL-avg Height derived from the average of the heights in a lidar point cloud. For buildings with sloped roof in practice it is usually between H2 and H4, however, for flat roofs without roof structures it may be below the elevation of the roof due to points on walls if not filtered.

HL-med Height derived from the median of the heights in a lidar point cloud, favoured over the average to filter outliers. In practice it is usually between H2 and H4, however, for flat roofs it may also be below H0 if points on the wall are not filtered.

HL-max Height of the highest elevation in the point cloud. If there is no vegetation, usually it corresponds to H6.

C. Non-elevation attribute references We have encountered a number of 3D city models obtained with the extrusion of footprints to an elevation that is available as an *indirectly derived* attribute such as the number of floors [141, 405, 549]. Such attributes may not overly reliable for determining the height of a building, but nevertheless they are not uncommon as they provide an approximate height that may be sufficient for visualisation and similar purposes [158, 163]. We jointly refer to this reference as:

Hx Height derived from non-elevation data, such as the number of floors (e.g. number of floors f multiplied by an assumed floor height h).

It should be noted that in Chapter 7 we revisit this topic by developing a method for reconstructing LOD1 models from non-elevation data, and we evaluate their quality.

Table 5.1: List of vertical geometric references (representations for the height of the top surface of the LOD1 block model). The equal sign means that the reference is re-used from INSPIRE. The asterisk (*) indicates that `generalRoof` could correspond in most of the cases, but not always (especially in the case of flat roofs).

Code (§5.3.3)	Height at	INSPIRE ref.	Our ref.
H0	Roof edges	`generalRoofEdge`	=
H1	Roof eaves	`generalEave`	=
H2	One third of the roof height	`generalRoof`	`oneThirdRoof`
H3	Half of the roof height	`generalRoof`	`halfRoof`
H4	Two thirds of the roof height	`generalRoof`	`twoThirdRoof`
H5	Top of the roof (i.e. ridge)	`topOfConstruction`	=
H6	Highest point of the building	`highestPoint`	=
HL-avg	Avg. height of the point cloud	*	`avgHeightLiDAR`
HL-med	Med. height of the point cloud	*	`medHeightLiDAR`
HL-max	Max. height of the point cloud	*	`maxHeightLiDAR`
Hx	Varies. E.g. $f \times h$	N/A	`NonEleAtt`

5.3.4 Horizontal geometric references (footprint of LOD1 and LOD2 models)

The list of horizontal geometric references is shorter, and it is closely related to the used acquisition technique. The two main references (accounting for virtually all models we have found) are:

F0 The footprint is modelled at its actual location. This reference is typical for terrestrial measurements, and it corresponds to the INSPIRE reference `footprint`.

F1 The footprint is derived as a projection of the roof edges of the building. In case there are no roof overhangs it corresponds to F0. How much a model with F1 deviates from the reality essentially depends on the length of the roof overhangs. INSPIRE labels this reference as `roofEdge`.

In Oude Elberink [584] and Schwalbe et al. [588] we have encountered an *artificial* reference that is derived by offsetting the F1 footprint by a fixed length to 'compensate' for the roof overhangs to produce models that attempt to resemble closer the reality. Such a reference applies to LOD2.(0,1,2) models—the measured roof edge is truncated by a distance d, and to LOD2.3—the walls are offset by a distance d, preserving the roof edges, and resulting in a LOD2 model with explicitly modelled roof overhangs of a predetermined fixed distance. In areas where buildings with

Table 5.2: List of horizontal geometric references (the footprint of LOD1 and LOD2 models). The equal sign means that the reference is re-used from INSPIRE.

Code (§5.3.4)	Footprint at	INSPIRE ref.	Our ref.
F0	Actual location	`footprint`	=
Fd	Roof edges offset by a fixed distance	N/A	`offsetRoofEdge`
F1	Roof edges	`roofEdge`	=

overhangs are predominant and the value of d is close to the average size, such a straightforward practice may provide models of a higher quality.

We use the code Fd to describe this reference, and in our experiments we use the value d = 20 cm as in [584]. It is important to note that this transformation is employed on all buildings, including the ones with no overhangs, potentially resulting in a smaller footprint than it is in reality. This reference is not discussed in the INSPIRE Building model.

On top of these three references, which in our experience cover all the models found in practice, the INSPIRE Building model defines three additional references `aboveGroundEnvelope`, `envelope`, and `lowestFloorAboveGround`, which define footprints for special cases of buildings and models, such as taking into account the underground structure when it is larger than the horizontal extent of the building above the surface. Because we have not encountered such cases in practice, we do not include them in our work.

Similarly as in the previous table, in Table 5.2 we give the list of horizontal references for the footprint of a building.

Figure 5.5 demonstrates an example of one building acquired with the considered representations. Observe the two instances of LOD2: the LOD2.1, and LOD2.3 with the explicitly available roof overhangs. In the latter, the roof edge is always represented at its actual location, but the body of the building varies.

5.4 CONCLUSIONS AND RECOMMENDATIONS

In this chapter we have researched geometric references in 3D building models, an important but frequently overlooked concept of the multiple representations of 3D city models that are of the same LOD. We have examined geometric references in LOD1 and LOD2 models that appear frequently in practice. In total, there is more than a hundred of ways to realise the LODs introduced in Chapter 4, indicating that this aspect is crucial when specifying 3D city models. Hence, we will revisit this

Figure 5.5: An example of a building modelled in 27 representations. The block models are 21 variants of LOD1.2 (3 horizontal references × 7 vertical references). The three models lined next to the LOD1.2 instances are three variants of LOD2.1 models, differentiated by the horizontal reference. The two models top left are LOD2.3 with references F0 and Fd (LOD2.3 with F1 is not possible). The bottom left is the LOD3.3 model for reference.

topic in Chapter 6 in which we generate a dataset with all combinations of LODs and geometric references. In Chapter 7 we demonstrate the generation of 3D models from non-elevation data such as the number of storeys (reference Hx). The topic of geometric references is further covered in Chapter 10 in which 3D city models acquired in the same LODs but with different geometric references is used in a spatial analysis to compare the results.

Finally, to conclude this chapter, we give recommendations related to the adoption and utilisation of geometric references, and discuss a few important points.

5.4.1 Extension of the INSPIRE Building model

The INSPIRE Building model provides extensive metadata for the vertical and horizontal geometric references, but our research has suggested that they are insufficient. We propose the following:

1. Supplementing the standard with additional references found in our research (Section 5.3).

2. Eliminating the ambiguous reference `generalRoof`, which indicates that the vertical reference may represent any point of the roof.

3. Enabling additional metadata on the references. For example, in the case of the horizontal reference Fd, we propose enabling the notation of the offset; and in case of the reference Hx, we deem that it would be beneficial to state the lineage of the data that is used to derive the extruded models. This is

especially beneficial for the increasing number of models derived by extruding footprints coupled with various information from cadastral datasets.

5.4.2 Extension of CityGML

CityGML does not provide a mechanism to store the geometric reference metadata, resulting in uncertainty and unknown lineage of the models. Therefore, we propose to extend the standard with INSPIRE metadata, and we have submitted a change request to the OGC to regard these geometric references (for more details see our paper [607]). In this conclusion we focus on two points of discussion that can aid the developers of the standard:

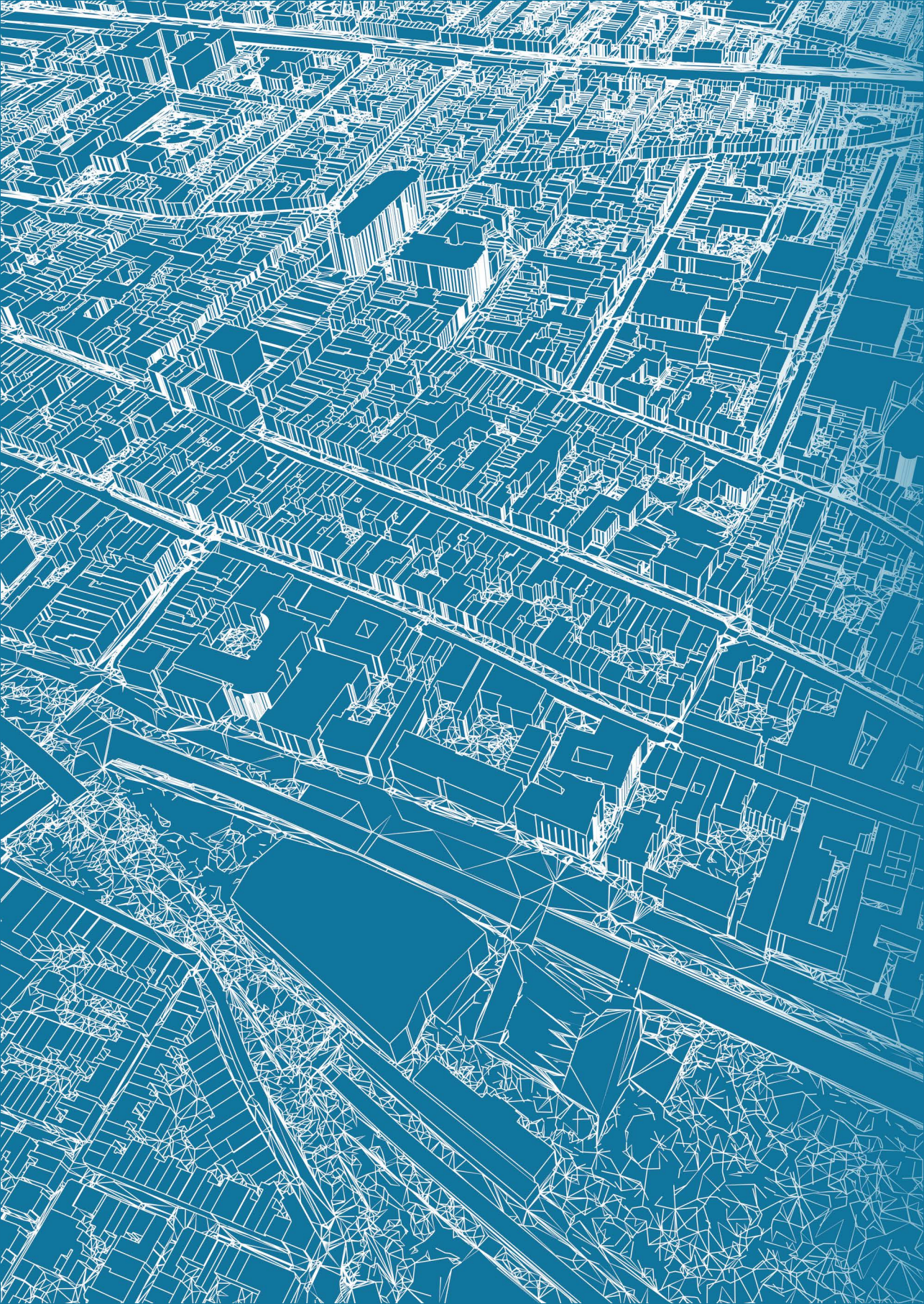
Cardinality of the representations The current version of the CityGML standard does not support storing multiple representations of the same LOD. Consequently, it is not possible to store two LOD1 models with different geometric variants (e.g. an LOD1.1 with F0/H3 and an LOD1.1 with F1/H0). We encourage the developers of the standard to take this into account, since each representation provides a different value for a spatial analysis. Hence, enabling the possibility of storing multiple models of different GRs might enable practitioners to switch the GRs and select the most suitable one.

Granularity of the metadata Nowadays, most of 3D city models are acquired with a consistent workflow, i.e. a city is surveyed by one party using one acquisition approach. This results in the geometric reference of buildings to be homogeneous across a dataset.

However, an increasing number of 3D GIS datasets contains models of heterogeneous lineage, e.g. see the work of Over et al. [549], Goetz and Zipf [627], and Goetz [405]. In contrast to most models, this approach potentially results in different geometric references in the dataset. For this reason, we argue that it is essential to provide the metadata on the building level, rather than on the dataset level.

5.4.3 Integration in quality control procedures

The geometric reference is usually not considered in spatial data quality documents, e.g. ISO 19157 [628]. We recommend the developers of quality standards to regard this important concept, by enabling the assessment of the geometric reference in the quality procedures, e.g. to express that the geometric reference in the dataset does not correspond to the one noted in the metadata.



II

Generation and management

Delft, The Netherlands. Generated with 3dfier using BGT and AHN datasets.

CHAPTER 6

Realisation of the specification

This chapter is based on my paper [629]:

> Biljecki F, Ledoux H, Stoter J (2016): Generation of multi-LOD 3D city models in CityGML with the procedural modelling engine Random3Dcity. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, IV-4/W1: 51–59. doi: 10.5194/isprs-annals-IV-4-W1-51-2016

Part I of the thesis proposed a comprehensive specification for LOD and geometric references. The goal of this chapter is to realise it by generating a compliant dataset. Such a dataset is essential for the subsequent chapters that involve experimenting with data in multiple LODs. The chapter first seeks to find a feasible method for generating data by revisiting Section 2.2, which overviews the acquisition techniques of 3D city models. Procedural modelling is identified as the most suitable approach for this purpose, and related work in this domain is presented. The chapter introduces an experimental procedural modelling engine developed during the PhD research in order to support the experimental portion of the research. The engine is designed to produce models in CityGML and does so in multiple LODs. Besides the generation of multiple geometric LODs, we implement the realisation of multiple levels of spatio-semantic coherence and geometric reference variants. As a result of their permutations, each building can be generated in 392 different CityGML representations, an unprecedented number of modelling variants of the same feature. The engine was also found to be of interest to other researchers, thus other uses of the software are showcased.

6.1 INTRODUCTION

The aim of this chapter is to realise the LOD and geometric references (GR) specification described in the previous two chapters in a sample dataset in CityGML 2.0. The LODs and GRs can be generated with a multitude of approaches (Section 2.2), e.g. an airborne laser scanning survey can result in both block models LOD1.2 and models with detailed rooftops (LOD2.2). However, most of those techniques are not feasible in the scope of this research, especially the ones resulting in LOD3, because of limited resources. The goal of this chapter is to find a suitable technique for generating 3D city models according to the new LOD specification, and to investigate how to generate multiple LODs in a feasible manner. In Section 6.2 we present related work and we justify procedural modelling as a potentially suitable acquisition technique to derive such data, which despite their synthetic nature may still be suited for various applications and experiments. In Section 6.3, we introduce Random3Dcity, an experimental procedural modelling engine developed in this PhD research to generate buildings in multiple LODs in the CityGML format, and according to our specification introduced in Chapter 4 and Chapter 5. It is the first engine of this kind, and we have released the code open-source for free public use. The engine is composed of two modules and it has been built entirely from scratch with a custom shape grammar. The datasets generated by this engine have already been used in several research projects (Section 6.4).

6.2 BACKGROUND AND RELATED WORK

6.2.1 Motivation for multi-LOD models

There are situations and use cases in which having multi-LOD data is useful. For example, (1) visualisation applications and spatial analyses in which, similarly to computer graphics, multiple LODs are switched for efficient visualisation and streaming, to increase cognition, and to decrease computational complexity [398, 568]. For example, when inspecting a large area it may be more beneficial to deal with a coarser LOD, or in certain stages in urban planning it may be detrimental to use fine LODs hence it is beneficial to use multiple representations depending on a purpose [494, 630]; (2) as a source of data for testing software implementations that are focused on structuring multi-scale data, e.g. compression methods; (3) to enable data producers to differentiate products by stripping down a high-quality model attracting different segments of the market (cf. quality discrimination and product versioning [631]), or for dissemination (e.g. city has an LOD2, but for some reason does not want to disseminate it, so it creates an LOD1); and (4) for assessing the suitability of a specific LOD prior to tendering and data acquisition ('LOD bench-

marking') and for serving different applications (e.g. noise estimation models only need LOD1 while a solar potential analysis may need finer detail)[18]; (5) intentionally creating a coarser LOD to increase privacy [632]; (6) in cases where two LODs come from different sources (e.g. see the example in Figure 3.1) that cannot be integrated, but it is beneficial to retain both since each may have its advantages[19]; and because (7) some software packages are limited to lower LODs, so multiple representations are generated. Software implementations for computational fluid dynamics and predicting the propagation of noise are such a case because they are frequently constrained to LOD1 representations.

6.2.2 Overview of acquisition workflows

In Section 2.2 we have analysed 3D production workflows and accompanying software support (see also Figure 2.2 for the taxonomy). We revisit the developed taxonomy in order to discuss an approach to facilitate the production of multi-LOD data. There are four categories of acquisition techniques: (1) direct acquisition; (2) reduction; (3) augmentation; and (4) design. When generating data in multiple LODs, the first approach can be laborious and expensive, and it is hindered by software limitations. Furthermore, the data is burdened with different levels of acquisition errors and it may include other inconsistencies such as invalid geometries, an unwanted but common outcome of 3D acquisition [227, 633]. Second, obtaining multi-LOD model with generalisation is viable only in theory because of the lack of implemented solutions, especially those that support CityGML. Furthermore, in this approach a dataset at fine LOD is required, which are in practice usually available for only a limited set of buildings. Third, as in generalisation, augmentation of a dataset results in a multi-LOD dataset. However, its capabilities are limited, require real-world data, and there is a lack of software implementations to use. The remaining group are design models, on which we focus in the continuation.

6.2.3 Designing data with procedural modelling

In this chapter we focus on procedural modelling, as a common augmentation approach and source of design models, but not previously considered as a source of

[18] For example, when planning the procurement of 3D city models for a specific application, it would be beneficial to have a sample dataset in multiple LODs, run a spatial analysis for each, and compare the accuracy of the results to their cost of acquisition. Such performance analysis can then serve as a decision factor for determining the optimal LOD to be acquired, prior to acquisition and to avoid procuring data of an unsuitable LOD (either too fine or too coarse). In this chapter we address the absence of such experimental ('pilot') datasets, and we use them for that purpose in Part III.

[19] Another example is having a 3D city model of coarse detail but high quality, and another dataset of a fine detail but poor quality. Such datasets will be further discussed and showcased in Chapter 12.

multi-LOD data. Procedural modelling involves generating 3D city models from scratch or based on an existing 2D dataset by using a set of rules [405, 634]. It is an important topic in computer graphics and GIS, and there have been several initiatives to develop procedural engines for modelling urban features. For example, buildings [49, 168, 635, 636], landmarks [637], roads [638], plants [639], monuments [640], and land parcels [641]. For a comprehensive list see the overview in Smelik et al. [169].

Procedurally generated models have proven to be an efficient technique for generating 3D models, either on its own without other data or by enhancing existing data—e.g. adding fenestration to an LOD2 model [71, 359, 642]. They find their use in 3D GIS in an increasing range of applications, e.g. in urban planning and simulation [643–646].

An example workflow is to take block models of buildings, and to *synthetically* augment their detail and appearance (e.g. adding a roof shape and textured façade) according to a predefined architecture typical for that spatial extent, resulting in a 3D city model at a notably finer LOD without a significant additional cost. Another example is to take building footprints as input, e.g. detected from a point cloud [647], and to generate buildings on top of it.

An advantage of this technique is that it is a quick and straightforward method to generate 3D city models, usually in large quantities. Furthermore, the models derived with procedural modelling are fine in detail [637], and because of their nature they tend to contain less topological inconsistencies (e.g. models obtained with automatic reconstruction from lidar point clouds are more susceptible to topological errors).

As a disadvantage, due to their generative nature, procedurally modelled datasets are not accurate from the GIS point of view: in fact, they may considerably deviate from the reality they purport to represent [502]. This is due to the primary goals of procedural modelling: to quickly generate 3D data, and to increase the LOD of existing models to improve their visual impression, achieved by *artificially* adding features (essentially embellishing existing models). Nevertheless, this inconsistency does not interfere with many applications, such as flight simulation, and gaming, where the focus is on visualisation, rather than on spatial analyses of real-world data [145]. As a result, many synthetic datasets have been generated completely from scratch, representing fictitious settings, with movies being such a case.

At the moment, one of the most prominent procedural modelling engines in the GIS world is Esri's CityEngine, used by urban planners and other practitioners. However, so far procedural modelling efforts have not been focused extensively on multi-scale representations and CityGML. Considering that a procedural engine can be programmed to produce data with a specific granularity, we take advantage of this idea in our work by defining different procedures to generate buildings in a series of

different representations based on the specification we defined in earlier chapters.

6.3 PROCEDURAL MODELLING ENGINE RANDOM3DCITY

In this section we present a CityGML compliant procedural modelling engine developed specifically to produce models according to our specification and in multiple LODs. We have formulated and developed a custom methodology, shape grammar, and rules that can be modified to suit the requirements of a user.

Besides the motivation of tackling the absence of multi-LOD datasets and shortage of diverse sample CityGML data, we have created Random3Dcity for other reasons. For example, to address the lack of CityGML procedural modelling software to easily create 3D city models in the respective format. The engine has two functions: generating an unlimited number of synthetic models that mimic a real-world setting, completely from scratch, and augmenting existing datasets.

The software prototype is composed of two independent and extensible modules (see Figure 6.1 for the workflow). The first module consists of a customisable set of rules that derives the configuration of the architecture in a parametric description encoded in an Extensible Markup Language (XML) format (Section 6.3.1). This means that the architecture and rules can be adapted to a specific setting. For instance, different rules can be encoded, such as to imitate a residential area with buildings that are between 3 and 6 storeys high and have predominantly flat roofs.

The second module reads the generated parametric data, and realises the parametric description of buildings as 3D city models in CityGML in multiple representations (Section 6.3.2).

Besides generating each building in multiple representations distinguished by geometric complexity (Chapter 4), the engine generates models in multiple geometric references (Chapter 5); in multiple levels of semantic structuring (e.g. LOD3.2 with and without the thematically enriched surfaces); with the difference in geometric type (boundary representations and solids); and their corresponding indoor geometry further in multiple LODs (indoor is not in focus of this thesis, but a couple of indoor LODs has been implemented to support future research). Permuting all these combinations results in 392 representations of the same building. To the extent of our knowledge, the models obtained with Random3Dcity present the most complete CityGML (and probably in general 3D building) datasets available to date, contributing to a multitude of application domains (Section 6.4).

The engine also supports thematic features other than buildings, and the generation of a basic interior of buildings (Section 6.3.3).

Each module of the engine is independent, facilitating the extensibility and potential integration in future. For example, since the modelled data is first stored

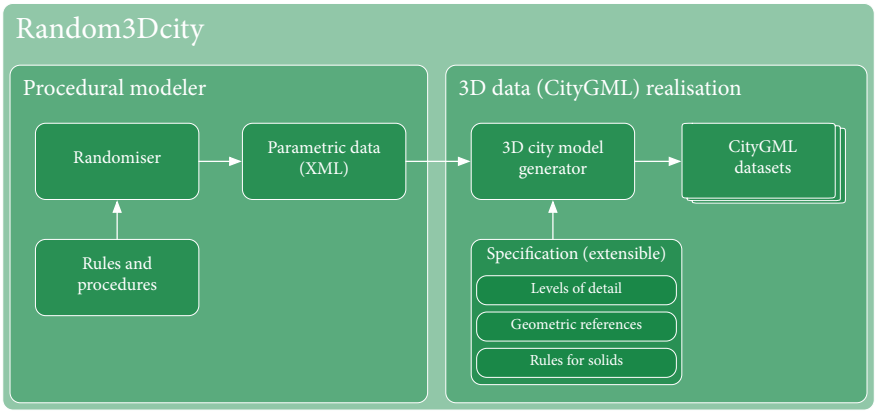


Figure 6.1: Modular workflow of the engine Random3Dcity. The architecture of the engine is composed of two parts: the procedural modeller resulting in a parametric description, and the generator of CityGML data from these parameters.

in a parametric form, it can be generated in formats other than CityGML. On the other hand, an open parametric building format brings two benefits: first, it facilitates the integration with existing data. This is demonstrated in Section 6.3.4 where we generate a 3D city model based on a real-world 2D cadastral dataset, similarly as in present-day commercial software. And second, the 3D data generator may be independently used to generate buildings derived from other sources with the same rules if stored in this form.

Figure 6.2 illustrates an example of the output of the engine: a setting with 100 synthetic buildings in four LODs that are randomly placed and rotated. All datasets (incl. the 388 others not shown here) have been generated in less than one minute.

The software has been implemented in Python, and it is available as open-source at Github. A set of generated test datasets is also available for public use on the website of the project.

6.3.1 Parametric description and rules

In the engine, each building and other real-world feature F_i consists of a set of n parameters p_i that define its architecture:

$$F_i = \{p_i^1, p_i^2, \dots p_i^n\}. \tag{6.1}$$

These parameters are in line with the ones used in photogrammetry (e.g. cf. [578,

Figure 6.2: Composite rendering of a few exemplary datasets generated by Random3Dcity. The example illustrates randomly generated buildings realised in CityGML in four LODs (1.2, 2.2, 2.3, and 3.3). The two representations on the left have their walls modelled as projections from roof edges as opposed to the third panel where the walls are modelled at their actual location and where roof overhangs are explicitly modelled, showing varying geometric references that are supported by the engine, besides multiple geometric and semantic LODs.

648, 649]), such as the width of a building, length of ridges and eaves, and roof height. The first part of the engine, described in this section, creates features described by these parameters, using an encoded grammar and a set of rules. The parameters are then stored in an XML schema that we have defined.

Our methodology of procedurally modelling the urban features consists of first defining a top-down hierarchy of parameters p_i. For example, before determining the number of windows on a wall, first the width and length of the footprint of the building are derived. Thanks to the hierarchical approach, the parameters are context-aware, i.e. the engine contains several procedural rules and constraints. For example, flat roofs cannot contain dormers, doors have to be located on the ground storey, buildings have to have at least one storey, and windows cannot be taller than the height of the storey. Second, a range for each parametre p_i is defined, e.g. the width of the window is between 0.5 and 1.5 m, from where the engine randomly samples a value according to a defined probability distribution function (also customisable). Third, the rules take care to generate a realistic setting, e.g. that multiple dormers are properly aligned on the roof. All these rules are stored in the code in

Figure 6.3: Support for different types of roofs. In this example 5 types (flat, gabled, hipped, pyramidal, and shed) are fitted on the same building body, with the automatically adjusted walls. This image illustrates the LOD2.3, and it also features an example of a building part.

a series of IF-THEN statements, and they are customisable to conform to a specific setting a user aims to (re)construct.

Random3Dcity supports five types of roofs, which are frequently described as the most common types of roofs [484, 577, 590, 650]. They are shown in Figure 6.3. Furthermore, the figure demonstrates that a building part such as a garage can also be generated.

In addition to the geometry and semantic representation, the engine is capable of generating a number of attributes, such as building age, number of storeys, use of building, which might be useful for some use cases. All together, these parameters are stored in an XML schema (Figure 6.4). This example reveals the underlying parametric description of the second building from the left in Figure 6.3.

6.3.2 CityGML realisation of the parametric building

The second part of our engine reads the generated parametric data p_i of each feature F_i and constructs CityGML 2.0 datasets in multiple LODs and GRs. The process of the construction of the geometric representation from the parametric representation is described in this section by highlighting a few important aspects.

6.3.2.1 Construction of the geometry and the CityGML file

Each of the representations is generated separately. For each, the engine reads only the set P_i of parameters p_i that it requires for constructing it. For example, for the LOD2, a set P_i^{LOD2} is defined and the engine fetches the size of the body of the building, roof type, and height of the roof, ignoring other parameters such as windows and dormers. These instructions are stored in the engine, and can be customised to create additional LODs outside the predefined specification, if required. For example, it is possible to generate an LOD3-like model that contains only roof openings, by just disabling the constructor for other features such as dormers. Different geometric references are realised in a similar manner (Figure 6.5).

```
<building ID="c209f43f-f137-4dd0-8816-cbc4dc4a407d">
    <footprint>Rectangular</footprint>
    <origin>173469.34 526427.95 0.0</origin>
    <rotation>34.3</rotation>
    <xSize>7.06</xSize>
    <ySize>8.91</ySize>
    <zSize>6.8</zSize>
    <floors>2</floors>
    <floorHeight>3.4</floorHeight>
    <embrasure>0.08</embrasure>
    <wallThickness>0.2</wallThickness>
    ...
    <buildingPart>
      <partType>Garage</partType>
      ...
    </buildingPart>
    <roof>
      <roofType>Gabled</roofType>
      <h>2.48</h>
      <overhangs>
        <xlength>0.5</xlength>
        <ylength>0.5</ylength>
        ...
```

Figure 6.4: Example of the building parameters stored in an XML.

In the construction of the geometry, the engine first creates a local Cartesian coordinate system X_i for the building F_i, and then a coordinate system X_{p_i} for each of its elements defined by p_i (e.g. wall). The vertices of the features are generated in this system (e.g. origin of the window on a wall), and the resulting surfaces are generated. For each different element class (e.g. chimney), an algorithm is designed for their geometric realisation from their parametric description. These 'sub-systems' X_{p_i} are then converted to the coordinate system X_i of the building, which is later converted in the global system defined by the user, depending on the location and orientation of the building in space.

The process of the generation of the geometry of the building elements is not equal for all buildings, it mostly depends on the type of the roof. That is, the vertices of the top of the walls are not equal for buildings with a flat and gabled roof (visible in Figure 6.3). Hence, separate algorithms are designed for each roof type.

In the process of the generation, geometries are given a universally unique identifier (UUID) according to [651], the recommended approach of CityGML and GML [652], and are then structured according to the semantic level of the representation. Furthermore, the attributes have been translated and stored according to the CityGML 2.0 standard. An excerpt of the CityGML data of the building exemplified through its parametric description in the previous section is given in Figure 6.6,

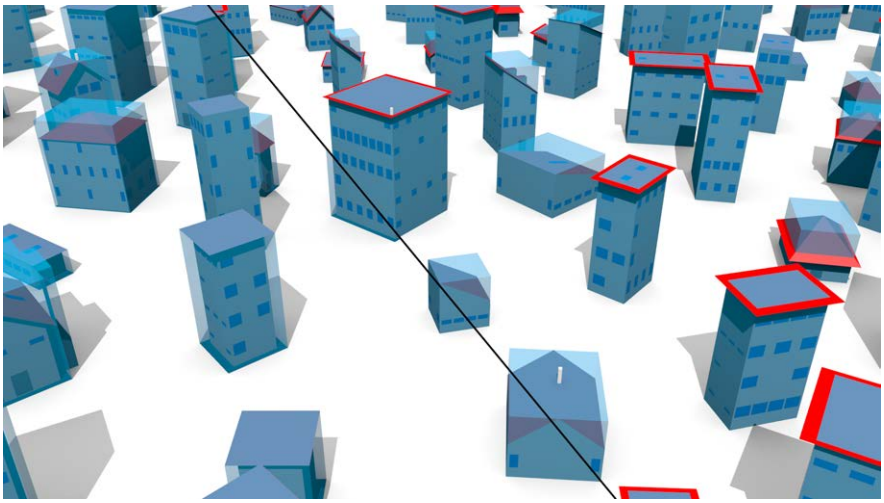


Figure 6.5: A composite rendering of three subsets of our procedurally generated datasets. Two LOD1.2 models with different geometric references for the footprint separated by the thick black diagonal (LOD1.2-H5-F1, left of the diagonal, and LOD1.2-H5-F0, right), superimposed on an LOD3.3 model. Note that where the building has no overhangs the models correspond. Observe that some of the LOD1 models deviate more than the others depending on their configuration (e.g. compare the building on the far left with the garage in comparison with a building that has a flat roof and no overhangs).

which also includes the realised attributes.

6.3.2.2 Generation of corresponding solids

Each model, with the exception of LOD0 models, is also stored as a `gml:Solid`. Solids are used to facilitate uses in application domains that require the usable volume of a building, such as for the estimation of property taxes [129], and the estimation of the energy demand of households [114].

In the construction of solids, features that do not contribute towards the usable volume of buildings are disregarded. This applies for instance to roof overhangs, and chimneys. In the `gml:MultiSurface` representation this is also regarded by the generation of a `ClosureSurface` to seal the open sides and to provide a representation as geometrically closed volume object, following the recommendation of Gröger and Plümer [408].

Figure 6.7 provides an example of the relationship between the semantic bound-

```
<cityObjectMember>
 <bldg:Building gml:id="c209f43f-f137-4dd0-8816-cbc4dc4a407d">
  <bldg:roofType>Gabled
   </bldg:roofType>
  <bldg:yearOfConstruction>2008
   </bldg:yearOfConstruction>
  <bldg:storeysAboveGround>2
   </bldg:storeysAboveGround>
  <bldg:boundedBy>
   <bldg:GroundSurface>
     ...
```

Figure 6.6: Excerpt of the generated CityGML 2.0 dataset. This is the output of the second module of the engine.

ary representation models and its solid counterpart. The solids generated by the engine have been geometrically validated according to the standard ISO 19107 [653] with the implementation of Ledoux [633].

6.3.3 Generation of indoor and non-building features

We have implemented multiple versions of the interior according to the refinement developed by [129]. Furthermore, thematic features other than buildings have been generated, such as vegetation and roads. Figure 6.8 illustrates an example with the interior of buildings (LOD2+ model as per [129]).

6.3.4 Augmenting existing data (2D footprints)

Besides generating all parameters of features from scratch, an experimental feature of the engine is to utilise existing GIS data such as footprints, and procedurally model the remaining features producing a 3D model. For example, Figure 6.9 illustrates a setting in which 2D footprints from an existing dataset have been used, and the 3D buildings have been procedurally modelled by specifying the architecture where only hipped roofs are present and where buildings must have large windows (as it is the case for that setting).

6.4 APPLICATIONS OF THE GENERATED DATASETS

Availability of freely available datasets with a large number of dissimilar buildings represented in multiple LODs opens a door for a multitude of research purposes. Another advantage is that the data is topologically correct and generated strictly according to CityGML guidelines. Instead of an experiment section in this chapter

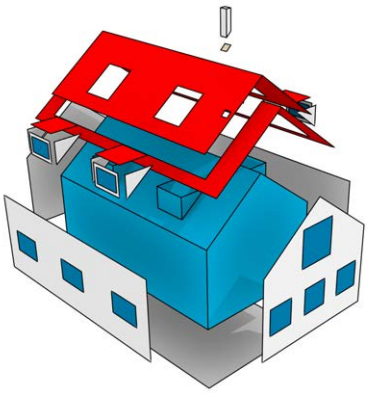


Figure 6.7: Two representations of LOD3 models. A `gml:Solid` and a thematically structured `gml:MultiSurface`. Note the `ClosureSurface` of the chimney (light brown) in order to separate the usable volume of a building.

(as the data generated by the engine will be used in Part III of the thesis), we showcase the uses of Random3Dcity from fellow researchers in the 3D GIS community. The generated datasets have already been tested and used in several application domains outside this research. For example, they have been used to test software that take CityGML data as input (e.g. viewers). A few notable cases follow:

1. Testing and improving CityGML validation software [512, 633, 654], primarily in the scope of the OGC CityGML Quality Interoperability Experiment (QIE) [655]. Considering that the engine creates topologically consistent datasets, they have been used as exemplary models of valid datasets. Furthermore, for this project the code of the engine was modified to intentionally produce data with topological errors, such as overlapping buildings and broken solids, which were used as input to test validation and repair software packages benchmarked in the QIE.

2. Optimising the coverage of geosensor networks [203, 656]. Researchers have used our synthetic datasets to test the implementation of their use case, which focuses on line of sight analysis.

3. As a data source in experiments with voxelisation of CityGML models to facilitate volume computation [657].

4. In an ongoing research as test datasets for simulating point clouds to benchmark the quality of surface reconstruction in multiple LODs—the obtained

Figure 6.8: A composite *street view* of a dataset generated by the engine. The image shows the vegetation (park), street network, and the basic indoor representation of buildings (a solid representing each storey).

3D model reconstructed from the simulated point cloud is compared to the original 3D city model generated by the engine [658].

5. In the ongoing research of Neuville et al. [659] the data is found useful for testing visibility scenarios, since it provides diverse cases of buildings.

6. In the research project of Tsigkanos et al. [660] the generated data was used to create an experimental setup bridging CityGML data and dynamic cyber-physical spaces, a new design of modern spatial environments from a software engineering perspective.

6.5 CONCLUSIONS

In this chapter we have realised the specification in Part I with procedural modelling, as opposed to orthodox techniques such as laser scanning and generalisation. In this way we also address the lack of procedural modelling engines that support CityGML, and the shortage of publicly available CityGML data. Furthermore, the chapter demonstrates that procedural modelling serves as a feasible choice of highly detailed (LOD3) models, since their acquisition with traditional techniques is expensive.

Figure 6.9: Procedural modelling from LOD0 to LOD3. Example of a CityGML model generated with Random3Dcity in conjunction with the existing real-world dataset of 2D footprints of buildings. This setting shows the *Beestenmarkt* in Delft, the Netherlands. The grammar has been adjusted to match the configuration of the buildings in that setting in order to resemble the reality as close as possible.

The experimental engine Random3Dcity, which we have introduced and built from scratch, is novel: it natively supports CityGML, and it is designed towards producing multi-LOD and multi-GR data. It yields an unprecedented number of variants of models, and does so according to an underlying set of customisable rules. The engine is open-source, and a set of example datasets is available for free public use. The reason why we have made this project open is that other researchers can benefit from multi-LOD datasets in their application domains, and that they could adjust its grammar to suit specific requirements. As a result, the generated datasets have already been proven useful by being featured in several research projects in different countries and application domains. Such interest suggests the need for open procedural modelling engines, and valid and diverse CityGML datasets. We invite other researchers to take advantage of the availability of these datasets to test their implementations and as a source for experiments.

The procedural aspect ensures that the data is realised strictly according to the LOD specification, but also it rigorously follows the CityGML guidelines. For example, it produces solids and b-rep, and the generated attributes are stored according the standard. Furthermore, the final two chapters of the thesis (Chapter 11 and

Chapter 12) demonstrate that procedural modelling can be leveraged for error propagation research. That is, a stochastic error engine is built as a third module, between the existing two modules of the software, simulating acquisition errors in the acquisition of 3D city models.

While the engine generates fictitious settings, it is still suited for applications where having real-world data is not important, and where different scenarios can be evaluated (e.g. to determine whether it is more beneficial to acquire an LOD2 instead of an LOD1 for a specific spatial analysis, by testing both representations before the actual acquisition; a topic of Chapter 10).

This experimental software does not aim to compete with advanced commercial solutions, such as Esri's CityEngine, which are capable of creating complex architecture. Nevertheless, it bridges the gap with respect to CityGML data and multiple representations to quickly obtain models suited for experiments and testing, and it is released as open-source.

For future work we plan to work in two directions. First, we plan to advance the shape grammar for generating more complex buildings, such as structures with less usual roof types and landmarks. This is a natural flow of the work, which is hampered by designing advanced rules that more complex features imply, such as complicated roof shapes and balconies. Second, we plan to increase the LOD of the interior. We have introduced the modelling of a basic indoor (storeys), and we intend to include the generation of rooms and openings (windows/doors), following recent efforts in procedural modelling of interior according to an indoor grammar [661–664].

CHAPTER 7

Generating 3D city models without elevation data

This chapter is based on my paper [665]:

> Biljecki F, Ledoux H, Stoter J (2017): Generating 3D city models without elevation data. *Computers, Environment and Urban Systems*, 64: 1–18. doi: 10.1016/j.compenvurbsys.2017.01.001

In this chapter the inverse approach to generalisation is further discussed, i.e. augmenting the detail of an existing dataset to enhance its LOD. Besides procedural modelling as discussed in the previous chapter, the most employed method to augment the LOD is extrusion—lifting the 2D dataset to the height of each feature to obtain CityGML LOD1 models. The principal ingredient needed to produce such block models via extrusion is the height of each urban feature. However, elevation datasets (e.g. point clouds) are often unavailable hindering extrusion. We investigate to what extent can 3D city models be generated solely from 2D data without elevation measurements. We demonstrate that it is possible to predict the height of buildings only from 2D data (their footprints and attributes available in volunteered geoinformation and cadastre), and then extrude their footprints to obtain 3D models suitable for a multitude of applications. The predictions have been carried out with machine learning techniques (random forests) using 10 different attributes and their combinations, which mirror different scenarios of completeness of real-world data. Some of the scenarios resulted in surprisingly good performance (given the circumstances): we have achieved a mean absolute error of 0.8 m in the inferred heights, which satisfies the accuracy recommendations of CityGML for LOD1 models and the needs of several GIS analyses. We demonstrate that our method can be used in practice to generate 3D city models where there are no elevation data, and to supplement existing datasets with 3D models of newly constructed buildings to facilitate rapid update and maintenance of data.

7.1 INTRODUCTION

In Chapter 2 several approaches to derive 3D city models have been overviewed. A taxonomy of acquisition techniques has been created, in which one category is to generate 3D data by augmenting the LOD of existing data. One of these approaches is procedural modelling, which was a subject in Chapter 6, e.g. to generate an LOD2 from LOD0 (Figure 6.9).

Another augmentation method is extrusion of footprints (LOD0) obtaining volumetric (LOD1) models. This chapter focuses on improving extrusion in an unconventional way—without elevation data, i.e. only using footprints, not involving often costly and infrequent acquisition procedures such as airborne laser scanning. This inherently results in the augmentation of the LOD without additional (3D) data—from LOD0 to LOD1.

The motivation is that while footprints are now widely available as open data from governments and volunteered geoinformation [666, 667], elevation datasets remain expensive and time-consuming to acquire, hindering the production and availability of 3D city models [668].

In addition, when elevation datasets are available, they may not always be suitable to generate 3D models. First, they may be outdated, resulting in a lesser completeness than footprints, which are fairly easy to update in contrast to point clouds, and are also usually produced at more frequent intervals. Second, their resolution and quality may not always be sufficiently adequate to produce 3D city models. For example, the popular Shuttle Radar Topography Mission (SRTM) offering free elevation data with worldwide coverage has been to a great extent taken advantage of in geosciences [669, 670]. However, such datasets are not suited for producing 3D city models, mostly because of their insufficient accuracy and coarse resolution (e.g. 30 m) [207, 671].

As the goal of this chapter is to explore and evaluate alternative ways to construct 3D city models in the absence of elevation data, it relates to the topic of geometric references in Chapter 5. In the work, a vertical geometric reference Hx was defined to describe LOD1 models derived with non-elevation data. Hence, this chapter exhibits such datasets.

In this chapter we first investigate what are the current alternative ways to obtain 3D models not involving elevation measurements (Section 7.2). A method we have commonly found is using the information of the number of storeys (floors, levels) of a building, which gives the impression of being a fairly good proxy for a building height. However, while several 3D datasets around the world have been constructed in such a way, this method was never evaluated, which we seek to accomplish in our work. Second, we investigate whether there are other *predictors* that hint at a building's height, and assess their usability with machine learning techniques (Sec-

Figure 7.1: Improving and refreshing existing 3D city models. Besides generating 3D models from scratch, our method can be used to maintain existing 3D datasets by mitigating outdated point clouds, i.e. generating 3D models of buildings constructed after the most recent acquisition of a point cloud. The supplemented buildings based on predicted heights from our approach are indicated in lighter colours in the updated 3D model on the right.[20]

tion 7.3). To train and test predictive models we use data from the Dutch authorities, and perform a proof of concept on the cities of Rotterdam (Section 7.4), and Leeuwarden (Section 7.5).

While a 3D building model derived with such an unorthodox method would obviously not be the most accurate one, we argue that it not only can serve as a provisional solution until an elevation dataset becomes available, but it can in fact be useful to carry out several spatial analyses. Hence the chapter also discusses the usability of such datasets, and performs experiments to demonstrate that they give a good indication of the urban morphology, which is useful for various applications (Section 7.6).

To give a hint of the possibilities of the work we provide a use case that benefits from the work we developed: our approach can be used to supplement existing 3D models by updating newer buildings for which footprints (LOD0) are available, but a lidar survey has not yet been carried out (Figure 7.1). This example illustrates 2D footprints of several buildings (left), but missing from the elevation data, since they were built shortly after the lidar campaign, which is not conducted as frequently as updates to the 2D cadastral database. Such a situation results in a 3D city model (LOD1) with the omitted buildings, since the footprint cannot be extruded (centre). Using our approach the heights of the missing buildings were inferred from 2D data and the local context, so that the footprints can be extruded to a predicted height approaching reality quite well to generate a provisional 3D dataset with the improved completeness (right). An example with a 3D model generated from scratch will be

[20] This example was generated using the following data sources: BGT, AHN, and the City of Rotterdam.

given later.

7.2 EXISTING UNCONVENTIONAL METHODS TO INFER THE HEIGHT OF BUILDINGS

The methods we have found to derive the height of buildings for producing 3D city models without elevation data are inherently different and require different data sources, hence we group them into three categories.

7.2.1 Using attributes (number of storeys)

Many 2D building datasets contain the information of the number of storeys of buildings. Because the number of storeys is generally accepted to be a self-evident proxy for the building height, it has been used for 3D city model generation, especially in volunteered geoinformation, e.g. OpenStreetMap [405, 509, 672–674]. In fact, we have found a substantial number of occurrences of 3D models generated by simply extruding a footprint to the height obtained by multiplying the number of storeys with an assumed storey height [7, 73, 106, 119, 141, 184, 206, 290, 675–680]. The storey height appears to vary among papers, ranging from 2.8 and 3.5 m [107, 250, 290, 681]. Models constructed in this way have proven useful in a variety of applications, such as visibility analyses and energy simulations [119, 677]. To a lesser extent, some of the obtained 3D models are used for interactive querying, e.g. browsing 3D cadastre registrations [682], rather than 3D models of reality for spatial analyses. Furthermore, the number of storeys has been used as an indication for the building's height in research where no 3D models are reconstructed, e.g. in investigating the relationship between price of construction and building height [683], meaning that it is widely used also outside 3D GIS.

An advantage of this method is that the number of storeys is available from open data of governments and volunteered geoinformation in many places around the world [122, 329, 684]. Alternatively, the number of storeys has been obtained from a quick visual inspection of buildings, using terrestrial or airborne imagery [493, 668, 685, 686].

A common observation about these papers is that they are reserved about the details: the generation of (accurate) 3D models is not their principal topic (it is merely described in a sentence or two), and none evaluates the quality of the predicted heights. A somewhat exception is the work of Over et al. [549] stating that 'the computation of building heights using the number of floors alone is not a reliable method to measure the height of a building'. Unfortunately an error metric is not directly given nor a quality analysis, but the paper mentions that the error is outside

the permitted margin for block models according to the CityGML standard (we will revisit this matter in Section 7.3.4).

The absence of quality analyses of 3D models derived in this way is unusual considering their amount, and taking into account that a potentially large uncertainty in the 3D data may substantially degrade the quality of a spatial analysis (this will become evident in our experiments in Chapter 11).

Number of storeys is the only characteristic associated with a building that we have found being used for this purpose. In our study we evaluate the accuracy of the 3D building models generated by using the number of storeys, and we investigate if several other characteristics can be useful in determining the height of the building for extruding its footprint.

7.2.2 Using local regulations

Another way to predict the height of buildings is to use maximum allowed height values as prescribed in building regulations. Many jurisdictions around the world impose building height restrictions for: (1) aesthetic reasons, (2) to maintain environmental quality, (3) to prevent increased traffic congestion, and (4) to limit the strain on urban infrastructure [687, 688]. In urban and densely populated areas with scarce land, height restrictions are usually exploited to the last centimeter, hence it is reasonable to assume that in such cases the height of most buildings corresponds to the maximum permitted height.

This reasoning has been capitalised on in Singapore by Chen and Norford [6][21]. Their method relies on the data on the maximum permitted floor area ratio (FAR) from the local master plan—the ratio of a building's total floor area and the area of land parcel on which it sits. Regulating the FAR effectively limits a building's number of storeys rather than directly the height [687], which can then be indirectly used to estimate the height, using the method described in the previous section.

A validation was performed against the number of storeys obtained with a visual inspection. On average the method overestimated the number of storeys by 20%, mostly due to the mix of building heights within a plot (an area with the FAR regulation), since the restrictions apply to multiple buildings in a plot.

Additional papers relevant in this context are the one of Brasebin et al. [689], which similarly proposes the generation of 3D buildings from footprints using different local rules and constraints; and the one of Allani-Bouhoula and Perrin [690], which presents a method to generate 3D buildings from local regulation and various architectural principles. The latter work is also interesting to mention because it attempts to deduce the shape of the roof from the same set of information.

[21] The resulting dataset is the one shown in Figure 1.1 in the Introduction.

We did not replicate these methods because in the Netherlands there is no corresponding dataset, but since in our work we examine the usability of the number of storeys and floor area we indirectly evaluate them as well.

7.2.3 Using sun ephemeris (shadows in imagery)

Shadows cast by buildings have been used as an indication of its height, with a large number of papers on this mature topic in remote sensing [691–694].

The height of a building can be estimated from the length of its shadow and solar altitude at a given latitude and the date and time at which the image was captured [695]. In contrast to the previously described method using the number of storeys, the quality of such predictions is often reported. For example, Lee and Kim [696] report an accuracy of 3 m, Cheng and Thiel [697] outlines 3.7 m, while Shao et al. [698] achieves an accuracy of 13 m in estimating the height of high-rise buildings. As main challenges, researchers report overlapping shadows, measuring them on slopped terrain, and the interference of vegetation [699].

This unconventional method continues to capture interest and has current uses [50]. For example, Tong et al. [700] and Tu et al. [701] detect partially and totally collapsed buildings after an earthquake by comparing the height of a building from a pre-disaster database with the one obtained from shadows from post-disaster imagery.

Because this method has been investigated to a great extent, and because of the nature of the data (we use vector building footprints and not imagery), we do not focus on it. Furthermore, this method requires satellite imagery, which suffers from the same problems as point clouds: acquisition and the fact that the most recent dataset available may be outdated.

7.2.4 Conclusions from the literature review

As it can be concluded from the literature review, common observations of these specific methods to predict the heights of buildings from non-elevation data are: (1) they are using only one predictor of the height; (2) not all of them are focused towards producing accurate 3D city models; and (3) comprehensive quality analyses are seldom performed, and when they have been performed—they are limited to smaller areas such as a neighbourhood, not investigating how the solution scales to larger area such as a city. In the subsequent sections we seek to overcome these limitations.

7.3 METHODOLOGY AND DATA

7.3.1 Overview and considerations

We have developed supervised learning models using different attributes of buildings (predictors) to estimate their heights to generate 3D city models in LOD1. Supervised learning involves training data to develop the predictive model. Hence for this purpose we use a subset of buildings in our study area.

Besides discussing the usability of the predictors alone, we combine different predictors, envisaging different scenarios where various combinations of attributes are available and to investigate the importance of each in combination with the other. This is also important in our method, because in our dataset we did not have all the attributes for all buildings. For instance, the information on the building use was not available for all buildings. Hence different predictive models can be used depending on what attributes are available for each building.

When developing the method, the following use cases are foreseen:

1. In case a 3D model of a city is not available: measuring the heights of only a small subset of a city and applying the inferred predictive relationships to the unmeasured buildings. This involves having heights of certain buildings, but they may be available from an old point cloud, from the cadastral database, or by measuring a small subset of a city, which can also be done from ground (e.g. with a total station). We focus on this case when developing and validating the method (Sections 7.3 and 7.4).

2. In the case of having an outdated 3D model (and outdated point cloud) which has to be supplemented, we can analyse patterns from existing buildings and infer the heights of the new ones, built after the elevation data has been obtained (the case shown in Figure 7.1).

3. Inferring the relationships from one city where the elevation data is available, and applying them to another city where elevation data is not available. We have run experiments in another city to investigate this possibility (Section 7.5).

We have used Random Forests (RF), a supervised learning method for classification and regression that works by creating a number of decision trees on random subsets of data, and uses averaging to improve the predictive accuracy and control over-fitting [702]. It has been used in remote sensing and GIS for various purposes [703], e.g. for classifying movement trajectories [704], for assessing fire risks [705], and for identifying the typology of buildings [666].

One of the method's strong points is that it can assess the importance of each used predictor (feature importance) [706]. This property is useful because it provides the ranking the importance of predictors in the regression, enabling designing predictive models by choosing only predictors that are important, minimising their amount.

In our experiments we have also evaluated Support Vector Machines (SVM) and Multiple Linear Regression (MLR). However, the results obtained with SVM and MLR have not been as good as with the ones achieved with RF. Hence in this chapter we focus on RF.

For the implementation, we have used Scikit-learn, an open-source Python module for machine learning [707].

7.3.2 Data and study area

Rotterdam is the second largest city in the Netherlands, and it contains a variation of more and less urbanised areas with large differences of building heights. Hence it is a good option for a case study. The city covers an area of 326 km^2, and has a population of 620 000 people. The extent of the municipality covers a sizeable industrial area, that is, it hosts Europe's busiest port (Port of Rotterdam), adding to the diversity of the analysed structures.

2D footprints and their attributes We have obtained the 2D dataset of building footprints from the City of Rotterdam. The geometry of the buildings corresponds to the LOD0.2 requirements of footprints[22] defined in the specification in Chapter 4. However, because the top surface is not present, the dataset is LOD0.1, and by extruding it we obtain LOD1.2 (this also demonstrates that the specification is not 'vertical'; see Figure 4.1). The geometric reference of the footprints is at the roof edge.

For each building we have the following attributes: (1) building use, (2) year of construction, (3) number of storeys above ground, and (4) the net internal area (sum of floor area in all units in a building). Researchers in several countries report the availability of these attributes from cadastral datasets [67, 90, 119, 120, 373, 416, 507, 567, 684, 709–712]. Furthermore, some of these attributes may be derived automatically and may also be available in volunteered geoinformation [327, 331, 666, 713], enabling a wider applicability of the method presented in this work.

Geometric properties On top of these attributes we investigate the following metrics derived from the geometry: (5) footprint area, (6) shape complexity, and (7) number

[22] For further reading about the specification of the cadastral database the reader is referred to [708].

of neighbouring objects. We will elaborate on them in the continuation, and explain the rationale behind their selection.

Statistical data In addition, we have obtained census data from Statistics Netherlands (CBS), a possibly useful piece of information to predict heights. Rotterdam is divided into 92 statistical neighbourhoods, and for each we have access to the following potential predictors of the height of buildings: (8) population density, (9) average household size, and (10) average income.

Completeness Our dataset contains 200 000 buildings. However, not all of the buildings contain all attributes. For example, about a third of the buildings does not contain the number of storeys, use, and net internal area; these were usually smaller buildings without a cadastral registration, such as sheds. This reflects real-world situations of varying completeness of data, hence it is important to investigate the performance of different kinds of attributes and different combinations of availability. Figure 7.2 gives an illustration of the study area and the datasets.

Heights for validation and ground truth 3D model For training the predictive models and for validating the results we have used a point cloud provided by the City of Rotterdam, from which we have calculated the ground truth heights of buildings.

The methodology (illustrated in Figure 7.3) is based on calculating the elevation of the bottom (ground) and the elevation of the roof of the building. The former is calculated as the 5th percentile of the elevations of the points that fall within the buffer of the footprint, while the latter has been determined as the 90th percentile of the points within the building footprint. This particular percentile value for the top of the building was taken in order to filter out not only outliers but also chimneys, antennas, and similar constructions (similar values have been used in related work, e.g. in [623]). The difference between the two elevations is taken as the reference height of a building. This method has been realised with our open-source software 3dfier[23].

For training we use a relatively small subset (10%) of randomly selected buildings, and validate the performance with the remaining buildings (90%). In order to minimise inconsistencies in this reference dataset, we have filtered out a minor number of buildings with obviously erroneous values, e.g. with the measured height lower than 2 m; these were caused by cases such as the buffer overlapping with other constructions.

[23] Available at https://github.com/tudelft3d/3dfier

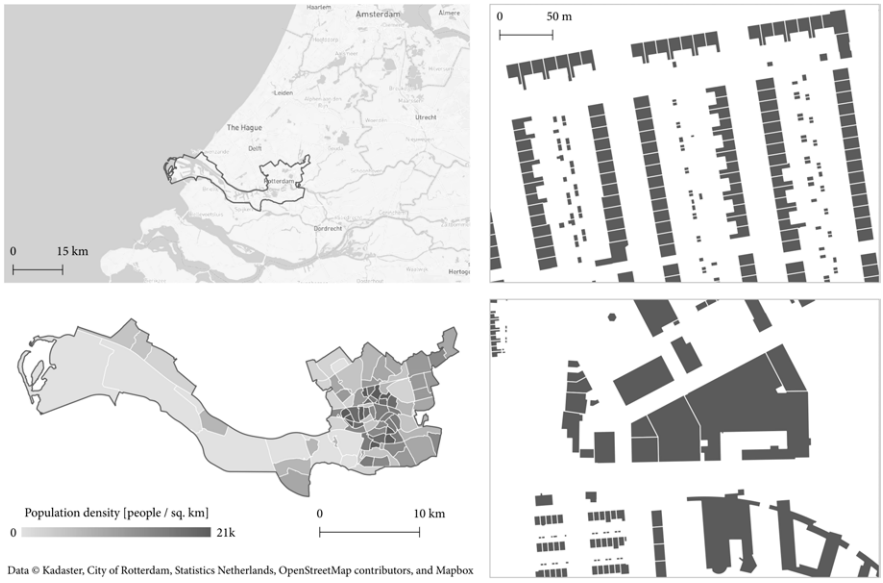


Figure 7.2: Some of the datasets used. Illustration of the study area (top left), the census neighbourhoods (bottom left), and the dataset of 2D buildings (right). Can you guesstimate the height of most buildings just by looking at the geometry of their footprints and neighbourhood?

7.3.3 Overview of the predictors

7.3.3.1 Building attributes (cadastre)

While using the *number of storeys* and multiplying it with an assumed floor height is straightforward, it is prone to many errors. Most importantly:

- Storey height differs between buildings due to different ceiling heights and slab thicknesses [714]. For example, a church has only one storey but it can be more than 25 times taller than a single-storey residential house. Furthermore, a building may have variable ceiling heights, e.g. a lobby may be considerably taller than the upper storeys.
- The building height that needs to be derived in our research is the overground height (from the base to the rooftop), excluding possible underground structures such as garages. The data on the number of storeys may include the total number of storeys, instead of only storeys above ground, resulting in overes-

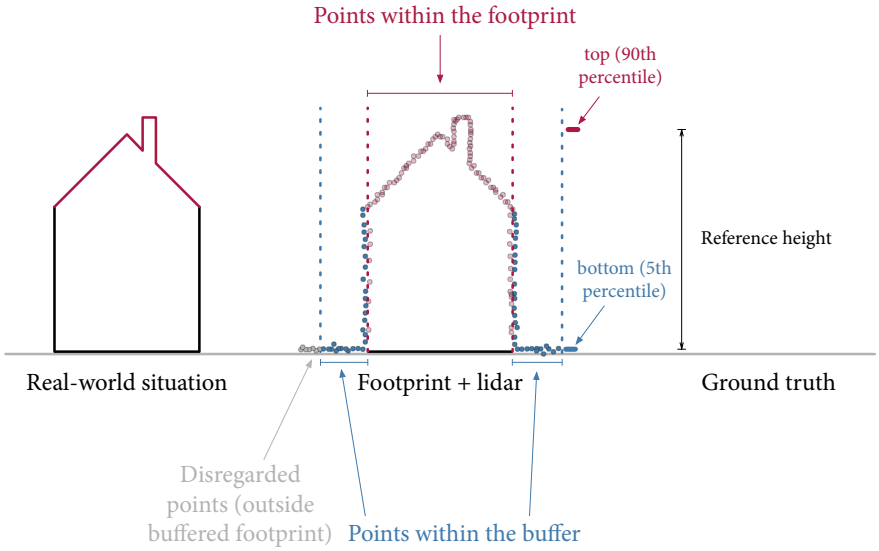


Figure 7.3: Anatomy of the approach to determine the reference (ground truth) heights of buildings. For each building, the software analyses separately two sets of points: those that are within the footprint (red), and those that are within the buffer of the footprint (blue).

timations. In our dataset we conveniently have the number of storeys above ground. But there are many datasets rather containing the total number of storeys.

- Buildings on sloped terrain may also contain an ambiguous number of storeys above ground (one side of the building may have a different number of overground storeys than the other side). While another convenience is that we are dealing with the flat topography of the Netherlands, we recognise that this could be a problem in other areas.

Furthermore, the complexity of the derivation of the height is indicated also by researchers in 3D geoinformation who work on the inverse problem: deriving the number of storeys from the height citing similar issues [129, 130, 715].

To give a general impression of the data on the number of storeys, Figure 7.4 indicates the relationship between the number of storeys of a building and its height. While there is a strong correlation, the plot also shows that there is a lot of variation in the heights of buildings with the same number of storeys. This is especially the case with buildings with one storey.

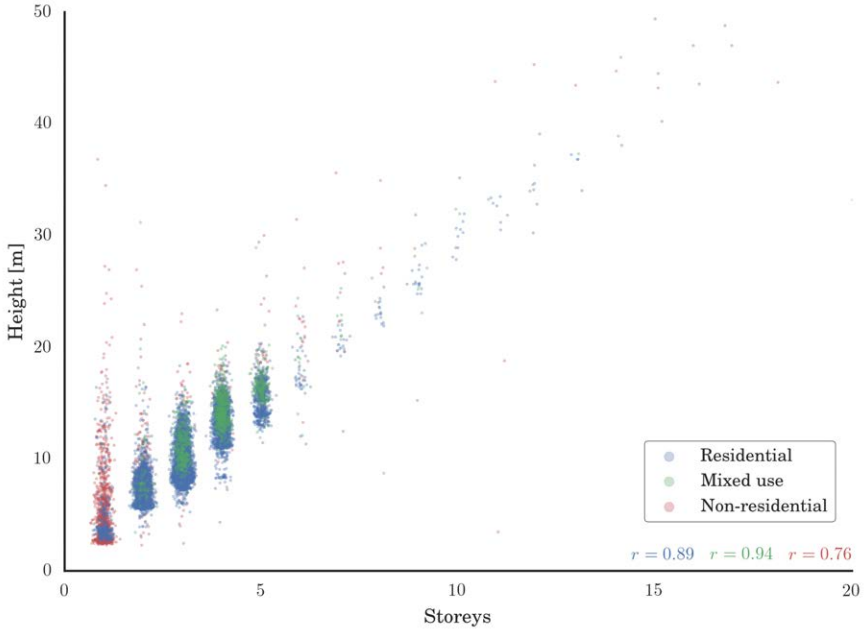


Figure 7.4: The relationship between the number of storeys and building height, coloured by building use. The number of storeys are integers, but jitter is added to expose the variation. The r values are correlation coefficients for each subset according to the building use.

The observations are coloured according to the *use of the building*, which gives more insight in the patterns. For example, it indicates that residential buildings have a slightly more consistent pattern than non-residential registrations. Hence we test if, along the information on the number of storeys, the building use may improve the predictions, as it influences the storey height.

Another predictor that should come in handy here is the *age of a building*: older buildings tend to have taller storey heights (Figure 7.5), so we will use this attribute to tackle the problem of the varying storey heights between buildings.

Before carrying out any experiments, it is obvious that the number of storeys appears to be a very useful predictor of the building's height, but it can be problematic to obtain in practice [509, 716]. This is also evident from our study area for which for a third of buildings we do not have this attribute.

Moving away from the conspicuous number of storeys, our dataset contains the information about the *net internal area* (NIA), which denotes the usable floor area of

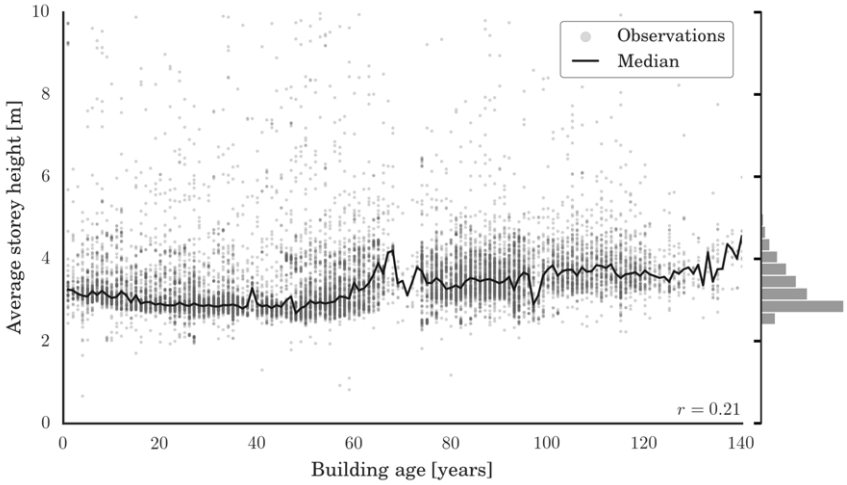


Figure 7.5: Storey height is somewhat associated to the age of the building. We take advantage of this relationship to improve the predictions. The plot also shows the distribution of storey heights in the dataset, indicating that multiplying the number of storeys with an assumed storey height is inherently subject to large deviations.

the units in a building. This area metric is the usual data recorded in real estate management and it differs from the gross area and other related measures. For instance, in the Netherlands the floor area of stairs and escalators and one that has a ceiling height lower than 1.5 m (e.g. in the attic below a sloped roof) is not counted in the NIA, meaning that it can differ substantially from the gross floor area [717]. Another limitation of the NIA, similarly to the number of storeys, is that underground floor area is counted.

Using the floor area of spaces inside a building could be a predictor of the height in two ways: (1) buildings containing more floor area may generally be taller; and (2) dividing the NIA with the footprint area may derive the number of storeys, which can then be used as a predictor when the number of storeys is not available directly.

Figure 7.6 depicts the relationship between the NIA divided by the footprint area (to indicate vertical extent), and the height of a building. It appears that here building use provides more distinction, but also that there are many outliers.

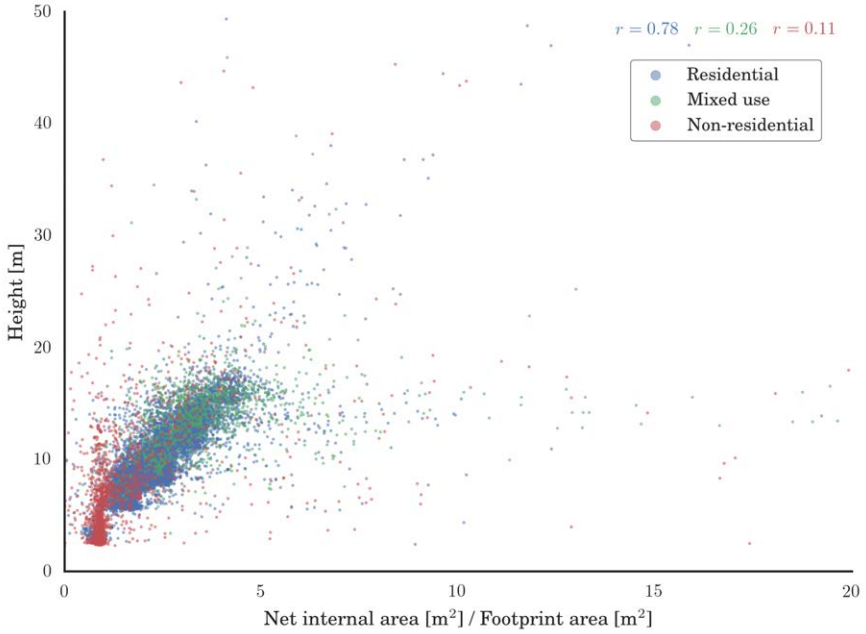


Figure 7.6: Relationship between the net internal area and the building height with the denoted building use. The association is influenced by the building use.

7.3.3.2 Geometric attributes

In our method we compute three attributes from the geometry of buildings that we seek to use as predictors. Researchers in related work have used similar attributes to derive building-related characteristics, e.g. building type and architectural style from footprints [132, 718]. Hence, our hypothesis is that these information may also be useful in predicting the heights of buildings.

First, the *footprint area* is computed. This attribute is introduced in the previous section to put the value of NIA in a vertical context. However, we investigate if the footprint area per se is commensurate with the height of the building.

Second, another shape metric worth considering is the *shape complexity* of a building footprint. Various shape metrics to quantify the characteristics of a polygon are frequently used in remote sensing to classify buildings and plots [620, 719]. Hence we investigate if the complexity of the building footprint can be similarly used to predict the height of a building by analysing their patterns. We use the Normalised

Perimeter Index (NPI)—the ratio of the perimeter of the equal-area circle and the perimeter of the shape: $\frac{2\sqrt{\pi A}}{P}$ (where A is the area of the polygon, and P the perimeter). It is normalised to make it independent of the size of the polygon, and the values range between 0 and 1 with lower values indicating smaller compactness of the shape [720].

Third, for each building we compute the *number of neighbouring buildings*. We have noticed that shorter buildings have more neighbours, so we deem that this metric can be used as a predictor. For this purpose we have selected a buffer of 30 m (we have also experimented with a few other larger values but they did not appear to make any difference so we choose a rather low value to speed up the computations).

The advantage of these three predictors is that no additional attributes and data sources are required to calculate these, as they can be always computed from the geometry, and should be available at all times. This is also valid for the number of neighbours, now that the completeness of volunteered geoinformation has significantly improved, especially in urban areas [509, 721].

In our estimations in Section 7.4 we also test whether it is possible to predict the heights using solely these three attributes, in cases when we have footprints without any attributes.

7.3.3.3 Census (demographics and socio-economic parameters)

During this research we have realised that in residential neighbourhoods there is an association between the average building height in the neighbourhood and its population density (Figure 7.2 illustrates the 92 neighbourhoods of Rotterdam and their population density), as more populous districts accommodate people in taller buildings. We take advantage of this finding to infer the vertical extent of neighbourhoods.

Figure 7.7 indicates the distribution of heights of buildings per neighbourhood of a specific population density in our dataset. The small white dots present the median of all building heights in the same class of population density.

However, while on average there is an association, there are several downsides about this predictor. First, this reasoning applies mostly to residential areas, since we are dealing with population. Second, this predictor encapsulates hundreds, if not thousands of buildings, from which the height of individual buildings cannot be predicted (due to privacy reasons it is not possible to obtain the number of residents down to the level of a building). Third, all neighbourhoods contain both short and tall buildings, preventing a generalised conclusion. Nevertheless, we have tested if this predictor is useful in conjunction with others.

In addition to the population density, our dataset contains the information of the *average household size* and *average income* in each neighbourhood. We also include

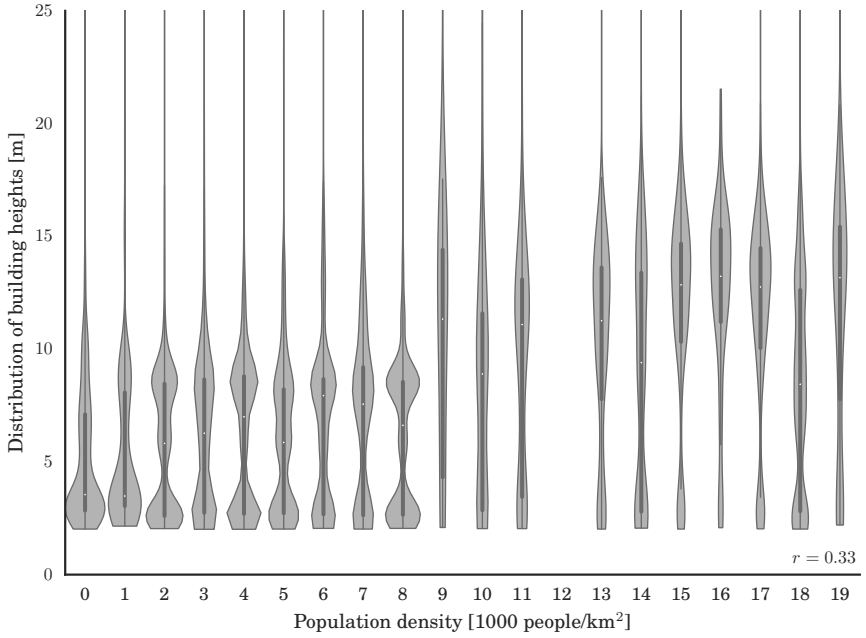


Figure 7.7: Violin plots [722] indicating the distribution of heights per class of population density of the neighbourhood in which the buildings are located. The average building height in a neighbourhood is moderately associated with the population density. We investigate if we can make use of this relation.

them in the training to investigate if they can improve the predictive models.

7.3.4 LOD1 quality measures

We assess the quantitative accuracy of the methods with both the mean absolute error (MAE) and root mean square error (RMSE). The latter metric is sensitive to outliers, but we include it as well following the reasoning of Chai and Draxler [723].

On the other hand, the qualitative assessment of the results is not simple. Technically, we can reconstruct an LOD1 model with any value of the height, and deem it valid, as there is no commonly agreed idea in terms of accuracy that would make an LOD1 model acceptable or not.

While CityGML [10] mentions an accuracy benchmark ('In LOD1, the positional and height accuracy of points should be 5m or less'), it is in our opinion far from

perfect. First, we are of the impression that this value has not really been picked up on and it is not widely accepted by the 3D GIS community. Second, a common misconception is that the standard imposes this requirement. On the contrary, the value is rather described as a recommendation: 'The accuracy requirements given in this standard are debatable and are to be considered as discussion proposals'. Third, the standard declares that the 'Accuracy is described as standard deviation σ of the absolute 3D point coordinates'. The standard deviation is not suited for non-normal distributions, which may appear in geographical data [724]. Finally, the standard states that the 'Relative 3D point accuracy will be added in a future version of CityGML', hence it is not clear whether we can consider the recommendation of 5 m when it comes to the height of buildings.

Despite the shortcomings, and because we are not aware of any national or international standard other than CityGML mentioning LOD1 accuracy requirements, in the results section we will come back to this recommendation as an indication of quality of the generated 3D building models. To bridge this uncertainty, we will tackle the results with two other descriptions that may indicate the quality of the results:

- Comparability: is the accuracy of the constructed 3D models comparable to some occurrences in research and practice? Section 7.2.1 describes a myriad of publications using 3D models constructed from attributes, hence our method already satisfies this requirement. However, we also compare the results to papers that expose the quality of the 3D models reconstructed with traditional techniques.
- Usability: can the derived 3D models be used in spatial analyses? In Section 7.6 we run experiments to support the accuracy with applications (error propagation).

7.4 RESULTS AND DISCUSSION

An illustration of the 3D model generated from the inferred heights is exhibited in Figure 7.8.

7.4.1 Overview of the models and their performance

We have trained predictive models based on different combinations of predictors. This is important because not all attributes covered are always available, but also to test the performance of each in combination with others.

We have selected 17 combinations of predictors. After training the models, we performed the estimations on the test dataset and we present the accuracy of the

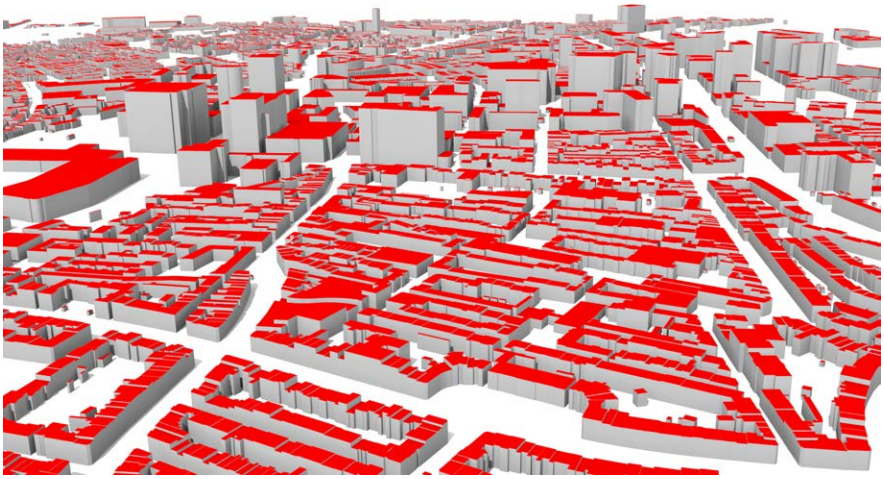


Figure 7.8: 3D model of Rotterdam generated without elevation data. An evaluation of these predictions will be shown later in Figure 7.11.

inferred heights in Table 7.1, along with an overview of the used models as different combinations of predictors.

The MAE varies between 0.8 m and 3.1 m (RMSE from 1.8 m to 4.3 m), depending on the used predictors, satisfying the first two aspects discussed in Section 7.3.4, as we have encountered a number of papers reporting accuracy of 3D models in that range or with an even larger error [22, 38, 725–727]. As a result, we argue that it is feasible to generate 3D city models with this approach.

In the continuation we will address each model separately. The reader should also follow Figure 7.9, which illustrates the importance of each predictor in a particular model, a value useful to indicate the relationship between different predictors, along with the accuracy of each predictive model.

Starting from the number of storeys alone (model 1) is already a convincing motivation for using attributes to generate 3D city models that roughly convey the size and shape of a building. Its MAE is 1.3 m, hence using the number of storeys to extrude building footprints in absence of elevation data is for a reason a relatively popular method to obtain 3D building models without elevation data—the accuracy of the heights obtained with this attribute can be sufficient for a number of spatial analyses. However, no study has critically evaluated these numbers so researchers have used such models without insight into the quality.

As a side experiment, as in related work we have taken two assumed storey heights

Table 7.1: Overview of the predictive models and their performance (accuracy).

	Predictors										Accuracy [m]	
	Cadastre				Geometry			Census				
#	S	U	A	NIA	FA	NPI	N	PD	AHS	I	MAE	RMSE
1	●										1.3	2.1
2	●	○									1.3	2.2
3	●	○	●								1.1	2.0
4				●							1.7	3.4
5	●	○	●	●							1.0	2.0
6		●	●	●							1.8	4.0
7					●						2.3	4.2
8					●	●	●				1.8	3.5
9	●	○	●	●	○	○	○				0.9	1.9
10				●	●						1.3	2.6
11				●	●	●	●				1.1	2.5
12	●				●	●	●				1.1	2.0
13			●		●	●	●				1.4	3.2
14								●	●	●	3.1	4.3
15					●	●	●	●	●	●	1.3	2.9
16	●	○	●	●	○	○	○	○	○	○	0.8	1.8
17	●		●	●							0.8	1.8

— Legend: S—storey, U—use, A—age, NIA—net internal area, FA—footprint area, NPI—normalised perimeter index, N—neighbours, PD—population density, AHS—average household size, I—income.
— The sign ○ denotes that a feature was used but in this particular combination it turned out to be marginally relevant for the RF regressor. Figure 7.9 illustrates the predictor importance in more detail.

(2.8 m and 3.5 m), and simply multiplied them with the number of storeys for each building to assess the approaches in related work. The errors are 1.6 m in both cases, indicating that even with an assumed storey height of a relatively large range without a training process it is possible to achieve a relatively good accuracy.

Adding the building use (model 2) does not improve the results. However, if age was added to the combination (model 3), the error was reduced down to 1.1 m.

Using only NIA (model 4) achieved an accuracy of 1.7 m, and when added to the previous combination the accuracy improved further to 1.0 m (model 5). Again, building use was the least important piece of information. The same combination without the number of storeys (model 6) did not fare well as the previous combinations (1.8 m).

Moving to the geometric predictors, using the area of the footprint (model 7) did not give good results. But when combined with the other two geometric attributes (model 8), the accuracy was 1.8 m. This value is not so accurate as model 5 consist-

Figure 7.9: Visual illustration of predictor (feature) importance for the predictive models, and their performance to predict heights (mean absolute error). The accuracy of each model is illustrated with the filled bar (see the end right for the values—each horizontal line indicates an error of half meter). Models containing only one variable have obviously always the importance of 1, but are given here for comparing accuracies.

ing of cadastre attributes. However, model 8 consists of predictors that are always available, hence it is pleasing to see that the worst case scenario of data availability results in an accuracy that can produce models that are for sure of interest, and might be useful in different application domains.

When combining all the 4 building attributes and 3 geometric predictors (model

9) we achieve sub-meter accuracy. This combination helped to put the predictors in perspective: again the number of storeys has been the most important predictor, but as Figure 7.9 illustrates, there is a variation in the importance of the others.

Model 10 combines the models 4 and 7, which use only NIA and footprint area, respectively. As discussed in Section 7.3.3, having the footprint area might help putting the NIA in vertical perspective. Indeed, as the accuracy improves to 1.3 m.

In model 11 we add other geometric predictors to achieve an accuracy of 1.1 m. Model 12 is similar, with the number of storeys instead of the NIA, achieving the same accuracy. In a similar fashion model 13 includes the age of the building, but with an error of 1.4 m not being as accurate as the previous two.

In model 14 we move to the neighbourhood attributes (census). Evidently, using only these predictors will give the same height for all buildings in the neighbourhood. The error is 3.1 m, due to large variations in the building heights in the neighbourhoods. Nevertheless, this model might be worth considering in cases in which aggregated data per neighbourhood are required (e.g. average height of buildings). In model 15 we add the three geometrical predictors (which are always available so there is no reason not to use them), resulting in an error of 1.3 m.

Finally, model 16 includes all 10 predictors achieving an error of 0.8 m. While the error of this predictive model is the smallest so far, it is interesting because it enables us to evaluate the importance of all predictors together, and to find the minimum combination of predictors that yield a comparably good accuracy.

It appears that the number of storeys, building age, and net internal area have been the most useful predictors. Hence we construct the final combination comprising these three predictors (model 17), resulting in the same MAE of 0.8 m.

The sub-meter accuracy of this model was achieved only with three attributes, all of which are from cadastre, suggesting that the geometric attributes and neighbourhood characteristics have not been of use when cadastral data was available. When the number of storeys was used in a predictive model, all other features were notably less important. However, there are other predictors that have proven useful, which is convenient because the number of storeys is not always available.

Using only the geometric attributes, which are available for all 2D datasets after trivial processing, have predicted the height below 2 meters. This makes sense, because just by looking at the building footprints in Figure 7.2, and judging the shapes of the buildings and their concentration it is possible to deduce that these are terraced houses, which always have about 2 to 3 storeys.

On the negative side, as evident in Section 7.3.3 we have put a lot of hope in the information on the building use while its contribution has been not of use.

7.4.2 Generation of the 3D city model

In Figure 7.8 we have shown the generated 3D city model based on the inferred heights. The prediction models were used based on the availability of data for each building: we have used the predictive model 17, but since for 31% of buildings there was no storey and NIA information, the next best available predictive model that does not require that variable was used automatically (model 15). The heights of the supplemented 3D buildings illustrated in Figure 7.1 in the Introduction were predicted in the same fashion, as sheds did not have cadastral information, and nevertheless they turned out to be fairly accurately predicted.

7.4.3 Cumulative errors

So far we have presented the MAE and RMSE, which aggregate all errors in one value. In Figure 7.10 we analyse the errors in more detail for five selected predictive models. The plot indicates that for most of the buildings in several predictive models the height was predicted within 1 m. In our best predictive model (no. 17) the heights of 77% buildings have been predicted with an error smaller than one meter.

7.4.4 Gross errors

Figure 7.11 supports these observations by exposing the errors of the 3D model given in Figure 7.8: the coloured portion of the walls represents the error in the estimation of the height, while the colours classify the errors into negative (underestimation) and positive (overestimation) errors. Some of the heights were predicted with a negligible error having the value almost the same as from the lidar measurement, however, for some buildings the error was substantial.

In this section we focus on the predictions that were significantly erroneous (e.g. overestimations of more than double the height). Figure 7.10 suggests that in all predictive models a certain share of heights has an error of more than 5 m. Such gross errors considerably affect the MAE and RMSE values.

Most of the large errors were caused by limitations of the method, mostly due to predictor-specific factors described in Section 7.3.3. For example, using the number of storeys substantially underestimated the height of tall constructions that contain only one or a few storeys (e.g. observation decks, religious buildings, and industrial objects such as warehouses, water purification and waste recycling plants). With the number of storeys there were cases of overestimation as well. For instance, the height of a car park of 5 storeys was predicted to be 17 m. But in reality it has a height of less than 8 m because the car park features split-level storeys each counted as a full storey. Moreover, the top surface of the building was also counted as a storey because of the parking function.

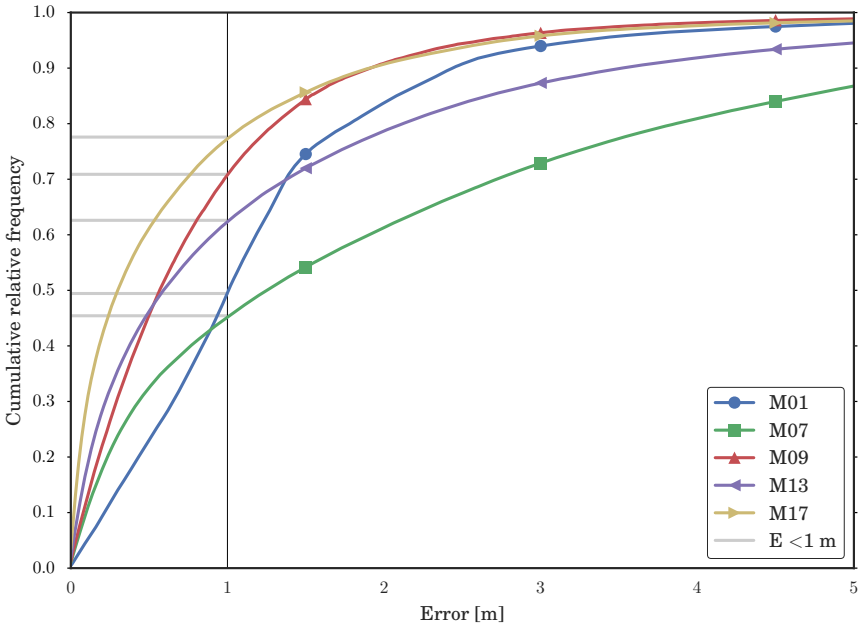


Figure 7.10: Cumulative plot of errors for 5 selected models. In several models the height of more than half of buildings was predicted with sub-meter accuracy.

Other attributes also had some problems as discussed, e.g. a building with a large net internal area that has a few underground storeys, was extruded too much because the method falls short with the NIA portion below ground. The large errors visible in Figure 7.11 were mostly caused by these reasons.

In general, if we consider the circumstances, we can conclude that the predictive models resulted in surprisingly good performance. However, due to gross errors researchers should exercise caution when using models derived from attributes. As we highlight in the literature review (Section 7.2), such models are not uncommon, and they are seldom critically evaluated being accepted as is, and sometimes used in spatial analyses where a relatively small error in the 3D model may cause a substantial error in the spatial analysis.

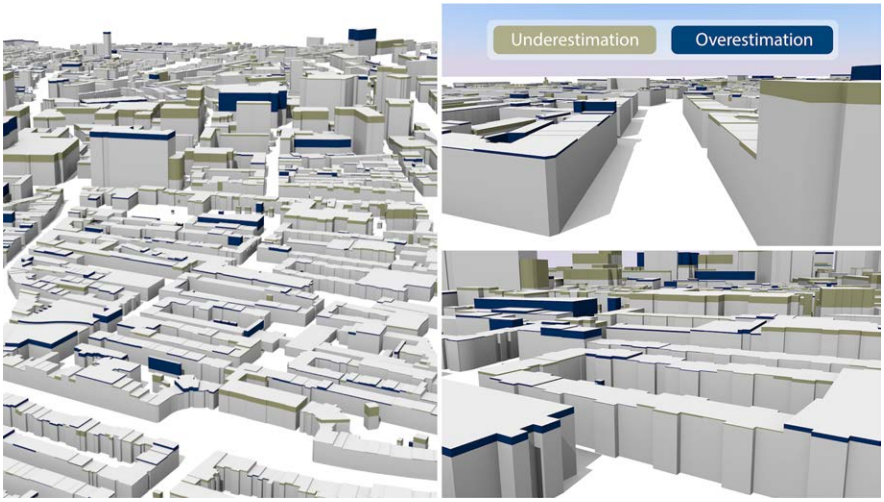


Figure 7.11: Evaluation of the 3D model presented in Figure 7.8. The colours represent underestimations (missing portions of buildings), and overestimations (excess heights).

7.4.5 Limitations of the validation

It should be noted that many gross errors were caused by limitations of the validation, as the reference data is not without faults.

First, we have noticed at least a few errors in the attributes, such as the incorrect number of storeys in the cadastral registration.

Second, the reference height of the building, which we have derived from the point cloud, contains some inconsistencies (either too low or too high). While the point cloud is accurate, the height might not be always represent the top of the construction, as factors such as adjacent tall buildings, may influence it. Rotterdam has also a fair share of buildings that overlap each other, which also influenced the height. Figure 7.5 hints at errors in the measurement of the reference height: some of the storey heights were unrealistically small.

Third, a notable problem with the validation are the shortcomings of the LOD1 models in which a building may have the notion of multiple heights, hence an LOD1 often does no justice in representing it [500, 575]. Several buildings for which the predictions resulted in a gross error are discontinuous in height. For example, a building with a sloped top that covers more than a few storeys. Buildings including setbacks also fall in this category (e.g. the floor area at the top is considerably smaller than at the bottom of the building; see Figure 3.1b for an example of such a building).

Figure 7.12: Examples of buildings in Rotterdam with an ambiguous notion of height. These cases illustrate that the 'credibility' of LOD1 models (extruding the building footprint to a single height) depends on the architecture of buildings.[24]

A few examples from Rotterdam are given in Figure 7.12. These cases suggest that the quality of extruded LOD1 models to represent reality depends on the morphology of constructions. In these cases an LOD1.3 would be more advantageous than lower LOD1 instances (e.g. LOD1.2) since it allows capturing such discontinuities in the height.

Finally, we have encountered several large footprints comprising arguably multiple buildings. While the dataset we have obtained contains buildings at a quite granular level (e.g. terraced houses—see the buildings in Figure 7.2), due to the modelling conventions of the Dutch cadastre some of the footprints encompass a large area containing multiple buildings. We have considered partitioning such buildings depending on the continuity of the height, following methods such as the ones of Commandeur [573] and Kada and McKinley [581]. However, in that case it would not be possible to meaningfully distribute attributes across partitioned buildings.

An example of such a case is illustrated in Figure 7.13. According to the number of storeys (33), in the predictive models involving this attribute, the height was predicted to be between 94 and 100 m, which corresponds to the height of the tallest tower. However, the height from lidar was predicted to be 17.5 m, flagging a large error in the validation.

[24] Imagery (c) Microsoft and Pictometry International Corp.

Figure 7.13: Example of a footprint covering multiple buildings of significantly different heights. The labels represent the number of storeys for each footprint, which we have used to predict heights, and in this case resulted in erroneous predictions. In cases such as this one it is ambiguous to discuss building heights.[25]

Predictive models not involving the number of storeys were more promising (e.g. model 15 predicted 15.5 m), however, here it anyway does not make sense to talk

[25] Footprints (c) Kadaster and City of Rotterdam. Imagery (c) Google and Aerodata International Surveys. Oblique imagery (c) Microsoft and Blom.

about the height of a building.

A byproduct of our method is that it can be also used to find errors in the underlying data: gross errors may indicate not the imperfection of the method, but rather inconsistencies in the reference data. Thanks to them we have found these errors and special cases.

7.5 APPLYING THE INFERRED PATTERNS TO ANOTHER CITY

In order to train our prediction models, we have taken a small subset of buildings in Rotterdam. While the heights were predicted without elevation data, training involves reference heights and the patterns may be limited to a unique local context. Hence one might argue that this method is not entirely without elevation data. Furthermore, training data might be applicable only for a specific context that is not applicable somewhere else.

As stated in Section 7.3, we argue that our method can be used in other areas far from the training area and where there are no elevation data. To prove this assumption we apply the predictive relationships on buildings in Leeuwarden, one of the farthest Dutch cities from Rotterdam (170 km distance). Leeuwarden is not as urbanised as Rotterdam, and it is also a city with a different cultural environment and history, so it is interesting to investigate whether the predictive models developed in Rotterdam would obtain comparable results in such a different place (Figure 7.14).

We have applied the inferred relationships of 9 predictive models (we could not test the models including number of storeys because we could not obtain that data for Leeuwarden), and validated the inferred heights with the heights obtained from the national height model of the Netherlands (AHN). The mean absolute errors of the evaluated models are shown in Figure 7.15. The heights in Leeuwarden estimated from the predictive models trained in Rotterdam are slightly less accurate, with the interesting exceptions of models 4 and 14 which are more accurate in Leeuwarden than in Rotterdam, possibly due to less variation in the heights of buildings and architecture.

This evaluation suggests that it is feasible to train a predictive model in one area where heights are available and use it for another one. Furthermore, we assume that we would obtain comparable results also beyond the border of the Netherlands, at least in Belgium and northwestern Germany.

For future work it would be beneficial to investigate the performance of the method in areas with a considerably different type of urban fabric, and to research whether the advancement of the sampling method can lead to an improvement in the results. In this research we have used simple random sampling. Perhaps using more advanced sampling methods (e.g. stratifying buildings based on different character-

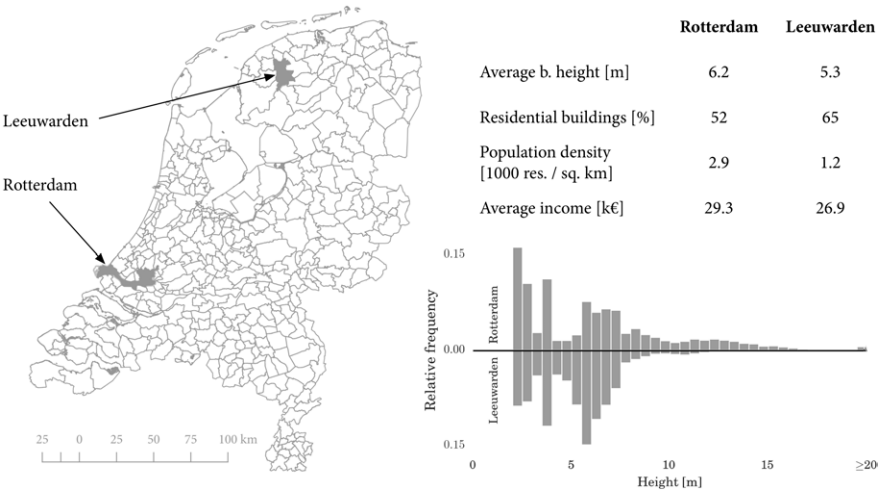


Figure 7.14: Location and comparison of Rotterdam and Leeuwarden. The two histograms reveal the differences in the heights of buildings.

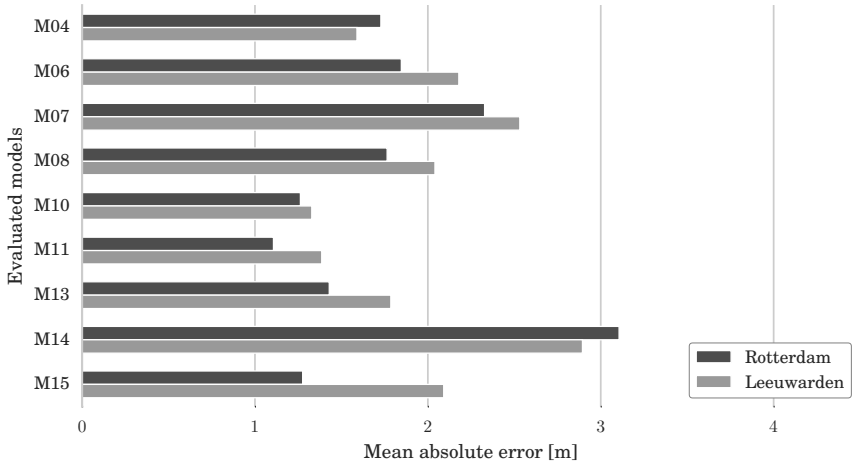


Figure 7.15: The performance of inferring heights far from the area where the predictive models were developed. Some predictive models fare even better than in the original context where they have been developed.

istics) would lead to a more effective training dataset.

7.6 POTENTIAL APPLICATIONS AND LARGE-SCALE EXPERIMENTS

3D city models derived from heights inferred from non-elevation sources are not as accurate as those derived from direct measurements such as lidar. Nevertheless, they can generally indicate the urban form that is useful in various urban applications [76, 678], and may be of use for various purposes where high accuracy is not crucial. For example, they may be used for visualisation purposes (e.g. for navigation), and their appearance may be enhanced with procedural modelling techniques.

This reasoning is also evident from Section 7.2 where we review several papers using 3D models generated from the number of storeys. In addition, these applications may include determining the urban heat island effect [728], computing the building roughness indicator [729], analysing the vertical growth of a city [730], seismic vulnerability assessment [731, 732], noise pollution predictions [259], waste management [733], estimating air pollution [623], storm surge vulnerability assessment [368], and analysing urban density [734]. It is our observation that most of these papers used crude data, hence our assumption is that the applications are not very demanding when it comes to accuracy of the underlying 3D models (it should also not be ignored that if such 3D models are constructed from very accurate footprints such as in our case, the planar accuracy of the 3D model still remains high).

Furthermore, some spatial analyses require aggregated data of buildings at a neighbourhood level. Such applications are calculating the volume of buildings in the neighbourhood, e.g. for energy demand estimation and thermal simulations [735, 736], population estimation [737], and estimating the material stock [738]. In fact, Rotterdam was already a subject of similar analyses [121] in which aggregated data on the energy demand of neighbourhoods have been estimated.

Calculating and presenting the volume of each building individually would not be very accurate, but our hypothesis is that on a larger area such as neighbourhoods, the computations would not be that far from the true value.

In order to test this hypothesis we compute the volume of buildings per neighbourhood, and compare it to the value obtained from lidar. For the experiments we have chosen the model 15, which performed moderately well in our analysis, and it does not require attributes of buildings. Using the estimated data, the error in the computation of volume of individual buildings is 20.2%, however, when summing the volumes per neighbourhood the error drops to 13.8%. The volume on the scale of Rotterdam was computed with an error of 3.8%. Predictive models with a higher accuracy have further yielded better results in computing the building stock volume of neighbourhoods (e.g. 8.4% for model 9).

With this section we also fulfill our third quality requirement stated in Section 7.3.4, as these results suggest that in absence of very accurate 3D city models, those obtained without elevation data are worth considering for certain spatial analyses.

The topic of error propagation, i.e. the influence of the acquisition error on the results of a spatial analysis, will be studied further in Chapter 11 and Chapter 12. The research will reveal that many datasets around the world are constructed with a combination of different techniques, resulting in different accuracy levels in the coordinates. The data generated in this chapter serves as an excellent example of that: the *xy* coordinates of the geometry have a considerably higher accuracy than the *z* coordinate.

7.7 CONCLUSIONS

Extrusion is the most well known technique to enhance the LOD of a dataset. It requires 2D datasets and numerical values of the vertical extent of each feature, resulting in LOD1 block models generated from LOD0 models. However, the corresponding elevation data is not available as frequently as building footprints.

The contribution of this work is advancing the extrusion without explicitly requiring the height values, since they are inferred from the 2D dataset: essentially, we have demonstrated that 3D city models can be generated automatically without elevation data. Hence the chapter demonstrates that it is possible to augment LOD0 datasets to LOD1, without additional data sources, employing machine learning techniques. This is namely for three reasons:

- Several attributes available solely from 2D data can hint at a building's height. The method works also with attributes other than the number of storeys, so 3D models can be generated outside areas with the availability of this convenient but not omnipresent attribute.
- The achieved accuracy is comparable to many other instances used in research and practice. In fact, in many of the predictive models, the majority of the heights was predicted with an error below 1 m, which fares better than a number of instances of 3D models acquired with traditional direct techniques.
- The experiments and the discussion reveal that these models could be useful in a number of spatial analyses.

Despite advancements in remote sensing, elevation data required to generate 3D city models will still not be available for many areas around the world for some time. In fact, many developed countries still lack national coverage of elevation data suitable for producing 3D city models [667]. In the meantime, the insights presented

in our chapter can serve to quickly obtain 3D models in LOD1 in such places where there are no elevation measurements available, but are rich with footprints from governments and volunteered geoinformation. Depending on the purpose, these 3D models obtained with approximate heights can be a reasonable provisional solution until elevation measurements become available, or if they are outdated—before they are refreshed.

While we have not invented the extrusion of footprints based on the number of storeys, no study has critically evaluated this method so researchers have used such models without the awareness of the quality. Furthermore, no research has been done on using machine learning techniques and whether there are other attributes that can be used to predict the height of a building, which is important especially because the information on the number of storeys still suffers from completeness issues. Our comprehensive study bridges these gaps: we investigate 9 other attributes, and perform experiments in 17 predictive models obtaining a thorough quality insight in what are the possibilities to generate 3D city models with non-elevation attributes. While the number of storeys remains the most useful attribute, including other predictors improves the predictive accuracy.

Related to our LOD specification presented in Part I, this chapter shows the enhancement of LOD0.1 to LOD1.2. This chapter also demonstrates how our specification is shaped after the capabilities of acquisition techniques: LOD1.3 models cannot be generated with this technique, because it is not possible to partition building parts without additional data sources, nor it it possible to independently predict the height of each part. The visual examples of buildings with an inconsistent height also imply the importance of having LOD1.3 models. Furthermore, the chapter gives an example of a dataset with a non-elevation vertical geometric reference Hx, as defined in Chapter 5.

For future work we plan to investigate whether it is possible to infer the roof type, possibly leading to the generation of LOD2 models without elevation data. The papers of Allani-Bouhoula and Perrin [690] and of Henn et al. [590] suggest that this idea is not unrealistic. For example, in the latter, in automatic mapping of buildings from point clouds researchers use attributes of buildings to improve the quality of the reconstruction of roofs. Perhaps the characteristics of buildings alone can give a reliable indication of the roof type.

In addition, it might be beneficial to further analyse the surrounding context of a building, which is readily available from volunteered geoinformation. For example, Bakillah et al. [739] analysed amenities around buildings to map population. Following a similar logic (taller buildings host more people, which results in more demand into supermarkets and other facilities) might result in improved predictions.

CHAPTER 8

Managing multi-LOD data

This chapter is based on my paper [740]:

> Biljecki F, Ledoux H, Stoter J (2015): Improving the consistency of multi-LOD CityGML datasets by removing redundancy. *3D Geoinformation Science*, Springer International Publishing, pp. 1–17. doi: 10.1007/978-3-319-12181-9_1

and partly on the paper [741]:

> Arroyo Ohori K, Ledoux H, Biljecki F, Stoter J (2015): Modeling a 3D City Model and Its Levels of Detail as a True 4D Model. *ISPRS International Journal of Geo-Information*, 4(3): 1055–1075. doi: 10.3390/ijgi4031055

Chapters 6 and 7 suggest that there are scenarios resulting in datasets in two or more LODs, i.e. multi-LOD datasets, as generalising or augmenting existing data (Chapter 7) inherently results in data in multiple LODs. Such datasets enable more flexibility, at the expense of their redundant maintenance and storage. A solution to managing such problems would be by integrating multiple LOD representations. This chapter mitigates these issues by introducing two ways of linking data stored in two or more LODs.

The first approach recognises identical geometries across LODs and links them to facilitate maintenance and update. We describe the possible topological cases, show how to detect these relationships, and demonstrate how to store them explicitly. A software prototype has been implemented to detect matching features within and across LODs, and to automatically link them by establishing explicit topological relationships (with XLink). The experiments ran on our test datasets suggest a considerable number of matched geometries. Furthermore, this method doubles as a lossless data compression method, considering that the storage footprint in the consolidated datasets has been reduced from their dissociated counterparts.

The second approach explores the possibility of linking multi-LOD data by considering LOD as the fourth geometric dimension. Such an approach integrates a multi-LOD dataset in a single hyper-dimensional (4D) construction.

8.1 INTRODUCTION

The OGC standard CityGML, and other 3D modelling standards such as COLLADA [742, 743] and ISO's X3D [744] allow the storage of multiple representations of a 3D model, in order to facilitate the multi-scale use of the models and to improve the computational efficiency of spatial operations.

Although 3D GIS datasets may contain multiple LODs (two or more 3D datasets of the same region may exist), they are seldom linked beyond an administrative link between object identifiers (i.e. they share the same building ID).

In this chapter we investigate the possible approaches to link multi-LOD datasets leading to improvements in the consistency and storage of multi-LOD datasets, but also with some benefits to single-LOD representations. It is our experience that in practice, besides exemplary models, single-LOD datasets do not contain the explicit representation of topological relationships, hence developing a joint method that is beneficial for both possibilities is important.

The chapter presents two approaches. First, we observe and take advantage of the practical fact that many of the stored geometries (primarily polygons) in 3D datasets are geometrically equal both within a single LOD and across multiple LODs. By determining the topological relationships between such reoccurring geometries and storing them explicitly, the consistency of 3D models can be increased, as we highlight in this work. The second approach is theoretical. It is briefly discussed integrating the LOD as a (4th) spatial dimension, resulting in a space-scale (4D) construction.

8.2 INTEGRATION BY LINKING MATCHING GEOMETRIES

As it is the case with the other chapters, this work is focused towards CityGML and its LOD concept However, most of the developed work is applicable to other formats as well. Our work consists of the following contributions: (1) we have investigated and described several cases of reoccurring geometries and introduce a terminology to distinguish them; (2) we have developed robust algorithms that efficiently index the geometries in CityGML datasets and that take advantage of the geometries that reoccur by explicitly storing their topological relationships; (3) we have developed a software prototype that analyses CityGML data and automatically computes explicit topological links between matching geometric features through the XML Linking Language (XLink) mechanism; and (4) we show with experiments that a considerable subset of data can be matched. We have tested the method with our synthetic dataset (Chapter 6). Because matched geometries are stored only once, the consolidated dataset is compressed without loss of information.

While developing the method, we have realised that in practice most of the geometries that are reoccurring are not identical and cannot be readily matched. Therefore, we have investigated other cases and covered them as well.

The implementation and the results demonstrate that after linking a higher degree of the consistency of the data is achieved, contributing to an efficient storage and maintenance. For example, if the geometry of a feature is altered in one LOD, thanks to the established explicit topological representations this change may propagate through other LODs.

8.2.1 Background and related work

Consistency of 3D city models is an important topic in GIS, in which topology plays a prominent role [63, 745]. Current research efforts focus on the relationships of features within the same representation, e.g. the validation of solids [633], and making use of topological data structuring to improve rendering performance of 3D city models on mobile devices [134]. To the extent of our knowledge, there is no related work to detect and link geometric features across multiple representations.

While this chapter generally describes a way how to increase the consistency and to compute topological relationships in a model, it is focused on the maintenance and storage of (3D) GIS datasets, which is a topical subject in academia and industry [67, 583, 746]. Updates of models often introduce errors [747], so increasing consistency is one of the prerequisites for an efficient workflow.

In this section we describe the redundancy and benefits of an established topology with respect to the scalability of the models: for single representations (single-LOD datasets), and for multiple representations (multi-LOD datasets).

8.2.1.1 Single-LOD datasets

Research that has been done in this topic is focused on the relationship of real-world features within the same representation [17, 748]. For example, the topology of two coinciding polygons, such as a wall shared by two buildings. The consistency that is achieved by establishing explicit topological relationships in practice simplifies the maintenance of the data and reduces the redundancy in the storage.

Figure 8.1 illustrates an example of the benefit with respect to the maintenance of a 3D model. The left model (Figure 8.1a) depicts a building with a wall that contains a window. The polygon representing the wall is given in darker colour, and it contains a hole (inner ring), which is filled by another polygon representing the window. The interior ring of the wall polygon corresponds to the exterior ring of the polygon representing the window. In a model without established explicit topological relationships, the two features are not linked in any way. When the geometry

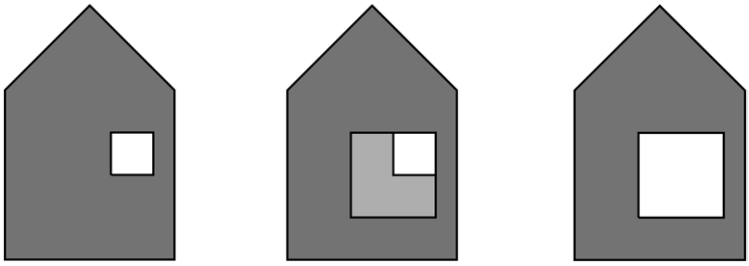


Figure 8.1: The determined explicit topological relationships in a dataset has a significant benefit to its maintenance. This example indicates the benefit on a wall with a window that is being enlarged: (a) a building model before the update (enlarging a window); (b) an updated model without links resulting in inconsistency; and (c) an updated model with established links (desired case).

of a part of the object is updated, e.g. the window is enlarged, the change does not affect the related geometry (i.e. hole of the wall), leading to redundancy in the process (see Figure 8.1b). In a model with established topological representations, the change properly propagates to the related features (see Figure 8.1c for the desirable outcome).

8.2.1.2 Multi-LOD datasets

On top of the redundancy in a single representation, which multiplies with each new representation in a multi-LOD dataset, there is also additional redundancy. For example, should a feature be changed, in practice the update must be done manually for each representation. Furthermore, because of the increasing complexity of the models, the size of the datasets substantially increases with the increase of the LOD, making the storage less feasible. The surge in the size is not only caused by the growth of the amount of details, but also because of the redundancy that could be removed, as we reveal later in the chapter.

Therefore, despite the fact that this option is available in CityGML, models are usually derived in a single LOD representation. While multi-LOD datasets are rare, when they are available they are usually produced by generalisation from finer LODs (e.g. see [529, 553, 594]), for example, as a bounding box of an LOD2 [555], or of an LOD3 including features such as antennas on roofs [579]. This is beneficial for this research, since it results in datasets where many of the geometries are preserved, and are identical in more than one representation. For example, the ground surface of a

Figure 8.2: The rationale of the method to link multiple LODs. If two or more geometries are found to correspond, links are created. In this case two polygons (walls) are identical in two LODs. However, in one LOD the polygon has a hole, hence only their exterior rings are linked. How and where to establish the links while balancing the maximisation of links and the topological structure is the main concern of the consolidation process.

building (i.e. `GroundSurface`) is usually identical in all representations.

Detecting and linking such occurrences would be a first step towards complementing the discussed practical shortcomings.

Furthermore, recent work suggests that the share of multi-LOD datasets is increasing, raising the importance of this topic. For example, Yang et al. [749] presents a method to reconstruct 3D building models at different LODs from airborne lidar data, suiting different use cases.

8.2.2 Methodology

We have developed a method that searches for matching geometries in the datasets and links them. Figure 8.2 illustrates the desirable outcome of the algorithms with an example of three LODs where some of the geometries are reoccurring and are consequently linked.

While examining multi-LOD datasets we have realised that there are different cases of corresponding geometry, not only polygons that are identical and that can be directly referenced. For example, polygons that share the exterior ring, but their interior is different (common in CityGML LOD3 where openings are allowed, see Figure 8.2). Furthermore, specific cases such as two equal polygons whose starting point is different should also be handled.

8.2.2.1 Terminology

In this chapter we focus on the two geometric feature types: polygons and linear rings. We consider two or more geometric primitives *identical* if they are topologically and geometrically equivalent, i.e. they can be readily linked and re-used. The geometric representations of the ground plane of a building in two LODs are usually identical. Two or more primitives are *partially identical* if they are not identical and if their relationship has one or more properties that prevent them to be identical:

- The orientation of their vertices is different, i.e. their normals are reversed. For example, two buildings share the same wall.
- They constitute different aspects of their parent primitive, e.g. two linear rings correspond, but one forms the exterior ring of its parent polygon, while the other forms the interior ring to describe a hole. A prominent example of this case is a wall with a hole that represents an opening—window or door that is stored separately.
- The number of points in their rings is not equal while the shape and location are identical. This is caused by *redundant* points p_i, where p_{i-1}, p_i, p_{i+1} are collinear. The removal of such points would not compromise the shape and location of the polygon.
- The starting point in the linear ring is different. This discrepancy might be easily detected and corrected *on-the-fly*, hence it will not be particularly emphasised in the continuation of the chapter.

Two geometries *match* if they are either identical or partially identical. When a match of two or more geometries is found, one is selected as the *resource*, and the rest are *linked* to it.

8.2.2.2 Topological relationships

In this section we outline the possible cases of the *matching* geometry, i.e. topological relationships between polygons, and their constituting components—exterior and interior ring(s). We have investigated the possible cases, and their occurrences in real-world datasets, which we highlight in Table 8.1. The sign = denotes identical primitives, – partially identical, ≠ non-matching, and ∅ denotes no geometry. The matched primitives in each case are shown in red. In the example column, when two objects are separated, the example refers to the multi-LOD case.

While more 'permutations' are possible, they are not valid according to GML, hence the list does not take into account invalid cases. Such an example are two

exterior rings that are identical, but their interior rings have a reversed orientation (which is shown in the last line as an exceptional example).

The case 0 is the (usual) case where two primitives have no topological relationship in the context of this research. Identical features rarely occur within the same LOD (in contrast to partially identical), so the cases 1 and 2 where the geometries of two features are identical are often present in multi-LOD datasets. One example is case 2 is when the ground surface of LOD1 and LOD2 contains interior (e.g. a courtyard). Case 3 is also typical in multi-LOD datasets (Figure 8.2), while case 4 is more common in single-LOD datasets where two buildings share the same wall[26]. The fifth case extends the previous with holes. Cases 6 and 7 are unusual in the real-world, and case 8 is similar to the fifth case in the occasion when the polygon is not identical. Cases 9 and 10 are cases of interchangeable roles of the exterior and interior rings, and might be rather considered as extended cases. Case 9 is uncommon, and case 10 is usually occurring in finer single-LOD datasets as the relationship between a wall and an opening.

8.2.2.3 Overview of the method

The workflow of the method to match regards both matching primitives within the same LOD and across multiple LODs:

1. *Indexing.* Index all points, linear rings, and polygons in all LODs for efficiency (Sec. 8.2.2.4).

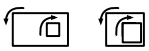
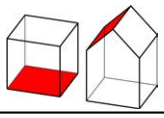
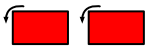
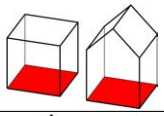
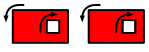
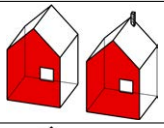
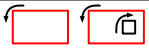
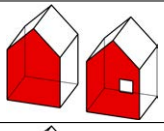
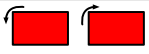
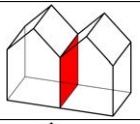
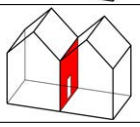
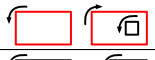
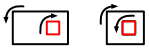
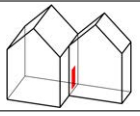
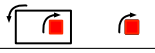
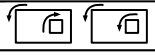
2. *Matching.* Detect matching geometries and flag them. Because of different topological relations, which have been introduced in the previous section, the detection of the matching geometries is done in multiple phases (Sec. 8.2.2.5).

3. *Consolidation.* Analyse the matched geometry and remove redundant data by replacing them with a link to one other matching representation resource, with modifications if necessary (Sec. 8.2.2.6).

Figure 8.3 provides the simplified workflow of the method and the relationship between the features when searching for matches. Because all features are indexed, the searching algorithm has a considerably reduced subset of potential matches. After ruling out non-matching features in the indexed subset, the algorithms detect the matched features and classify them according to the cases presented in the previous section.

[26] Note that it is also common to split the polygon into two polygons if the buildings are not of the same size [750].

Table 8.1: Cases of topological relationships of rings and polygons. The curved arrows denote the ring's orientation, while the long horizontal arrows indicate that there is a relationship between one polygon's exterior to another polygon's interior.

Case	Ring		Polygon	Graphical explanation	Real-world example
	Exterior	Interior			
0	≠	≠	≠		
1	=	∅	=		
2	=	=	=		
3	=	≠	≠		
4	–	∅	–		
5	–	–	–		
6	–	≠	≠		Unidentified
7	≠	=	≠		Unidentified
8	≠	–	≠		
9	=		≠		Unidentified
10	–		≠		
Invalid	=	–	?		Not possible

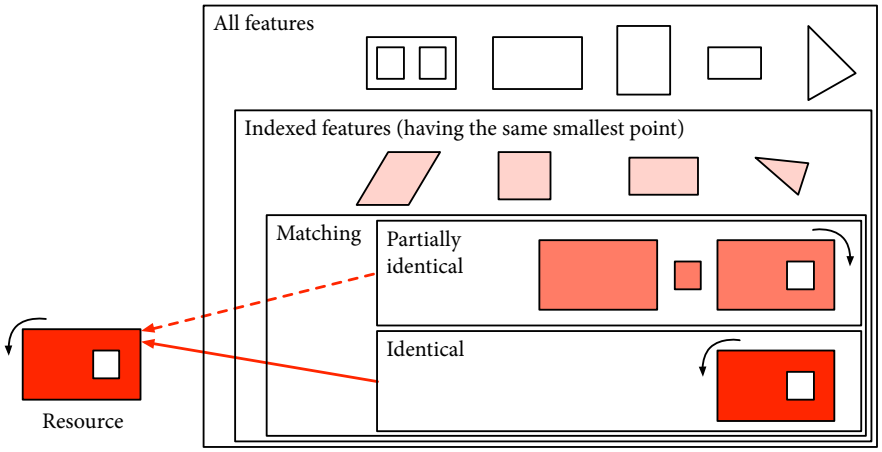


Figure 8.3: Simplified classification of the relationships and workflow of the method. The primitives in the dataset are first indexed, and then tested for matches, of which there are different categories.

The developed algorithm and the implemented software prototype are focused towards ideal cases where the vertices of the geometry are identical across LODs and where the geometries fully correspond. This is useful for datasets produced with generalisation, however, when used on datasets with a different lineage it might not produce results to the same extent. This could be solved by introducing the snapping of the points according to a tolerance threshold, and more sophisticated matching of similar geometries. The automatic matching of the representations that are acquired with different techniques would require employing more advanced algorithms such as [751, 752], extending related work done in cartography (e.g. [753, 754]) to 3D GIS, and would probably result in a lossy compression (partial data discarding).

8.2.2.4 Algorithm for indexing the geometry

In order to make the search for the matching geometry more efficient and enable the consolidation of larger datasets, the polygons and their rings are first indexed. We have decided to build an index where each polygon's ring is indexed according to its smallest point. The smallest point is the point with the smallest coordinate value, i.e. the one that is closest to $(-\infty, -\infty, -\infty)$. Each valid ring has one such point, hence, it can serve for the purpose of indexing. This point should not be confused with the starting point of the ring, which is not relevant here. This step considerably reduces the search time for matches since in practice only a few non-matching polygons

share the same smallest point.

8.2.2.5 Algorithm for the detection of matching geometries

After indexing, the rings (both exterior and interior) are queried for their relations. The algorithm first removes vertices that are redundant (i.e. being collinear with its preceding and succeeding points). The algorithm is given in Algorithm 8.1.

8.2.2.6 Algorithm for the consolidation of the data

After detecting matching geometries, the last phase of the method involves consolidating the data (linking), i.e. analysing the relationships and determining the level of the relationship that can be linked. A straightforward solution would be to directly link the matching rings. However, because of the different cases and hierarchies, this phase is not forthright, and it can be solved in multiple ways. For example, if two rings that form the exterior of two polygons match, this does not necessarily mean that the polygons can be matched right away, since the interior may be different (e.g. see case 3 in Table 8.1). Furthermore, one of these rings may be an interior of another polygon, that is further related to another polygon in another way. Therefore, maximising the number of links that are established is the main concern when designing such an algorithm, and cannot be solved easily by determining the frequency of the occurrences and selecting the topmost ring. Furthermore, the development of the algorithm is associated with the content of a targeted dataset, as an algorithm might not be equally beneficial when employed for consolidating two different datasets. This problem is related to the field of data compression [755, 756].

We have designed a top-down approach that first iterates polygons with holes, comparing the matched rings, and builds a hierarchy of features, continuing to polygons without holes. This is particularly beneficial for cases 1, 2, 3 and 10, which are the most common. The algorithm is given in Algorithm 8.2.

8.2.3 Implementation and results

8.2.3.1 Test data

We have used a dataset in multiple LODs that was automatically generated with the method presented in Chapter 6. The dataset contains 100 buildings. Some of the LODs are represented in two ways, as a `<gml:Solid>` or as semantically structured surfaces (`<gml:boundedBy>`). This is done in order to extend the experiments by comparing the different variations of the models that are valid [434, 436]. We have used six different representations for the experiments: LOD1.2 (solid), LOD2.1 (b-rep and solid), LOD3.2 (b-rep and solid), and LOD3.3 (b-rep). The finest represen-

Algorithm 8.1: Algorithm for the detection of matching geometries.

Input: Indexed features, where each ring has its smallest point indexed (Sec. 8.2.2.4)
Output: Topological relationships between the geometries

```
Iterate all linear rings and remove all redundant points
for each ring r_j of each polygon do
    for each ring r_k that shares the same smallest point do
        relation = 1 {Assume identicalness until proven otherwise}
        if r_j and r_k are identical features then {Avoid comparing the ring to itself}
            Go to 3
        end if
        if the relationship between the two rings has already been checked then
            Go to 3
        end if
        if the number of points of r_j and r_k are equal then
            Reinstate the points{i.e. reorder them so the first point coincides}
            while relation = 1 do
                for each point p_j^i in r_j do {Start from the second point because the first one is identical
                in any case (index)}
                    if p_j^i ≠ p_k^i then {Points not identical, so there is no match}
                        relation = 0
                    end if
                end for
            end while
            if relation = 0 then {If the match has not been found, try for the other orientation}
                reverse = True {Assume that the linear rings are inverted until proven otherwise}
                n = number of points
                while reverse = True do
                    for each point p_j^i in r_j do {Start from the second point because the first one is
                    equal in any case (index)}
                        if p_j^i ≠ p_k^(n−i) then {Points not identical when the rings are reversed}
                            reverse = False
                        end if
                    end for
                end while
                if reverse = True then {the relationship between the rings is reversed}
                    relation = -1
                end if
            end if
        else {If the number of points is different, the features are not identical}
            relation = 0
        end if
        if relation = 1 or relation = -1 then {If the rings are identical or partially identical, store this
        information}
            Store the information about the relation r_j ↔ r_k
        end if
    end for
end for
```

Algorithm 8.2: Algorithm for consolidating the data.

Input: Topological relationships, obtained from the previous algorithm in Sec. 8.2.2.5
Output: Consolidated dataset with established links

1: **for** each polygon P_0 that contains the interior and has at least one ring matching to a ring of at least one other polygon P_i **do** {Both exterior and interior rings are taken into account}
2: **if** P_0 and P_i were already matched **then**
3: Go to 1
4: **end if**
5: **if** P_i comprises only polygons with holes **and** are identical to P_0 **then**
6: Establish P_0 as the resource, and link P_i
7: Remove the primitives from further consideration
8: **end if**
9: **if** the exterior of P_i is identical to P_0 **then**
10: Establish the exterior of P_0 as the resource, and link P_i
11: Remove the primitives from further consideration
12: **end if**
13: **if** P_i is partially identical to P_0 **then**
14: Establish the exterior of P_0 as the resource with necessary modifications, and link P_i
15: Remove the primitives from further consideration
16: **end if**
17: **end for**
18: **for** each polygon P_0 that does not contain the interior and has at least one ring matching with a ring of at least one other polygon P_i **do**
19: **if** P_0 and P_i were already matched **then**
20: Go to 18
21: **end if**
22: Repeat the steps above without the operations on the interior
23: **end for**
24: Store the consolidated dataset with determined topological relationships

tation is interesting to consider because it presents a considerable increase in the storage footprint comparing to the coarser representations, e.g. roughly twice the size of LOD3.2.

8.2.3.2 Implementation

The implementation was done in Python. The XML Linking Language (XLink) mechanism was used to realise the links between the features. It allows elements to be inserted into XML documents for creating and describing links between resources [757].

XLink has been used already in many research projects that employ CityGML [412, 758–760]. It is also mentioned in the GML standard [652, 761] and in the CityGML standard [10] as the preferred way of explicitly storing the topological relationships. In CityGML, it is primarily used for referencing mutual geometries of two objects

(e.g. a building and a building part), and for the re-use of the geometry in the semantically enriched boundary representation for the construction of solids [510].

The example below reveals a data excerpt for a resource and a link. When two or more matching geometries are detected, the prototype adds an identifier to the `gml:id` attribute of the resource, i.e. the preserved single instance of the multiple matches. For the unambiguous identification of resources and links, the Universally Unique Identifier (UUID) standard has been used [651].

```
<gml:surfaceMember
  <gml:Polygon gml:id="0127875b-a2a8-498e-8024-770fd661aef6">
    <gml:exterior>
      <gml:LinearRing>
        <gml:posList>
          173469.0 526442.0 0.0
          173469.0 526446.64 0.0
          173478.81 526446.64 0.0
          173478.81 526442.0 0.0
          173469.0 526442.0 0.0
        </gml:posList>
      </gml:LinearRing>
    </gml:exterior>
  </gml:Polygon>
</gml:surfaceMember>
```

Afterwards, the contents of each matching geometry is removed and a link is added pointing to the resource that contains the geometry:

```
<gml:surfaceMember xlink:href="#0127875b-a2a8-498e-8024-770fd661aef6"/>
```

Geometries that have opposite normals (reversed ring orientation) can be referenced with `<gml:OrientableSurface>`. For example, the reversed match of the polygon above may be stored as:

```
<gml:OrientableSurface orientation="-">
  <gml:baseSurface xlink:href="#0127875b-a2a8-498e-8024-770fd661aef6"/>
</gml:OrientableSurface>
```

For reversing the linear rings, `<gml:OrientableCurve>` can be used.

These examples not only indicates the achieved consistency, but also the decrease in the storage footprint.

8.2.3.3 Results

The implemented software prototype was run on the test dataset that contains diverse cases of matching geometries. After running, the results suggest considerable improvements in the consistency of the dataset. A significant number of polygons was found to match, both within single-LOD and multi-LOD data. Furthermore, examining the results in details exposed that some of the rings reoccurred as many

as 9 times in the 6 representations, which reinforces the importance of linking such geometries.

The total number of polygons in the datasets is 13 021 (of which 8% has interior) with 17 029 rings. Indexing reduced the amount of queries by 99.88%. Most of the points in the index referred to less than 10 rings (i.e. the number of rings that share the same smallest point), and within each comparison, an average of 42% of rings were found to match. The geometries were linked, and the consolidated data has been stored.

A useful insight in the established links is that all interior rings in the dataset are found to be matched to an exterior ring of another polygon. This is due to openings such as windows and doors, and closure surfaces.

Obviously, the obtained results such as the number of matched primitives (88%), strongly depend on the used dataset. However, they give a good impression of the quantity of reoccurring geometries, for which there is no reason to store them more than once.

While the consolidation algorithm provides a good balance between simplicity and the end result (amount of consolidated data), its computational complexity is exponential, rendering it less feasible for larger datasets.

From the storage perspective, after the consolidation the size of the dataset was reduced by 20%. Due to the highly repetitive structure of an XML schema such as CityGML, we do not expect that the consolidation of recurring geometries can further contribute to the reduction of the storage footprint so further data compression of multi-LOD datasets should be sought in methods such as XML clustering [762].

8.2.4 Discussion

The first part of this chapter presents a method to analyse multi-LOD CityGML datasets by detecting matching geometries, and to automatically adapt them by storing the explicit representations of their topological relationships. This enhancement improves the consistency of the model, leading to more efficient maintenance and storage, and the experiments indicate significant improvements in that respect.

As a foundation of the work, we have developed a theoretical framework that describes cases of topological relationships occurring in reality, and we investigate more closely the XLink mechanism that provides the explicit modelling of some topological relationships in (City)GML. The work can be used for both within a representation and across multiple representations, doubling its purpose. Since our implementation is automatic, we hope that these improvements will contribute to the increased creation of multi-LOD datasets and their deeper integration.

While the tests involved only buildings, the presented algorithms are intended to work for other thematic classes as well. However, because of the lack of such datasets,

testing possibilities are limited, and had to be done on a synthetic dataset produced by our procedural modelling engine.

The matching algorithm is suited for ideal cases such as generalised datasets that contain identical primitives across LODs, and may fall short in datasets constructed with different acquisition techniques. Improvements in this direction are one of the aims that we plan for future work. Furthermore, we plan to improve the algorithm that consolidates the data and to offer a few alternatives with different advantages and suitability.

8.3 FOUR-DIMENSIONAL MODELLING

8.3.1 Introduction

Another solution to link two or more 3D representations is to integrate them in 4D space, following the research of Arroyo Ohori [763] on higher-dimensional geographic information, and related work in 2D [764, 765]. That is, the LOD is modelled as a geometric dimension perpendicular to the three spatial dimensions, in which a 4D object would comprise multiple linked 3D representations. In such a case a specific LOD could be obtained by slicing the hyper-dimensional construction at a particular point along the LOD axis, resulting in continuous LODs and a refined number of 3D representations. The method of linking multi-LOD datasets by storing them in a 4D construction consists of multiple steps, briefly discussed in the continuation. For more details on the methodology and results the reader is referred to the paper [741].

8.3.2 Identifying corresponding features

In order to establish a 4D structure first it is necessary to establish links between corresponding objects across multiple LODs, similarly as in the previous section on matching geometries, or according to a hierarchy such as the one presented in Figure 3.10 in the chapter on the LOD formalisation. This is akin to map matching in cartography [766], and it is impeded by similar challenges. For example, as discussed in Chapter 3 and Chapter 4, multiple LODs may not have corresponding features in such a way that a 1:1 relationship is possible. CityGML LOD1 and LOD2 are such a case, in which the roof surface of a building in LOD2 does not exist in LOD1. Another hindrance is the dimensionality metric as discussed in Chapter 3: the dimension of features across multiple LODs may not always be equal. As an example, a tree may be modelled in high detail in 3D, but in another LOD it may be modelled in 1D (as a vertical line).

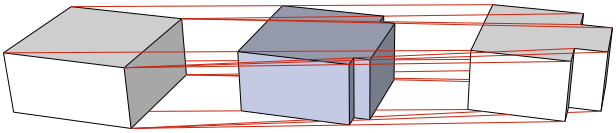


Figure 8.4: Two LODs of a house (LOD1.1 and LOD1.2) with differing geometry and topology are integrated into a 4D model by collapsing cells in the model at the highest LOD.

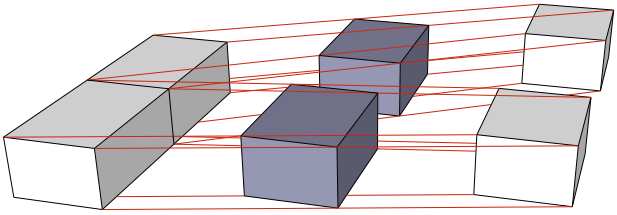


Figure 8.5: Two LODs (LOD1.0 and LOD1.1) of two houses being aggregated are integrated into a 4D model by modifying the topology of coarser LOD so as to match the topology of the model at the highest LOD.

8.3.3 Linking corresponding features

After identifying corresponding features, they may be linked with each link forming an nD object embedded in 4D space. The links may be established according to multiple schemes. Here two examples are given.

The first linking scheme involves collapsing features that do not exist in coarser LODs on related features. For example, edges may be collapsed to vertices: a dormer in LOD2.2 may be collapsed to a point of a roof surface in LOD2.1. This scheme is illustrated in Figure 8.4 showing two block models of a building, with a different feature (geometric) complexity. The 4D model was constructed by first matching the obvious correspondences (e.g. walls that are equal in both representations), and then by collapsing other faces based on their adjacency with the matched faces.

The second linking scheme (Figure 8.5) is based on modifying the topology to ensure that there is a mapping between the linked features. In the illustrated example, the topology of the coarser LOD is adjusted (the single volume is split into two) to match the one of the higher LOD, accommodating the links between the two representations.

8.3.4 Slicing

The 4D object may be sliced (i.e. its dimension may be reduced to 3D by intersecting the 4D object with another lower-dimensional object [767]) at a particular level to obtain a 3D model of specific properties. Owing to this reasoning, slicing inherently results into a higher number of ('intermediary') LODs, a possibly interesting method to refine multi-LOD data.

The two figures also illustrate the slicing at a point between the two original LODs. This approach enables continuous LODs, since moving the intersecting object along the LOD dimension enables slicing at any point between the coarsest and finest representation. Furthermore, the approach results in a smooth transition between the original LODs, something that is desireable in visualisation.

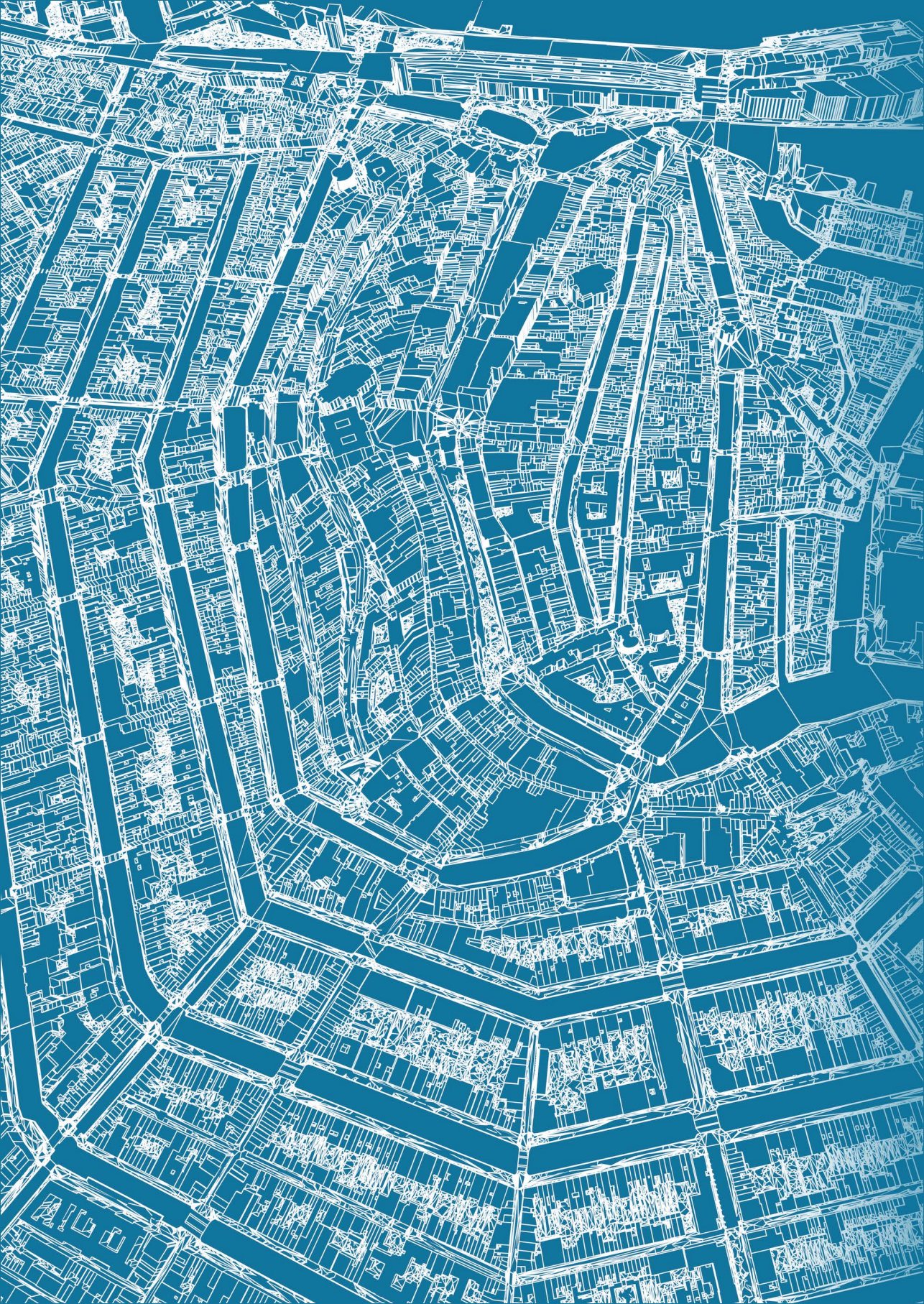
However, the derived representations may not be very useful in practice. Figure 8.4 indicates that the number of faces of the 3D model resulting from slicing is the same as the finer one, but from the quality point of view the model has a discrepancy to its real-world counterpart. Hence such representations may be of questionable value.

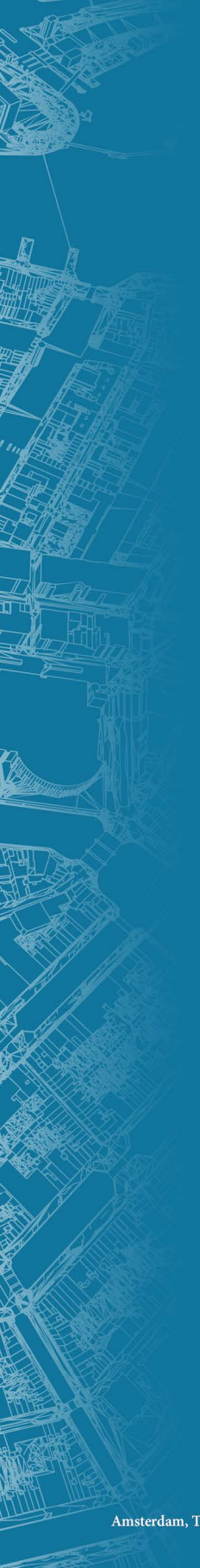
8.3.5 Conclusion

This theoretical methodology offers an alternative approach to integrate multiple LODs of the same region by storing them in a 4D construction. The LOD is modelled as the fourth spatial dimension, each feature having a set of 4D coordinates (x, y, z, l), where l is a specific LOD. While the approach offers a new perspective on integrating multi-LOD data, major challenges persist. For example, it is difficult to automatically store multi-LOD models in 4D, and the results suggest that these intermediary LODs may not be useful in practice.

Future work is desirable, with the following points proposed. First, linking corresponding geometries is a challenge because multiple LODs do not always have corresponding features that are straightforward to link. While the LOD framework presented in Chapter 3 enables a hierarchy of elements, it is not always clear how to link the geometries in the context of this method. Second, the examples of slicing the 4D construction shown in this chapter result in homogeneous models. However, allowing the slices to be non-orthogonal to the LOD axis may result in mixed-LOD 3D city models. It would be interesting to investigate whether it is possible modifying the slicing operator to produce application-specific LODs, i.e. representations that are suited for a specific application. For example, when using 3D city models for pedestrian navigation, details that are far up from the ground (e.g. roof details) are not important and could hinder the performance of the application. Hence, a customised slicing operator deriving a 3D model where buildings would have more

details closer to the ground could be developed. Third, while the above examples focus on the geometry, 3D city models include other aspects such as attributes and texture. It would be interesting to investigate how to link such data, and how the results of the slicing operator would look like.





III

Experiments and uncertainty

Amsterdam, The Netherlands. Generated with 3dfier using BGT and AHN datasets.

CHAPTER 9

Influence of LOD on spatial analyses (I)

This chapter is based on my paper [768]:

> Biljecki F, Arroyo Ohori K, Ledoux H, Peters R, Stoter J (2016): Population Estimation Using a 3D City Model: A Multi-Scale Country-Wide Study in the Netherlands. *PLOS ONE*, 11(6): e0156808. doi: 10.1371/journal.pone.0156808

and partly on my paper [769]:

> Biljecki F, Ledoux H, Stoter J (2017): Does a finer level of detail of a 3D city model bring an improvement for estimating shadows? *Advances in 3D Geo-Information*, Springer International Publishing, pp. 31–47. doi: 10.1007/978-3-319-25691-7_2

Finer LODs allegedly bring more accurate results in a spatial analysis, at the expense of a higher cost of procurement. However, the performance of different LODs in a spatial analysis has not been investigated thoroughly. This chapter investigates this perennially repeated hypothesis by performing the same spatial analysis with multiple datasets of the same area in different LODs. First we give an overview of related work in 2D and 3D GIS that analyses the use of data in different scales. The spatial analysis selected for experiments is the estimation of the population with 3D city models, leading to a census without field surveys. Based on the assumption that housing space is a proxy for the number of its residents, we use two methods to estimate the population with 3D city models in two directions: (1) disaggregation (areal interpolation) to estimate the population of small administrative entities (e.g. neighbourhoods) from that of larger ones (e.g. municipalities); and (2) a statistical modelling approach to estimate the population of large entities from a sample composed of their smaller ones (e.g. one acquired by a government register). Population is estimated with 9 datasets of different spatio-semantic LODs. The work reveals that in this spatial analysis both the semantic and geometric detail have an influence on the accuracy of the results. A byproduct of this research is that it advances the method of population estimation with GIS techniques, and that it gives a comprehensive overview of this 3D use case.

9.1 INTRODUCTION

In the initial chapters of this thesis it was argued that the concept of level of detail in 3D city modelling was borrowed from computer graphics. In computer graphics the LOD concept is used to balance the computational complexity and the visualisation quality (fidelity). The latter, i.e. how similar the object looks like to the original one, can be assessed with error metrics such as the deviation between the geometry of the original model and the geometry of its simplified counterpart [400].

Such rationale can be applied to geoinformation science, in order to analyse whether it is worth to invest funds and computational resources in dataset of a particular LOD. In appropriate scale ranges [770], data modelled in finer detail are inherently believed to benefit spatial analyses, at the expense of an increased cost of acquisition, storage complexity and maintenance, and hindering the speed of spatial analyses. Hence, the benefit may not always justify the investment.

Experiments that we present in this chapter are relevant for understanding how different LODs affect the accuracy of a spatial analysis. This has a twofold meaning. First, as modelling data at finer LODs comes at a higher cost, a relevant question is whether a certain spatial analysis can take advantage of this finer detail. Second, the problem may be approached from the generalisation perspective; 3D geoinformation at fine LODs may in fact be too complex for certain spatial analyses. Hence, the data is occasionally generalised to reduce complexity while attempting to preserve usability [771]. Insights into the performance of the LODs may help to achieve that balance: by generalising the models to a point at which their complexity is sufficiently reduced but at the same time to ensure that their usability is not compromised by the reduced LOD.

However, in GIS, and particularly in 3D city modelling, this topic has not been investigated thoroughly. In this context, one of the main differences between GIS and computer graphics is that the models are often used beyond visualisation for several purposes that significantly differ from each other—they have different behaviour, different requirements for LODs, and an outcome of different nature. For example, 3D city models are used to estimate the real estate net internal area in m^2 [129], energy demand of buildings in kWh/year [78], and noise in dB [259], resulting in different kinds of errors, differently influenced by a particular LOD.

In this chapter we carry out experiments to determine the influence and performance of different LODs performed for the same spatial analysis. The method that we employ is predictable and straightforward, consisting of three steps:

1. Sourcing data of the same spatial extent in multiple LODs.
2. Running the same spatial analysis for each LOD.
3. Comparing the results between LODs.

While the method looks straightforward, each step is not without obstacles. Starting from the first step, obtaining data in multiple LODs is difficult because most of the data is produced at a single LOD, leaving multi-LOD datasets not widely available. In the context of this research it is also difficult to ensure that a particular dataset is acquired strictly according to an LOD specification.

Running a spatial analysis entails multiple problems. First, not all spatial analyses are suitable for experiments because the outcomes of only a subset of analyses result in a quantifiable result, which is a prerequisite for error analyses as it provides a measure to compare results. For example, using 3D city models for analyses listed earlier in this section ends up with quantifiable values, but using them for urban planning and many other purposes does not result in figures that can be objectively assessed.

Furthermore, several spatial analyses derive results that are subject to different interpretations. For example, 3D city models may be used for different kinds of visibility analyses, and therefore quantified in different ways: binary (a point in space is visible or not), distance (range) of visibility, the area or volume visible from a point, number of buildings that have visual access to a feature, and population that has visual access to a point [677, 772–774]. Each one might have different behaviour. Second, from the implementation aspect, it is not always easy to automatically run a spatial operation for a large dataset and extract results in a format that is suitable for error analysis. For example, the shadow and visibility analyses presented in Figures 1.2 and 2.4d, respectively, result in a visual impression (a rendered raster image), rather than a vector dataset that can be further analysed (e.g. to calculate the area of the shadow and the visible volume).

The last step—comparing the results of a spatial analysis—may not always be straightforward as well. Spatial analyses that have quantifiable outcomes may lead to different conclusions because of different interpretations of the error. This ambiguity will become clearer in Chapter 10 on the case of predicting shadows cast by buildings, as there are different ways to quantify shadows and their errors. Furthermore, if multiple spatial analyses are compared, it may be complex to put their results in the same context.

The spatial analysis that appears suitable in this context and that is used for the experiments in this chapter is the estimation of population using 3D city models for refining census maps. Besides addressing the influence of different LODs on the accuracy of a spatial analysis, this chapter also analyses the use case of population estimation in depth, and it gives a few contributions adding to the body of knowledge in this topic.

9.2 RELATED WORK: INFLUENCE OF DATA GRANULARITY ON SPATIAL ANALYSIS

9.2.1 2D GIS

Studies on the influence of the scale, LOD, and resolution to the quality of a GIS operation are focused on 2D and to raster representations. Hengl [775] discusses the importance of considering the resolution of a raster, and underlines that in GIS projects the resolution is usually selected without any scientific justification. Usery et al. [776] determined the resolution effects on watershed modelling by resampling input rasters, and concluded that the resolution has a significant effect on the accuracy of the result. Guo-an et al. [777], Booij [778], Chaubey et al. [779], Ling et al. [780], and Pogson and Smith [781] performed similar analyses with similar results. For example, the study of Chaubey et al. [779] indicates that the resolution of DEMs affects the output of a hydrologic spatial analysis.

There has been a considerable amount of research on the influence of different representations in cartography and remote sensing across multiple scales [611]. For example, Veregin [782] and Cheung and Shi [783] study the effect of the simplification of lines (e.g. roads) in maps and their propagation to positional displacement.

In another relevant study, Ruiz Arias et al. [784] estimated the solar irradiation of several locations with DEMs of different resolutions (100 m vs 20 m grid). The results demonstrate that there is an improvement in the results with the improvement of the resolution of the DEMs, but that it is minuscule in comparison to the error induced by the spatial analysis.

9.2.2 3D GIS

In 3D GIS such studies are uncommon. A possible reason is that 3D city models cannot be easily 'resampled' in the same way as rasters to easily obtain multiple LODs for analysis.

The influence of different representations has mostly been evaluated in urban planning and related domains in which the visual impression is the main decisive factor [235, 495, 570, 785]. For example, Kibria et al. [282] survey the perceptual value of a few LODs in spatial planning. It turns out that in some planning phases a finer LOD is actually undesirable.

In the past two years there have been a few related analyses comparing the results of a spatial analysis utilising data of the same area modelled at different LODs [209, 560, 786]. While the results are mixed, most researchers appear to agree that having finer LODs may even be detrimental, as the potential benefit may be countervailed by cost and complexity [787]. Brasebin et al. [227] tested the fitness for use

of LOD1 and LOD2 models of two French datasets in the determination of the sky view factor (SVF). The result is that for 75% of the analysed samples the improvement in accuracy of the predictions is less than 2%. Besuievsky et al. [231] carry out a similar study with SVF focusing on the windows of buildings, with a few variants of high-detailed architectural models, and find a significant difference in the results.

The main shortcomings of such analyses are that most of these analysis are performed on only two LODs, and on real-world data, thus preventing the focus on the LOD-induced error alone. Moreover, some of the datasets used, are generated from different sources, containing different magnitudes of errors. Some of the datasets are quite small (e.g. containing merely a few buildings) limiting the conclusions to a specific architectural type.

Other mentions of the influence of the LOD are merely anecdotal, and they differ. For example, Strzalka et al. [114] investigate the use of 3D city models for forecasting energy demand, and argue that the suitability of an LOD depends on the configuration of buildings (i.e. for an area with predominantly flat roofs they suggest that an LOD1 suffices). However, experiments with other LODs are not documented. In an energy demand estimation, Malinverni and Tassetti [466] note that with a finer LOD the results of their analysis would improve, but there is no proof for such a statement. In a similar analysis, Barbano and Egusquiza [788] note that an LOD2 is more than sufficient. The same goes for the paper of Zwolinski and Jarzemski [490] dealing with a shadow analysis. Brédif [789] states that LODs higher than LOD1 are not necessary for navigation. Fai and Rafeiro [790] mention that in building operation and maintenance, 'a lower LOD may in fact be more desirable'.

9.3 USE CASE: POPULATION ESTIMATION USING 3D CITY MODELS

GIS and demography have long been closely related, and GIS techniques are ubiquitous in mapping, analysing, and filling gaps in demographic data. In particular, geostatistical techniques are often used to estimate a region's population in the absence of reliable or complete census data [791, 792].

2D GIS datasets (e.g. satellite imagery and maps) have been used extensively in the past 50 years for this purpose, as several of them have been found to be reasonable proxies for population [334, 792–810]. For example, Bakillah et al. [739] estimate the population based on the concentration of surrounding points of interest (e.g. restaurants); Anderson et al. [811], Sutton [812], and Doll et al. [813] use night-time imagery following the hypothesis that city lights indicate the magnitude of the urban extent, which in turn indicates the population. Pozzi and Small [814] infer the population density from a vegetation cover map, based on the idea that less vegetation means more people; Xie [815] finds the relationship between the density

of road network and population; Steiger et al. [816] analyse georeferenced Twitter data to locate clusters indicating home- and work-related social activities that can serve as a proxy to estimate the residential and workplace population census data; and Lwin et al. [711] present a similar work using geolocated mobile phone usage data.

Among all these methods, many successful approaches rely on 2D datasets (maps) containing *building footprints* (e.g. derived from cadastral records or satellite imagery). The most straightforward approaches rely on the total number of buildings in a region or the total area of building footprints in it [330, 336, 817]. These methods perform reasonably well in homogeneous areas, but they exhibit significant errors in areas where buildings have a great variation in the number of storeys.

With the advancement of remote sensing technologies, such as lidar and aerial photogrammetry, it is now possible to automatically and remotely measure the height of a building, which can be used to obtain a volumetric representation of a building (3D city model) that is useful for population estimates. In fact, several researchers have indicated that the volume of buildings and the floorspace provide a strong cue for its population [79, 329–335, 337, 339, 715, 739, 804]. For example, Lu et al. [79] use multiple regression models to perform a study in Denver, Colorado, based on both footprint areas and building volumes. Lwin and Murayama [329] and Alahmadi et al. [338, 715] estimate the number of floors from an elevation dataset, and multiply it with the footprint area to get the approximate internal area of the apartments. Their results indicate that the volume-based approach gives more accurate results than the area of the footprints due to heterogeneous building morphologies.

However, despite the frequent indication that volume-based methods can improve on the estimates of area-based methods, there has been no large-scale study that conclusively proves that this is true. Existing studies have several gaps: they usually focus on single metropolitan areas, which can be relatively homogeneous; they seldom compare the accuracy of different approaches within the same region; they derive a building's volume based on a raster dataset, which limits its accuracy; they do not consider how this approach scales between larger and smaller areas; and they do not consider how the level of detail of the used volumetric representation affects the accuracy of the result.

The goal of this research is to bridge these gaps. We investigate *to what extent 3D city models can be used to estimate the population of a region* by performing a multi-scale country-wide study in the Netherlands. As the Dutch government provides both highly accurate census and building data, we consider that the Netherlands serves as an excellent case study, both for the experiments and the validation of the methods.

We therefore evaluate the use of 3D city models in population estimation in two directions: (1) disaggregation (areal interpolation) to estimate the population of

small administrative entities (e.g. neighbourhoods) from that of larger ones (e.g. municipalities); and (2) a statistical modelling approach to estimate the population of large entities from a sample composed of their smaller ones (e.g. one acquired by a government register). We compare the population estimates obtained by both methods with the actual population as reported in the census, and use it to evaluate the quality that can be achieved by estimations at different administrative levels. We also analyse how the volume-based estimation enabled by 3D city models fares in comparison to 2D methods using building footprints and floor areas, as well as how it is affected by different levels of semantic detail (information on building use) in a 3D city model. We conclude that 3D city models are useful for large scale estimations (e.g. for a country), and that the 3D approach has clear advantages over the 2D approach.

In the context of this thesis, this is essentially an analysis involving datasets of different LODs—LOD0, LOD1, and indoor data. The latter has not been subject of this research, but it is nevertheless an interesting addition worth taking into account because such data is available in this case, and it has not been a subject of related work. Because several different methods are used to estimate the population, it is also interesting to investigate whether the used LODs achieve consistent results.

9.4 DATA AND METHODOLOGY

9.4.1 Census data

The Netherlands is decomposed hierarchically into 12 provinces (not covered further in this study), 393 municipalities, 2816 districts, and 12237 neighbourhoods. The population of each of the entities is known thanks to the open dataset of Statistics Netherlands—CBS (*Centraal Bureau voor de Statistiek*). As illustrated in Figure 9.1, the dataset consists of sets of polygons representing statistical units—the population within each polygon is stored as an attribute. We use this dataset to validate our results, and its subset to train one of the methods. The properties of statistical units across the country vary (see Figure 9.2), covering widely heterogeneous household sizes, population densities, and dwelling sizes, among others.

9.4.2 3D city model of the Netherlands

The supposed key advantage of using 3D city models over 2D maps is that they provide volumetric data, which is beneficial for applications that take advantage of the height or volume of buildings. Population estimation is clearly such a case, as high-rise residential buildings are very likely to contain more inhabitants per unit area than low-rise buildings.

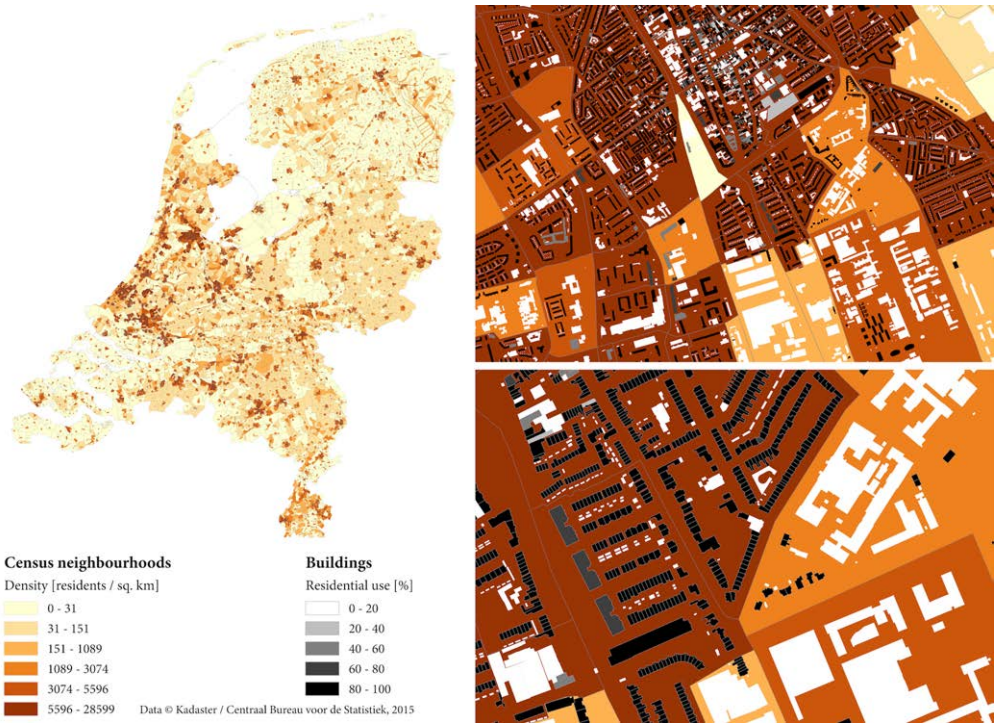


Figure 9.1: Datasets used in this research: census neighbourhoods with building footprints. (Left side:) The Netherlands divided into more than 12 thousand neighbourhoods; and (right side:) two zoomed-in urban areas, where building footprints are visible along with the information on their use (residential share). Note that the maps on the right side reveal large variations in population density despite neighbourhoods being similarly urbanised. The less populated areas have many non-residential buildings, e.g. industrial and university buildings, showing that information on their use is crucial, and it significantly impacts the quality of the population estimation. The population density classes are divided into quantiles.

In this study, we generate a country-wide 3D city model in LOD1.2 by combining two open datasets from the Netherlands government: (1) building data from the national register of addresses and buildings (BAG—*Basisregistraties Adressen en Gebouwen*, which is collected and maintained by each municipality, and disseminated as country-wide dataset through the national portal of *Kadaster*, the national mapping agency of the Netherlands; and processed by the NLExtract project)—containing the base geometry, building use, and floorspace information (see Fig-

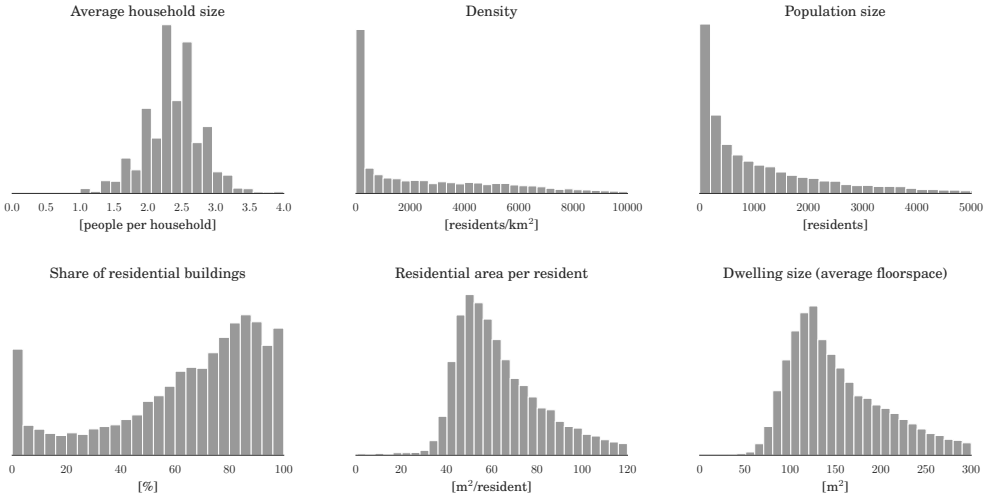


Figure 9.2: Census neighbourhoods statistics. The plots reveal substantial housing differences among the neighbourhoods across the country.[27]

ure 9.1); and (2) elevation data—the Height Model of the Netherlands (AHN—*Actueel Hoogtebestand Nederland*), which contains 639 billion elevation points [818], covering the whole country at the accuracy of a few centimetres [819]. The same combination has been used for the study presented in Chapter 7 on inferring the heights of buildings, and for other purposes in the thesis (e.g. for some figures in Chapter 1 and Chapter 2).

In order to test how different LODs of a 3D city model affect the population estimations, we construct 9 different LODs using various combinations of different LODs in a building's geometric and semantic information (Figure 9.3). In this way, we can directly compare the quality of the estimations given by the area-based (footprint and floor area) and volume-based approaches.

We consider three geometric LODs: (LOD0) 2D building footprints (the traditional area-based approach without height measurements); (LOD0+) building floorspace (area-based approach in which the vertical extent of the building is available); and (LOD1) volumetric 3D block models (from which the volume of a building can be calculated). The latter is comparable to LOD1.2 as defined in Chapter 4, since the footprints provided by Dutch authorities are quite detailed and match our specification [708].

[27] Derived from the data from Kadaster / Centraal Bureau voor de Statistiek, 2015.

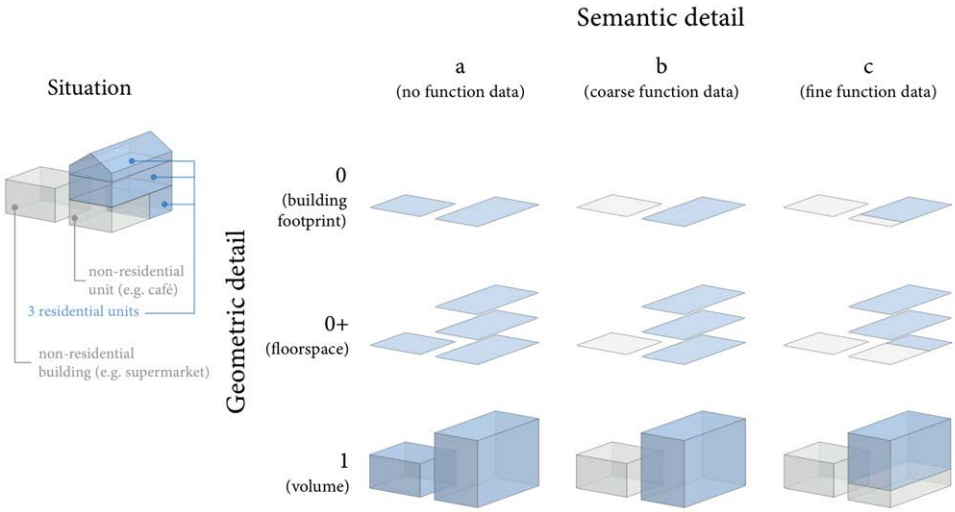


Figure 9.3: Multi-LOD data used for the experiments. Different granularities, which reflect the different grades of data available in practice. The coloured space indicates residential space (proxy for population) as considered for each LOD, which differs depending on the geometry and semantics, and ultimately affects the performance of the methods. In our work we benchmark the performance of each grade of the data for the purpose of estimating the population.

On the other hand, indoor has not been a focus of this thesis. However, in order to test how the volumetric approach fares in comparison to the area-based approach, we use indoor data that we dub as LOD0+, as introduced above. For LOD0+, we rely on accurate indoor measurements from the Dutch cadastre, which is a dataset that is rarely available elsewhere. However, it should also be noted there is recent work focused on its automatic reconstruction [129, 130].

The general hypothesis used in this chapter, and in related work, is straightforward: the larger the building, the more people reside in it; and the larger the living capacity of a district, the more populous it is. However, we argue that other building properties should be taken into account as well. The occupancy of a building also depends on its type, e.g. a cathedral, indoor arena, or a factory can be very large but at the same time they house zero inhabitants. Therefore, only residential buildings must be taken into account. This is further complicated by mixed-use buildings, which are composed of non-residential and of residential units, e.g. a three-storey building, where the ground floor is occupied by non-residential space (e.g. a restaurant and a shop), and the remaining two floors by residential units (fairly common

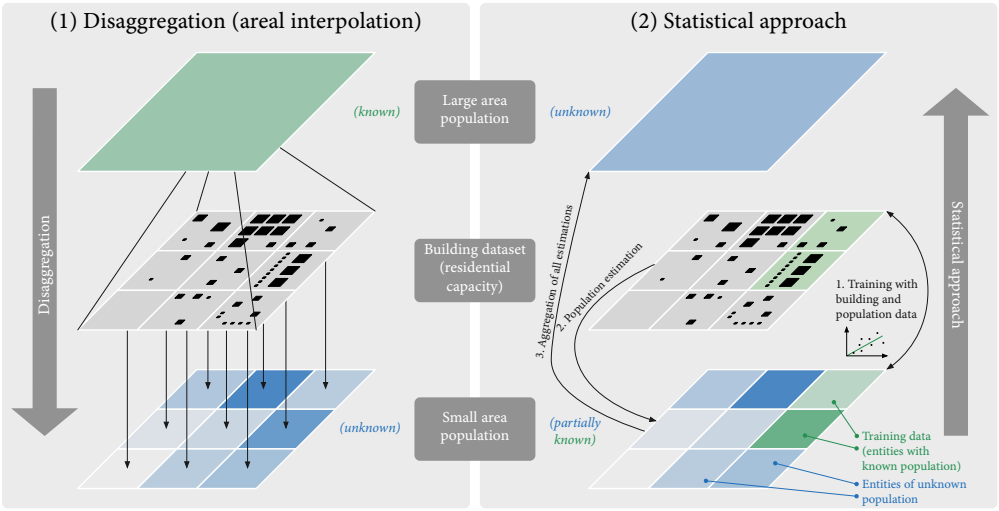


Figure 9.4: The two population estimation methods used. In this study we employ both methods, and for the residential capacity we use three different indicators in parallel: building footprint area, floorspace area, and building volume. Our work determines the usability of each of the type of geographic information for this purpose.

in the Netherlands). However, such information is not always available, hence we pay special attention to the semantic aspect of data. Therefore, for the semantic part, we distinguish three LODs: (a) no data about the function of the building, and hence all buildings are treated equally; (b) a building is either residential or non-residential [331, 737]; and (c) fractional building use, where the share of the residential use within a building is known.

The possible combinations of the three geometric LODs and the three semantic LODs result in the 9 LODs used in this study, e.g. LOD1^{b} denotes a block model with the singular information on the building use.

9.4.3 Existing methods for population estimation

The estimation of population with GIS data and techniques has been extensively reviewed by numerous authors [328, 332, 739, 820, 821]. Generally two groups of methods are recognised [328], both of which are used in this chapter (Figure 9.4):

1. Disaggregation (areal interpolation): this is a top-down approach where the population of a larger administrative unit or zone (e.g. region, municipal-

ity, census district) is distributed across smaller units (e.g. neighbourhood), usually by weighting it according to different factors that hint at the population [822–827]. This approach is typically used when the population of a large entity is known (e.g. a city), but the one of its composing entities is not known (e.g. its neighbourhoods). Hence disaggregation inherently refines population maps (i.e. it increases their level of detail).

The disaggregation can be done by simply distributing the population among administrative sub-zones, but it also can be aided by dasymetric mapping to shape smaller surfaces in such a way that variation within each surface is minimised [828]. This is especially useful when the smaller units are political subdivision of the larger (parent) unit often found in choropleth maps (e.g. interpolation from a province to the containing municipalities), because such regions may contain variations in the population density. Dasymetric mapping therefore results in (sub-)units that are more homogeneous [829–832].

2. Statistical modelling approach: first the relationships between population and socio-economic and morphological variables associated with the population density are inferred, e.g. land use [798, 833], proximity to transportation network [820], and distance from the central business district [834]. The deduced relationships are then applied to estimate the population count of unknown areas. In this approach multiple linear regression is most commonly used. The advantage of this bottom-up approach is that a sampling census has to be carried out for only a small area. It is useful in the scenario when only the population of a subset (e.g. a city) of a large area (e.g. a province) is known.

9.4.4 Our proposed method using 3D city models

For our population estimation study, we test three indicators to determine the disaggregation weights and the statistical relationships: (1) area of the 2D building footprints (in m^2), (2) area of the building floorspace (in m^2), and (3) building volume (in m^3). Each of these is tested at three levels of semantic detail, resulting in the 9 aforementioned LODs of the input datasets.

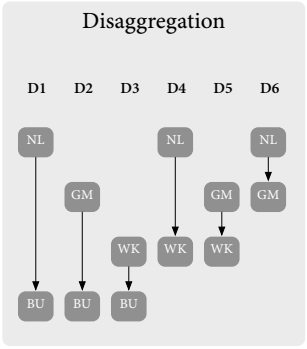
In order to diminish residential and socio-economic variations across a large area, but also to test the performance of different estimation scenarios, we use multiple scales of estimations, as shown in Figure 9.5. In the disaggregation approach 6 scales are analysed:

D1 Disaggregation from the country level to its 12237 neighbourhoods.

D2 Disaggregation from each of the 393 municipalities to their 12237 neighbourhoods.

Netherlands administrative hierarchy
(scale of estimations)

Administrative unit	Median pop.	Median no. of buildings	Median area [km^2]
The Netherlands	16.9M	9.9M	41 540
393 municipalities (*gemeente*)	26 190	18 418	73
2 816 districts (*wijk*)	3 237	2 180	7.7
12 237 neighbourhoods (*buurt*)	655	466	0.8



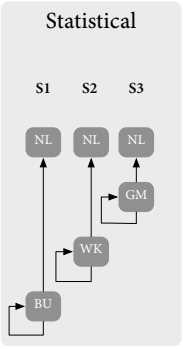


Figure 9.5: The Dutch statistical hierarchy, and our hybrid multi-scale approach. The hybrid approach refers to both the disaggregation and statistical approach, while multiple scales refer to the level of the statistical units.[28]

D3 Disaggregation from each of the 2816 districts to their 12237 neighbourhoods.

D4 Disaggregation from the country level to its 2816 districts.

D5 Disaggregation from each of the 393 municipalities to their 2816 districts.

D6 Disaggregation from the country level to its 393 municipalities.

On the statistical side we use a random subset of 10% of each statistical level to determine with ordinary least squares the relationships between building space and population, and apply them for three different experiments:

S1 Estimation of the population of the test neighbourhoods (i.e. the remaining 90%).

S2 Estimation of the population of the test districts (i.e. the remaining 90%).

S3 Estimation of the population of the test municipalities (i.e. the remaining 90%).

Furthermore, in the statistical approaches (S1, S2, and S3) we also estimate the population of the Netherlands. This means that we test the suitability of carrying out the census for 10% of the country (training dataset), and estimating the population of the rest of a country (test dataset).

[28] Statistics of the units obtained from data of Kadaster / Centraal Bureau voor de Statistiek, 2015. The provinces are not shown here because they have not been considered in our work, and the data refer to the situation in 2015.

In each of the 9 approaches we carry out separate experiments with the data in the 9 different LODs. This results in a total of 108 experiments.

As in related work [329, 835], we ignore very small buildings (footprint smaller than 20 m^2) such as sheds, garages, etc. which are unlikely to be inhabited (visible in Figure 9.1 as tiny white footprints in the predominantly residential areas).

9.5 RESULTS AND DISCUSSION

9.5.1 Performance and observations

We perform the experiments, and compare them to the actual values, as observed in the governmental census dataset (CBS). We use percentage error because we are dealing with different scales of data (e.g. an error of 1000 residents is not of the same magnitude on the neighbourhood or city level). Furthermore, because of large errors in some statistical units (explained later), instead of the usual mean absolute error and RMSE we use the median absolute error. As in related work [820, 836], we observe that estimations in areas with small populations is prone to a high relative error (see Figure 9.6), hence medians are a good option here. The two histograms in the figure show the data divided in two bins (the left one of the statistical units with the population smaller than the median value of all units (710 residents), and the one on the right the units with the population higher than the median). This plot should not be confused with Figure 9.8, which illustrates the relation of errors to the population density (however, notice that in this case as well the methods tend to underestimate the population in more populated areas).

The results of all experiments are given in Table 9.1. Because of many different models and types of data, we focus on the most important results only, however, the elaborated observations are similar with the rest of the models. It should be noticed that both the disaggregation and statistical approach exhibit congruent behaviour in most cases.

The results exhibit a large degree of variation between the accuracy depending on the approach, LOD of the data, and the scale of the estimations. The smallest error of the volume-based disaggregation approach is in D5/LOD1^b (the disaggregation from municipalities to districts) and it equals 11.8%. The smallest error in the statistical approach was observed in S3 (estimation of the population of cities), resulting in an error of 9.3%. We observe and conclude the following:

- 3D city models (LOD1) and the volume-based approach provide a substantial advantage over traditional 2D geo-datasets (LOD0) and the area-based approach because they capture the vertical extent of the building. However, the estimations carried out with 3D models are still less accurate than when

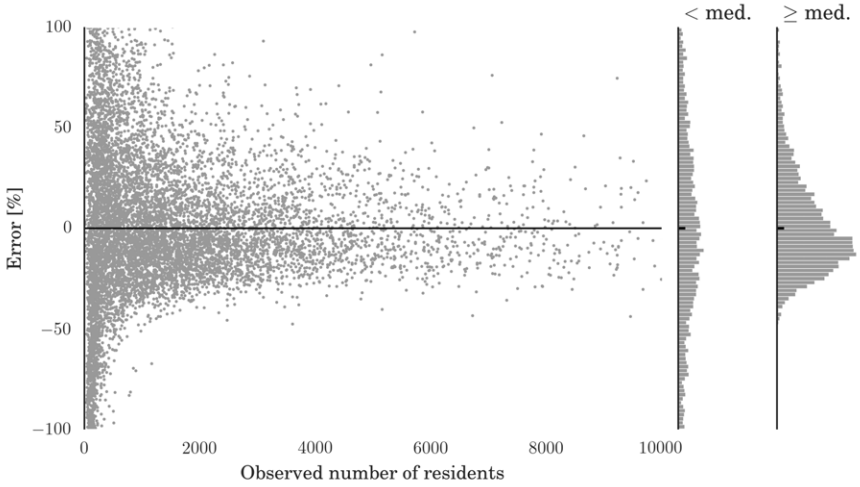


Figure 9.6: Less populated districts exhibit large relative errors, promoting the use of medians. In relative terms, the estimation is more accurate when carried out in more populous areas. These are the results from the experiments S1/LOD1[c].

using floorspace information (LOD0+). We think that volume does not add value over floorspace because two flats of the same floorspace but of different volumes (e.g. ceiling height of 2.5 m vs 4 m) generally do not host a different number of residents, unlike what the method would predict. It should be noted, however, that floorspace information is difficult to acquire automatically and it is generally not available.

- In most cases, semantic information on the use of buildings provides a substantial improvement in the estimations over data without such an information. This helps to exclude non-residential units, which can significantly skew the estimations. Such a behaviour is visible as outliers in the scatter plots in Figure 9.7 (other observations will be discussed in the continuation). Population estimation without information on the building function is practically unusable in most cases, especially in industrial neighbourhoods (in our experiments we have seen overestimations of more than 5000%). In fact, the results reveal that in this use case, semantic detail is typically more important than the geometric detail (e.g. cf. error of 41.9% in D1/LOD0[b]—semantically enriched 2D footprints vs error of 56.4% in D1/LOD1[a]—plain 3D buildings).

Table 9.1: Median absolute percentage errors in the population estimates resulting from our experiments. The order of errors in each 3×3 matrix is expressed in the same order as the LODs in Figure 9.3.

	a	b	c	a	b	c	a	b	c
	(1) Disaggregation								
	D1 (n=12237)			**D2 (n=12237)**			**D3 (n=12237)**		
0	61.9	41.9	42.4	53.9	25.5	25.7	42.7	17.7	17.7
0+	39.8	20.8	20.8	37.2	16.2	16.4	29.1	12.0	12.0
1	56.4	25.5	25.8	53.0	20.8	20.7	42.4	15.6	15.3
	D4 (n=3237)			**D5 (n=3237)**			**D6 (n=393)**		
0	56.5	37.7	38.2	34.3	15.5	15.5	32.0	25.3	25.5
0+	25.8	16.9	16.5	21.3	9.3	9.2	13.2	11.5	11.4
1	43.5	20.0	20.5	32.0	11.8	11.9	22.1	13.2	13.2
	(2) Statistical approach (local units)								
	S1 (n=12237)			**S2 (n=3237)**			**S3 (n=393)**		
0	85.4	42.0	42.2	56.9	53.1	52.8	74.0	38.7	38.8
0+	35.4	18.3	18.5	41.5	28.9	28.5	20.6	12.6	12.2
1	66.8	24.3	24.8	49.8	26.1	28.6	28.9	9.5	9.3
	(2) Statistical approach (country level)								
	S1 (n=1)			**S2 (n=1)**			**S3 (n=1)**		
0	0.6	1.4	1.4	2.7	5.6	5.7	21.5	1.3	1.7
0+	9.3	2.0	2.2	2.7	0.6	0.5	7.9	1.9	1.9
1	4.1	1.2	1.3	3.1	1.9	1.9	11.7	2.0	1.8

- While semantic data is crucial, it appears that there is inconsistent added value of the detailed (fractional) semantic information versus only binary information. It appears that the difference between binary and fractional semantic information becomes negligible at the neighbourhood level. In fact, in some estimations (e.g. D4/LOD1) the estimations with fractional semantic information (D4/LOD1^{c}) are slightly less accurate than when using binary semantic information (D4/LOD1^{b}).

  In the floorspace data (LOD0+) there is generally a small improvement of using fractional semantic information rather than binary. A possible reason is that the volume-based estimations are more sensitive to errors in the input dataset.

  For the purposes described in this chapter, it does not appear worthwhile to collect detailed building usage, as the binary information suffices. Because such information may be automatically derived from the building morphology, aerial imagery, and land use maps [132, 327, 620, 666, 712, 719, 837],

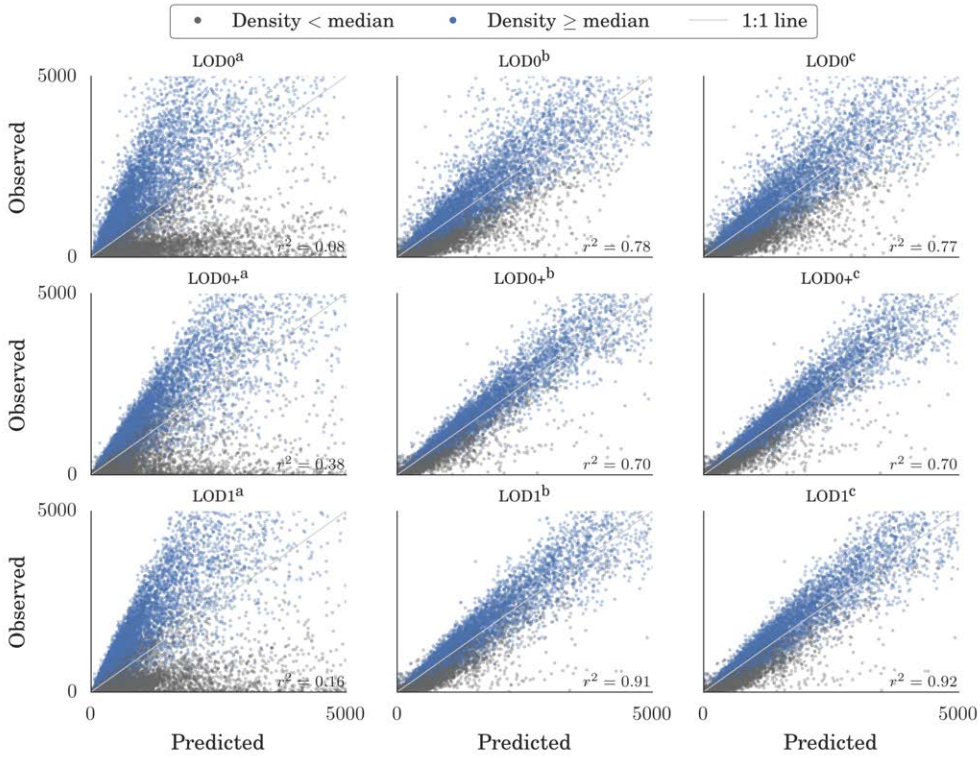


Figure 9.7: The performance of the estimation models depends on the population density of the target area. The plots reveal observed (actual data from the government census) vs predicted values of the 9 input datasets in the D1 method. The lower density refers to areas with the population density lower than the median of all neighbourhoods, and the higher those areas that are denser than the median, indicating urbanised areas. Notice the outliers in the estimations ([a]) that do not take advantage of the semantics—those represent highly industrialised areas without inhabitants or with sparse population. Furthermore, in the experiments carried out with fine-grade data most of the outliers are caused by input data (e.g. mislabelled residential use of a non-residential building) and by districts in which housing standards highly deviate from the average.[29]

this insight is beneficial for estimations that need to be carried out on a large extent where cadastral data is not available.

- Different scales of estimations exhibit different performance and different suit-

[29] Observed data courtesy of Centraal Bureau voor de Statistiek, Den Haag/Heerlen, 2015.

ability for the different methods. In the disaggregation, the method works best in hierarchically close units: compare D3 (districts to neighbourhoods) with D1 (country to neighbourhoods). This is because such relations exhibit less difference in housing variations. Furthermore, it appears that disaggregating data to units higher in the hierarchy is more accurate than to units of a finer scale, because larger units such as districts and municipalities capture larger residential differences than small neighbourhoods, i.e. the variation among smaller units is greater than that among larger units. For example, two municipalities may have equal population but within municipalities the population differences among districts may be relatively large (e.g. rural vs urban zones). On the other hand differences among neighbourhoods in a district may be small.

- The statistical approach is of comparable accuracy to the disaggregation because it is also based on coefficients uniform for the whole country, which hide massive disparities among different neighbourhoods and provinces, and it is therefore equally affected by the differences in living standards and residential choices.

  However, for the largest extent (country), the statistical approach is impressively accurate: in the S1/LOD1^{b} experiment (statistical approach applied on neighbourhoods with the semantic volume-based LOD1 block model) the population of the Netherlands based on a subset of 10% neighbourhoods has been estimated to 17 100 292, just a 1.2% overestimation from the true figure. The floorspace-based (LOD0+^{b}) data fares even better with a deviation of 0.5% in S2. This finding gives confidence in the use of 3D city models for estimating the population of large areas such as countries, especially in developing countries since the data required for such estimations can be derived automatically and remotely from airborne sensors [838]. However, it should be noted that the model S1/LOD0^{a} (building footprints without information on the building use) performed best with an error of 0.6%. It is hard to explain the reason why in this particular model lesser data gave better results, because all errors (induced by different LODs, uncertainty in the input data, different residential choices, etc.) are aggregated in a single number that cannot be decomposed. Shortcomings such as this one will be subject of the remaining chapters of the thesis.

- We have noticed that the models tend to overestimate the population in rural areas, and underestimate it in urban areas (see the coloured points in Figure 9.7). This finding is similar to the observations in related work [335]. The differences are caused by the varying utilisation of living space, which differ

between less and more densely populated entities. We use this finding in the succeeding sections for additional insights and we take advantage of it to improve the statistical approach (models S1, S2, and S3).

9.5.2 Sources of error

After analysing the errors we observe different causes of errors. The residential differences (e.g. residential space per resident) is the principal cause of the residuals (the errors very strongly correlate with the average space per resident; $r > 0.99$). There is a variable level of occupancy and variable utilisation of space within each building, i.e. living space per inhabitant considerably varies based on social, economical, and other factors. Some households live in large houses, while others in small studios and dormitories, rendering significant differences in the residential density [339], and presenting a problem for population estimation with remotely sensed data [715]. Furthermore, these differences are also caused by non-residential space within residential units, such as storage rooms, utility rooms, common rooms, gyms, garages, etc., which increase the building size and considered dwelling space, but due to the shortcomings of the data cannot be accounted as non-residential space. It is usually not possible to assume that these characteristics are equally distributed in each entity, as they are not constant among different neighbourhoods and also on larger extent such as among municipalities [130, 839]. This fact is also visible in Figure 9.2. Therefore it is important to consider different environments when calibrating the method, and accept imperfections as one model cannot fit all situations within a large area such as a province or country.

We had expected that these differences would cancel out within the statistical entities (since one typically contains hundreds of houses, see Figure 9.1), however, the difference between units, including larger ones such as cities, is still gross. One would assume that a city contains a fair diversity of different configurations, but it turns out that each city has a unique setting, which cannot be applied to another one.

Furthermore, another variation of the dwelling density is caused by vacant residential buildings (e.g. empty houses for sale, vacation homes). In our method we can only assume that the vacancy rate is homogeneous in our area of study, consistently with other researchers (e.g. when estimating the energy demand [121]), however, that assumption might deviate from the reality.

When using the data without information on building use (i.e. LODx^{a}) many large errors were found in industrial neighbourhoods with huge building volumes, highlighting the importance of using semantics. When using the semantically enriched buildings, the results improved substantially. However, errors in the input data on building use have also caused errors in the estimation of the population. For exam-

ple, we have noted that in an industrial neighbourhood a large factory was mislabelled as a residential building, so the population has been pointlessly disaggregated in an empty building, inducing a substantial error. The input datasets that we used were very accurate [708, 819], but occasional small errors induced gross errors in the estimations. Furthermore, it is worthwhile to mention that there were peculiar cases that also caused discrepancies, such as a small neighbourhood with a prison as its sole building. Its inmates are counted as residents in the CBS dataset, but the prison building is not classified as residential in the cadastral dataset, hence the estimate of the neighbourhood exhibited a large error—the population was predicted to be 0, while in reality it is 75.

The 3D geometric aspect (calculated volume) may induce errors to the estimations as well. However, related work in analysing error propagation in population estimation is limited to 2D [840–842]. For future work it would be interesting to investigate the influence of errors in the input data when using volume-based approaches.

9.5.3 Analysis of errors and enhancement of the statistical approach

We have analysed the errors with demographic and other indicators for each statistical unit in order to understand them better and to potentially improve our methods.

We have analysed the income of each neighbourhood and did not find a correlation with errors. We had presumed that low-income neighbourhoods might have less space per resident, as income may be related to living standard and may drive residential purchasing choices. However, that is not the case, because there are cheap but large country-side properties, and expensive small flats in cities such as Amsterdam, invalidating our assumption.

In the previous section we have noted the particular behaviour of errors with respect to the different population densities of estimated areas. There is a clear difference between more and less urbanised areas caused by the different utilisation of dwelling space (see Figure 9.8). It is clear that the data on the population density could be used to improve the estimations, but as such it is not available prior to the estimation of population (otherwise we would not need to conduct the estimations).

In Chapter 7 we have found that there is another indicator that it is associated with the population density, and which is available prior to the estimations: the average building height in a neighbourhood is associated with the population density (recall Section 7.3.3.3 and Figure 7.7), and consequently to the living space. Therefore, for each neighbourhood we have calculated the average building height (easily available since we have 3D city models), and we have incorporated it in our regression model (which now contains two variables: the total building space in the statistical unit, and the average height of buildings in the unit). We have not applied this enhancement

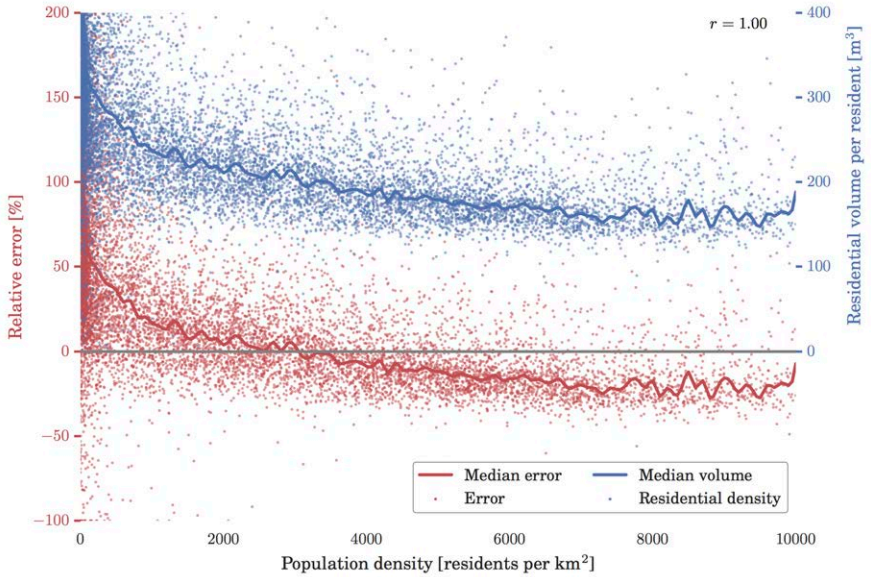


Figure 9.8: The relations between the errors, population density, and living space per statistical neighbourhood. The errors in the model are from the experiment D1/LOD1^c.[30]

to the LOD0 approach in which vertical measurements are not available.

The statistical experiments reveal that there is an improvement to the models: a reduction of errors by a few percent on average has been observed in the models S1, S2, and S3. Note that the results presented in the previous section are of those with the enhanced models, and that the disaggregation method was not enhanced because of its inherently different approach in which there is no training data.

While we believe that the presented prediction models *might* be further augmented to improve the estimations with additional variables and 2D GIS data such as land use, in this work we have used only 3D models to determine how accurate the predictions can be if relied solely on them. Adding such additional variables is avoided because of a contradictory situation: if such data is available, it is likely that accurate census data is also available, rendering such estimations unnecessary.

[30] Derived from data of Kadaster / Centraal Bureau voor de Statistiek, 2015.

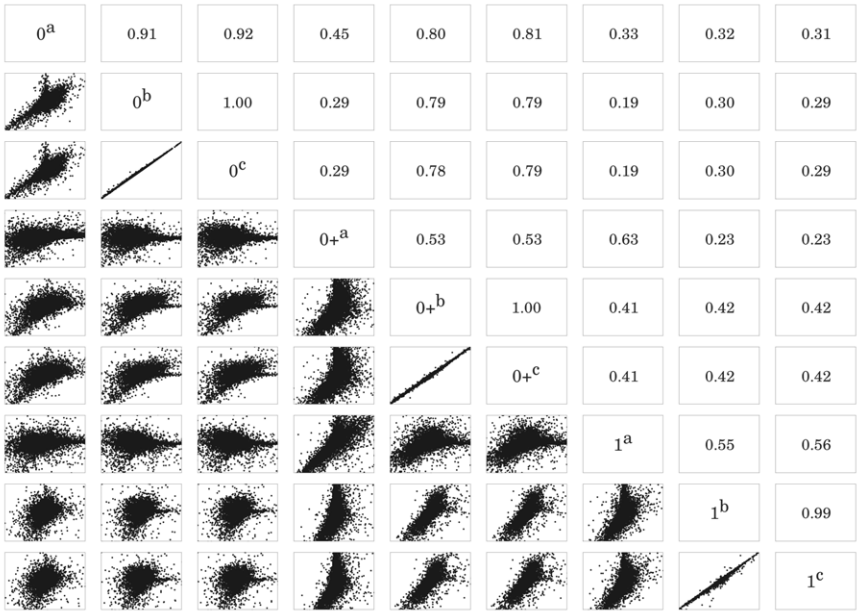


Figure 9.9: Scatter plot matrix revealing the relationship between errors of different spatio-semantic LODs. The numbers are correlation coefficients, and the errors are those of the disaggregation model D1. The plots expose a stronger association between semantic LODs, rather than geometric counterparts.

9.5.4 Analysing errors between LODs

Figure 9.9 illustrates the pairwise relations between errors of each LOD. The plots offer different observations that confirm the ones in the previous section, but also provide a deeper insight in the errors. First, different LODs induce different errors, with the exception of the difference between the semantic levels b and c, which are nearly identical. Second, the relations between datasets of comparable provenance do not reveal a consistent relationship. For example, the correlation between 0^b and $0+^b$ is 0.79, and between 0^b and 1^b is 0.30. While the LOD0 (footprint) and LOD1 (volumetric model) are obtained from the same data source, the errors between them do not exhibit a high correlation, in contrast to the relationship between LOD0 and LOD0+ (floorspace), indicating the stronger role of semantics in this spatial analysis.

9.6 CONCLUSIONS

In this study we have used geoinformation in multiple levels of detail to estimate the population of 12.2 thousand neighbourhoods, 2816 districts, and 393 municipalities in the Netherlands, and of the Netherlands itself. On a side note, considering the aims of this thesis, it is also interesting to note that disaggregation refines the LOD of the population dataset, and that the study has been carried out at different levels of the Dutch administrative hierarchy (e.g. disaggregation from national to municipal level).

Our experiments demonstrate that the LOD of the input data significantly influences the performance of this spatial analysis. The semantic aspect of the LOD appears to have a somewhat larger influence than the geometric aspect. However, several questions remain unanswered due to limitations of the method. For example, the provenance of the used datasets is mixed, as they are derived from different acquisition techniques containing different magnitudes of acquisition errors. In this case it is difficult to set apart the error induced by the LOD from the error introduced by the acquisition error. Another shortcoming is that due to the lack of datasets at finer LOD (e.g. LOD2), it is not possible to test whether using LOD2 would bring an improvement over LOD1. Because LOD0 and LOD0+ are not 3D volumetric models, this study remains limited with regard to the performance of different LODs, as only one volumetric LOD (LOD1.2) was used. These shortcomings will be subject of the next three chapters.

Focusing on the application aspect, this chapter presents a comprehensive nationwide population estimation study carried out at multiple scales. The results indicate that in certain circumstances 3D city models can give a good approximation of the population, and that in most cases 3D city models add value over traditionally used 2D datasets, but also that they are not accurate enough to replace accurate census techniques employed by governments. Furthermore, there were certain instances when 2D data (even without the information on building use, e.g. S1/LOD0[a]) performed better than 3D data, which is beneficial because such data is easier to acquire. The main reason why this method is useful is because it does not require expensive and time consuming field surveys and other means of collecting population counts as the data can be acquired automatically and remotely, and it can be carried out more frequently, in contrast to official censuses (usually conducted every decade).

One of the strengths of our work over previous studies is that we carried out a country-wide analysis, in which differences between neighbourhoods are more emphasised. Beside the multi-LOD aspect of the work (both area-based and volume-based approaches have been evaluated, along with multiple grades of semantic information), our study is multi-scale (for assessing the suitability of mapping statistical units of different sizes), and multi-method (both the weighted disaggregation and

statistical approaches have been employed).

Remote estimation of population with GIS could be applied in areas where census information is not available or it is not reliable, and serves two purposes: (1) as a potential solution to estimate the population count of large areas where a census is not available, or as an intercensal estimate; and (2) for refining the population on a finer scale (e.g. disaggregation of an accurate census of a city among its neighbourhoods).

Our approach is applicable in other countries. Governments have started to publicly release building footprints and other GIS data [666], and where data is available many 3D city models have been generated [67, 416, 507, 843]. Alternatively, 3D city models may be generated from volunteered geoinformation [405], ensuring the applicability of our method elsewhere. While in this study for the building use we used datasets from the cadastre, it is worth noting that such data can also be derived manually from aerial images, and automatically from the building morphology and other characteristics, or from volunteered geoinformation [132, 327, 620, 666, 712, 719, 837]. Such an approach provides an enhancement over previous research, since in related work coarse datasets have traditionally been used, e.g. Kressler and Steinnocher [334] and Silván-Cárdenas et al. [332] distinguish residential buildings from non-residential ones with a zoning map.

Concerning the first application, estimating the population count of large areas where a census is not available, in the 21st century there are still many places around the world where the census has not been carried out in decades, and such remote sensing methods can help to bridge the gap [844, 845]. For instance, Myanmar did not have a reliable census until two years ago, and in the meantime the authorities were dealing with information that turned out to be significantly erroneous [846, 847], something unthinkable in developed countries nowadays. Obviously, low income countries cannot boast about 3D city models, however, with the development of remote sensing technologies, and surge of volunteered geoinformation and their quality [721, 848], the generation of 3D city models is becoming increasingly easier and cheaper [405, 843, 849]. Therefore we expect that in the near future country-wide 3D city models will not be a luxury exclusive to developed countries.

With respect to the second purpose, the derived data on the number of residents on a finer scale is beneficial for a multitude of applications [850], such as disaster management (e.g. in flooded areas) [341, 342], predicting the cascading effects on population in the event of a failure of critical infrastructures [113], analysing accessibility [851], public health [791, 852], crime mapping [853], environmental risk [829, 854], infrastructure planning and transportation sustainability [713], epidemiology [855], territorial classification [856], assessing exposure to noise [249, 259, 857], optimising network coverage (e.g. television) to cover more people [340, 617], for finding areas for landing of stratospheric balloons [617], marketing strategies [329],

estimating the quantity of waste [858], estimating energy consumption [859], and in urban simulations [860].

We have also discovered that this method can also be used to detect potential errors in authoritative census and building data (e.g. we have detected erroneous semantic information for some commercial buildings by analysing the large errors in population estimates). Furthermore, we envisage that such a method could be used for detecting false residencies (e.g. a large number of people registered in a particular neighbourhood for tax-related reasons, triggering an alert by the population that exceeds the housing capacity in that area).

The results indicate that the estimations are hampered by socio-economic disparities between neighbourhoods, and that population estimation is more reliable when focused on statistical units with a closer proximity. However, this limitation does not seem to affect the estimation of the national population, in which case our method has particularly excelled.

For future work it would be worthwhile to advance the sampling method of the training data in the statistical approach to investigate whether that leads to more accurate estimates. For example, stratified sampling [861] could be employed instead of the simple random sampling, which is used now. Such a sampling method could stratify entities based on different characteristics obtainable from 3D city models, such as predominant building types in a neighbourhood, and apply different statistical models to each stratum.

It would also be interesting to investigate the performance of 3D city models derived with the method introduced in Chapter 7. 3D models derived with heights inferred from footprints may give better results than using footprints alone, hence they might be useful in areas where no elevation data is available.

CHAPTER 10

Influence of LOD on spatial analyses (II)

This chapter is based on my paper [769]:

> Biljecki F, Ledoux H, Stoter J (2017): Does a finer level of detail of a 3D city model bring an improvement for estimating shadows? *Advances in 3D Geo-Information*, Springer International Publishing, pp. 31–47. doi: 10.1007/978-3-319-25691-7_2

and partly on my papers [606] and [607]:

> Biljecki F, Ledoux H, Stoter J, Vosselman G (2016): The variants of an LOD of a 3D building model and their influence on spatial analyses. *ISPRS Journal of Photogrammetry and Remote Sensing*, 116: 42–54. doi: 10.1016/j.isprsjprs.2016.03.003

> Biljecki F, Ledoux H, Stoter J (2014): Height references of CityGML LOD1 buildings and their influence on applications. *Proceedings of the 9th 3DGeoInfo Conference 2014*, Dubai, UAE. doi: 10.4233/uuid:09d030b5-67d3-467b-babb-5e5ec10f1b38

The previous chapter presented an analysis to determine the relative contribution of each LOD when data is used in spatial analysis. However, the analysis presented in the chapter is burdened by shortcomings caused by using real-world data. This chapter addresses these limitations by using synthetic data generated with the procedural software prototype developed in Chapter 6, in different use cases. First, multi-LOD dataset are used to estimate shadows. The experiments indicate that the differences between LODs are in most cases subtle and depend on the architecture of buildings. For example, using LOD2.2 instead of LOD1.2 reduces the RMSE of shadow estimation by just 3.1%. Furthermore, the experiments suggests that measuring the benefit of finer LODs is open to more than one interpretation, due to multiple outcomes of a spatial analysis and ambiguity of metrics to measure error. The second part of the chapter takes geometric references into consideration. For this purpose it also introduces two additional spatial analysis to add to the diversity of the experiments, and to investigate whether the influence of LOD and geometric references (i.e. representation) varies from one spatial analysis to another. Several conclusions are drawn. The numerical experiments reveal the interesting fact that, when used for a spatial analysis, models with different geometric references have a substantially higher impact than those with different LODs, and that a coarser LOD with one GR may give better results than a finer LOD modelled with another GR.

10.1 INTRODUCTION

Chapter 9 opens the experimental part of the thesis by presenting a study on the influence of the LOD onto the results of a spatial analysis. The study uses a straightforward three-step approach to determine the influence: (1) source multi-LOD data; (2) run them in a spatial analysis; and (3) evaluate the results. As much as the chapter provides interesting insights to answering one of the research questions (Q8), each of these steps in the method is affected by shortcomings, and these limitations are characteristical to all the related work listed in Section 9.2. Hence, the previous chapter serves as an introduction to this part of the thesis to demonstrate a multitude of errors that plague related work and prevent an objective view on the topic.

The continuation of this section describes the limitations of using real-world multi-LOD data, while the other shortcomings are discussed in the succeeding sections.

Limited number of LODs Multi-LOD datasets are at the moment seldom available, and when they are available usually they come in not more than two LODs. Hence such analyses are limited to only a few LODs, and the previous chapter on population estimation is no different. An alternative to producing data in multiple LODs would be to take a fine LOD and obtain coarser counterparts with generalisation. However, data modelled at a fine LOD are scarce. Moreover, while research in 3D generalisation is plentiful [630], there is no implementation we are aware of.

Acquisition errors Datasets usually contain a series of errors, e.g. positional and topological errors [566]. These errors affect analyses such as this one because a deviation in the spatial analysis may be misinterpreted as an error induced by the LOD. Moreover, the LOD may not be consistent across the dataset and it is not possible to ensure that it has been acquired according to a particular LOD specification. This limitation prevents in analysing whether an error in the output of a spatial operation has been caused by the LOD or by an acquisition error.

Varying lineage When multi-LOD datasets are available, they are usually sourced from a different lineage, usually with different levels of accuracy (Figure 10.1). This different provenance intrinsically inhibits the comparison between LODs, since it is hard to isolate the error induced by the degree of simplification from the error caused by the acquisition.

Lack of fine LODs The absence of data at fine LODs makes experiments less interesting and results in lack of a reference to gauge the results. When data at fine

Figure 10.1: Two datasets of different LODs of the same area but of different lineage. The models (LOD1 and LOD2) are produced in separate campaigns, resulting in different positional accuracies and varying completeness.[31]

LOD is available, it is usually restricted to a small area, insufficiently large enough for experiments.

Geometric references Besides the scarcity of multi-LOD data, datasets with multiple geometric references are non-existing. This is a significant shortcoming, because besides the influence of different LODs we are interested in investigating the influence of different GRs.

All of the above shortcomings can be solved by the use of procedurally generated data as procedural modelling offers a sterile and controllable environment suited for this problem. Hence for these experiments we use the datasets produced by the experimental procedural modelling engine developed in Chapter 6.

In general, synthetic data have already been in use in 2D GIS when experimenting with error and spatial analyses [231, 862, 863], and were proven powerful in testing diverse configurations. An advantage of procedural models is that they can be generated in a straightforward manner and for a large area, and such an approach minimises inconsistencies. Furthermore, the nature of procedural modelling warrants that the models are produced according to a strict specification that introduces

[31] Data courtesy of the Swiss Federal Office of Topography.

Figure 10.2: This chapter in a nutshell. The estimation of the shadow differs between the different LODs, and the accuracy of the prediction appears to increase with each LOD. However, this should be investigated numerically and it is not straightforward as it may appear.

no additional errors.

The aim of this chapter remains the same as the previous one: to conduct a similar analysis determining the influence of the LOD on a spatial analysis, but with alleviating the problems listed above.

A more suitable spatial analysis is selected for the experiments: the estimation of shadows in an urban environment (Figure 10.2). This use case is frequent in 3D GIS, and it is used in several application domains, e.g. to assess the shadow impact of new buildings to their surroundings (Section 10.2).

It is important to note that the study presented in Chapter 9 contains an analysis about both the semantic and geometric LOD of the data. From now on we focus on the geometric part of the LOD, as it is the focus of this thesis, but also because of the lack of other suitable spatial analyses that involve the use of semantics.

10.2 BACKGROUND

10.2.1 The role of shadow in GIS

The estimation of shadows cast by buildings is a common topic in geoinformation science, and it is important for a number of application domains, such as analysing the thermal comfort [230, 244], in solar energy [90, 99, 864], and in developing strategies to mitigate heat [865]. This use case may be related to the visibility analysis, as the estimation of shadow is a variation of the visibility problem, with the particularity that the sun is practically an infinitely distant point resulting in parallel rays, and that its position is variable.

Herbert and Chen [235] underline that understanding shadow is crucial in urban planning, for assessing the effect new buildings induce on existing ones. They perform a survey among urban planners on the quality of the visualisation of the shadow based on different visual representations (e.g. level of transparency, 2D vs

3D), and also include a query about the suitability of the LOD. However, only LOD1 was given as an option, and the participants were given the opportunity to perceptually assess whether a 3D model of LOD1 is sufficient or not for such an analysis.

Estimating shadows is also important for determining solar envelopes, the subset of urban space with a certain period of assured access to sunshine [246]. These are defined in terms of discrete numbers of hours of sun, but they can also be defined in terms of solar irradiation [245]. Solar envelopes are to a degree enshrined in local and state laws, where residents are protected with the right to solar access (e.g. the façade of houses must receive a certain amount of hours of direct sunlight per day; see [236] and [237] for exemplary regulations).

In urban planning, shadows are not analysed only for buildings, e.g. Lange and Hehl-Lange [247] study shadow casting from a proposed wind turbine, and Kumar et al. [80] forecast occlusion by terrain.

Accounting for shadows is common when estimating the solar potential of buildings [82, 96, 99, 240–243]. Strzalka et al. [90] develop a method to determine the shadow projected on a roof surface in order to account for the reduced yield of solar panels when estimating the feasibility of their installation. The method is designed as a visibility problem between small triangular partitions of a surface and the sun at various timestamps. The centroid of each partition is tested for visibility to the sun, and if the sun's ray intersects any of the other surfaces, the partition is marked as shaded at that timestamp.

In a related research, Alam et al. [239] note that while LOD1 block models are sufficient for shading, higher LODs will inherently bring different results. However, this is not supported empirically, and in this chapter we bridge this gap.

Jochem et al. [100] develop a method for estimating the solar irradiation of roofs from point clouds, taking into account shadows. While they deal with a single-LOD representation, they highlight that roof overhangs and chimneys may play an important factor in the magnitude of the shadow. This is important to note because in our research we investigate their supposition.

Shadowing plays an important role in the research of Helbich et al. [109]. Their premise is that solar radiation is significantly capitalised in flat prices, and they consider the shadow effect in order to improve the accuracy of a hedonic house price model. They highlight that such simulations should be conducted for different positions of the sun because of the considerable differences in the results.

Finally, shadows are crucial in geo-visualisation to increase the quality of the visual communication [283].

10.2.2 The role of shadow in computer graphics

Shadows have a longstanding underpinning in computer graphics where they play a significant role, as they enhance the realism of the scene and provide cues of spatial relations such as depth [866, 867]. As a result, many algorithms have been developed to estimate shadows and enhance realism, e.g. recursive ray tracing [868]. Furthermore, many other computer graphics algorithms are closely related to this topic and frequently applied, e.g. the determination of shaded portion of the roof when estimating the insolation may be considered as a ray-triangle intersection problem [869]. We consider such algorithms in the design of our experiment.

10.3 METHODOLOGY

Our methodology for estimating whether finer LODs bring improvements in the estimation of shadows is analogous to the one in the previous chapter:

1. Sourcing multi-LOD data: procedural generation of 3D building models.
2. Shadowing: rendering shadows from the 3D models. We consider the shadow a building casts on the ground. While this use case has been presented in the introduction of this thesis (Figure 1.2), it is relevant to note that in this chapter the shadow is generated in a GIS form suitable for the analysis.
3. Analysis: quantification of shadows, and measuring their error for each LOD. This chapter will expose challenges with quantifying the results of a spatial analysis, and hinting at difficulties of comparing the performance of different LODs when used for the same spatial analysis.

For these experiments 400 diverse buildings are generated in multiple LODs. In order to estimate the influence of the different specifications, we use each model as an input of a spatial analysis. Each representation r_i of a building b is used in a spatial analysis A producing a result $R = A^b_{r_i}$. For each building, also its ground truth $GT = A^b_{gt}$ is computed. Because LOD3.3 is the finest representation available, we consider it as ground truth ($GT = A^b_{\text{LOD3.3}}$). A minor caveat here is that due to the absence of ground truth data we do not have proof that in this case the dataset at the finer LOD brings more accurate results. However, it is reasonable to assume that such comparative differences bring more accurate results (or at least equally accurate) in comparison to analyses using their coarser counterparts.

Afterwards, for each building b, and for each of the used representations r_i, the error in the result of a spatial analysis A is calculated as

$$\epsilon(A)^b_{r_i} = R - GT = A^b_{r_i} - A^b_{gt}. \tag{10.1}$$

To put this residual in perspective, we also calculate the relative error

$$\rho(A)^b_{r_i} = \frac{R - GT}{GT} = \frac{\epsilon(A)^b_{r_i}}{A^b_{gt}}. \tag{10.2}$$

Then for each of the representations the root mean square error (RMSE) values are derived:

$$\text{RMSE}_{\epsilon(A)_r} = \sqrt{\frac{\sum_{b=1}^{n} \left(\epsilon(A)^b_r\right)^2}{n}} \quad \text{RMSE}_{\rho(A)_r} = \sqrt{\frac{\sum_{b=1}^{n} \left(\rho(A)^b_r\right)^2}{n}}, \tag{10.3}$$

where n is the number of buildings.

10.3.1 Considered LODs

For our tests we benchmark the performance of the LODs defined in Chapter 4. Because LOD0 cannot be used for this analysis, models LOD0.0–LOD0.3 are not included in the experiments. Furthermore, because we are interested in shadows cast by the individual buildings, we exclude the aggregated LOD1.0 from the analysis as well, leaving us with 11 LODs.

10.3.2 Sun position and location on Earth

Estimating shadow is considerably influenced by date, time, and location on Earth because of the complex sun paths that differ day by day (Figure 10.3). Figure 10.4 signifies a substantial difference in the behaviour of this spatial analysis when it comes to the different relative position of the sun. In order to diversify our experiments, we consider two locations: Delft (Netherlands) and Kuala Lumpur (Malaysia)[32], and several timestamps during three characteristic days in 2015: spring equinox (20 Mar), summer solstice (21 Jun), winter solstice (22 Dec), and an arbitrary day: 27 April. This variety results in 81 different positions of the sun spread over daytime. Since there are 400 buildings, this means that there are 32 400 samples to be evaluated for each representation.

[32] Kuala Lumpur is chosen because it was the location of the 3D Geoinfo 2015 conference in which the paper behind this chapter was presented [769]. Moreover, it is close to the equator, having a substantially different sun path than Delft.

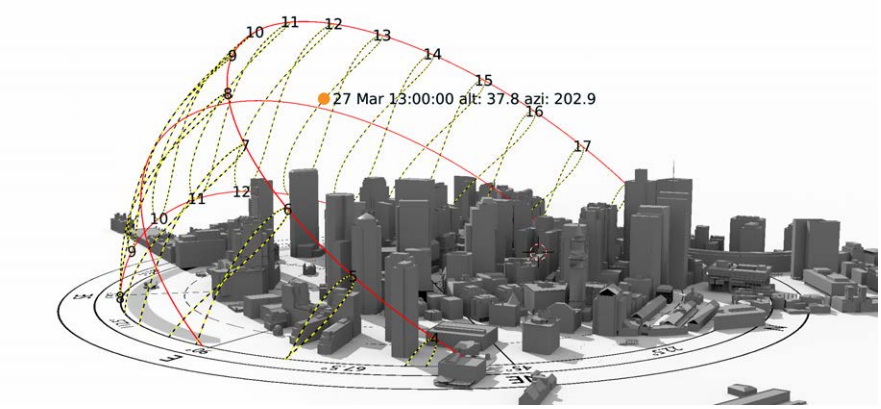


Figure 10.3: Complex sun paths necessitate a multitude of samples. Estimations of the shadow should involve a large number of samples, hindering experiments.[33]

10.3.3 Computation of the shadow

We define the shadow S^{B_i} as the subset of the $\mathbb{R}^2$ (the ground, a horizontal plane in our case, considered as the shadow receiver) that is occluded by a building B_i. When 3D city models are utilised $S_r^{B_i}$ denotes the shadow forecast with a dataset of the representation (LOD) r. We compute the shadow by projecting each polygon of a building model B_i to the plane with a perspective transformation [870], according to the position of the sun. The union of the projected polygons represents the shadow (Figure 10.5), however, in the final step we adjust the polygon by removing the footprint of B_i.

10.3.4 Selection of error metrics

The shadow cast on the ground is a polygon, thus the first measure that comes to mind to quantify a shadow is its area $a(S^{B_i})$, and to compare it to the ground truth:

$$\epsilon_1^r = a(S_r^{B_i}) - a(S^{B_i}). \tag{10.4}$$

As announced in the introduction of the previous chapter, even if a spatial analysis results in a numerical result, the outcome may not be perfectly suitable to be used in the evaluation of error. The shaded area is a good example here, because the deviation of the estimated area may be inconclusive, as it appears in Figure 10.6. Therefore

[33] Data used for the illustration is courtesy of the Boston Planning and Development Agency. The analysis has been carried out with Vi-Suite and Blender.

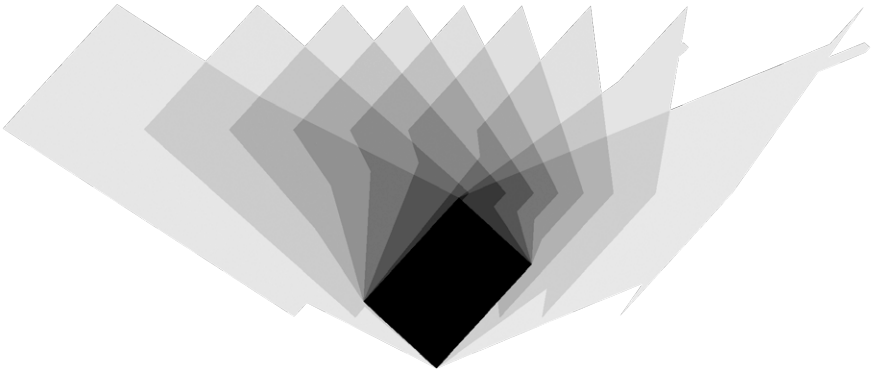


Figure 10.4: Orthogonal top view composite of 9 shadows during a day for an LOD3.3 model. This image is also known as the butterfly shadow diagram (the outline of each shadow is drawn at hourly intervals). It suggests the different degree of influence of LOD-related details on the shadow depending on the date, time, and location. The building footprint is shown in black. Not to be confused with Figure 1.2 in the introduction, which illustrates the annual amount of shadow.

we introduce a new metric: the area of the symmetric difference (the union without the intersection, see the light red area in the same figure):

$$\epsilon_2^r = a(S_r^{B_i} \ominus S^{B_i}) = a((S_r^{B_i} \cup S^{B_i}) \setminus (S_r^{B_i} \cap S^{B_i})). \tag{10.5}$$

In the GIS context this non-negative metric appertains to commission and omission errors, and to false positives and false negatives: the union of the (1) subset that is estimated as shaded but in reality it is not with the (2) subset of the inverse case.

Because area is only one of the aspects that quantifies the extent of a shape, we compute the similarity between the two shapes, which is a fundamental subject in computer science and GIS [751, 871, 872]. There is a variety of methods and metrics to express the correspondence of two shapes in GIS [873, 874], one of the prominent being the Hausdorff distance [448]. It has been widely used in geoinformation science and 3D city modelling for diverse purposes [875], such as to assess the quality of GIS data [876, 877], to assess the performance of 3D generalisation [878], to aid map matching [879], to analyse movement trajectories [880], and to detect changes between two CityGML models [370].

The Hausdorff distance quantifies the mismatch between two geometries by identifying the point on one shape that is the maximum distance from the other shape, therefore we define the third shadow error metric as

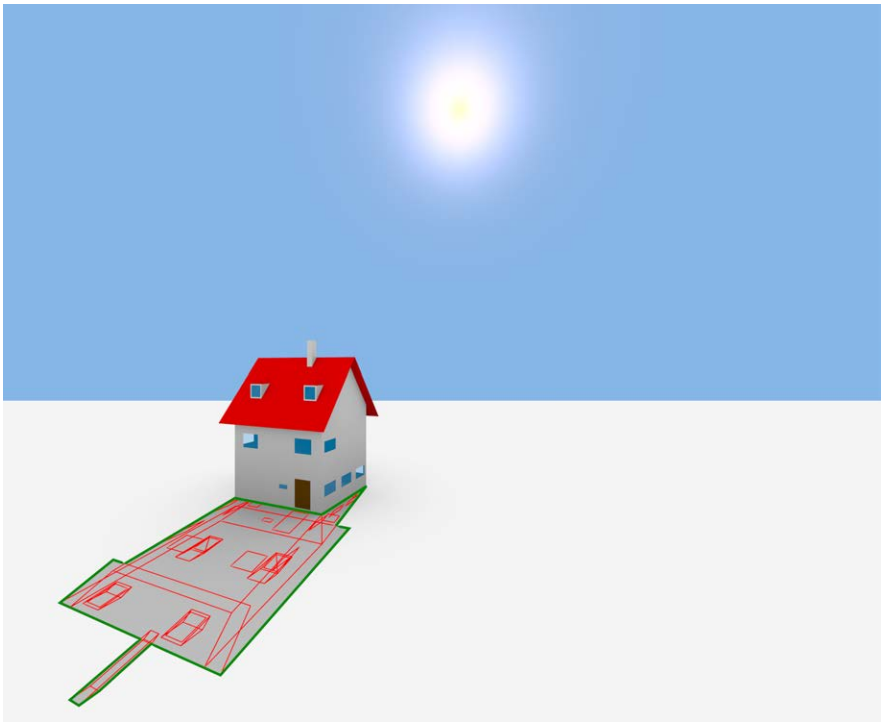


Figure 10.5: The approach to derive shadows. In our approach and our software implementation, the shadow on the ground is derived as a unionised set (green) of projected polygons (in red; 51 polygons in this LOD3 case) from the CityGML model, and accounting for the footprint.

$$\epsilon_3^r = H(S_r^{B_i}, S^{B_i}) = \max(h(S_r^{B_i}, S^{B_i}), h(S^{B_i}, S_r^{B_i})), \tag{10.6}$$

where $h(A, B)$ is a function that finds the point $a \in A$ that is farthest from any point in B and measures the distance from a to its nearest neighbour in B:

$$h(A, B) = \max_{a \in A} \min_{b \in B} ||a - b||. \tag{10.7}$$

For the three error metrics we compute their RMSE. While the Hausdorff distance technically is not an error, it is not uncommon to compute its root-mean-square value [881, 882].

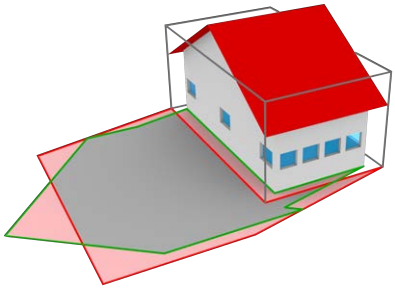


Figure 10.6: Ambiguity of the metrics to quantify the error. A case of two shadows from an LOD1 and an LOD3 (outlined in red and green, respectively) that have the same area, indicating that this error metric can be ambiguous. The area of their symmetric difference (light red area) is 28.7 m^2 (33.9% in relative terms), and the Hausdorff distance in this case is 2.47 m.

For ϵ_1 and ϵ_2 we compute also the relative error (with respect to the true size of the shadow) to put the derived numbers in perspective, which is not possible for ϵ_3.

10.3.5 Implementation

Available implementations do not fully support our methodology, necessitating the development of an own solution. For example, several GIS tools contain a module to forecast shadows at a specific timestamp, however, such a functionality cannot be exploited to our advantage: the tools cannot be automated nor the shadow can be exported as a vector geometry (see Figure 1.2 in which the shadows are rendered as a raster). Therefore, we have implemented in Python a software prototype that reads CityGML data, estimates their shadow for a particular location and timestamp, and exports it as a polygon.

The sun positions are taken from PyEphem/XEphem[34], the implementation of the ephemeris of Bretagnon and Francou [883] and Meeus [884]. The shadow polygons operations (e.g. union and symmetric difference) are achieved with Shapely[35]. The Hausdorff distance has been calculated with PostGIS[36]. For validating the correctness of shadows, we have first converted a CityGML model and its calculated polygon shadow to the OBJ format (with CityGML2OBJs, a software that I have developed during this PhD research; see [885] for the methodology), and imported

[34] https://pypi.python.org/pypi/pyephem/ and http://www.clearskyinstitute.com/xephem/
[35] https://pypi.python.org/pypi/Shapely
[36] http://postgis.net

Table 10.1: Numerical results of the experiments. The error metrics are expressed as RMS values, both relative and absolute measurements.

LOD	ϵ_1		ϵ_2		ϵ_3
	[m^2]	[%]	[m^2]	[%]	[m]
1.1	27.6	16.2	40.3	30.1	2.5
1.2	27.6	16.2	40.3	30.1	2.4
1.3	27.2	16.0	39.9	29.9	2.4
2.0	25.1	13.1	33.3	20.7	1.8
2.1	25.1	13.1	33.3	20.7	1.6
2.2	25.1	13.1	33.2	20.6	1.6
2.3	0.5	0.7	0.5	0.7	0.4

it in Blender[37], an open-source 3D computer graphics software. We have rendered the setting for the same date, time, and location, thanks to the Blender add-on Sun Position[38]. The shadows matched—this is evident in Figure 10.5 where the rendered shaded area and the shadow polygon are conflated from independent workflows.

Some of the shadow polygons were found to include long tiny spikes due to floating point errors, which was inhibited with snap rounding [886], and triangulation-based polygon repair with the tool prepair [887].

In addition to calculating the error metrics, the computational cost was recorded (computation time and number of projected polygons), in order to suggest the load of each LOD. The earlier Section 3.3 argues that quantifying the cost in GIS is not as straightforward as in computer graphics, however, in this particular analysis the cost was added into the results to spark a discussion.

10.4 EXPERIMENTS

We have conducted experiments on 400 different buildings in 11 LODs ($400 \times 11 \times 81 = 356$k shadows in total). We present the errors in Figure 10.7 and Table 10.1, and discuss them in Section 10.4.1. The results of LOD3.x models will be given later.

[37] https://www.blender.org
[38] http://wiki.blender.org/index.php/Extensions:2.6/Py/Scripts/3D_interaction/Sun_Position

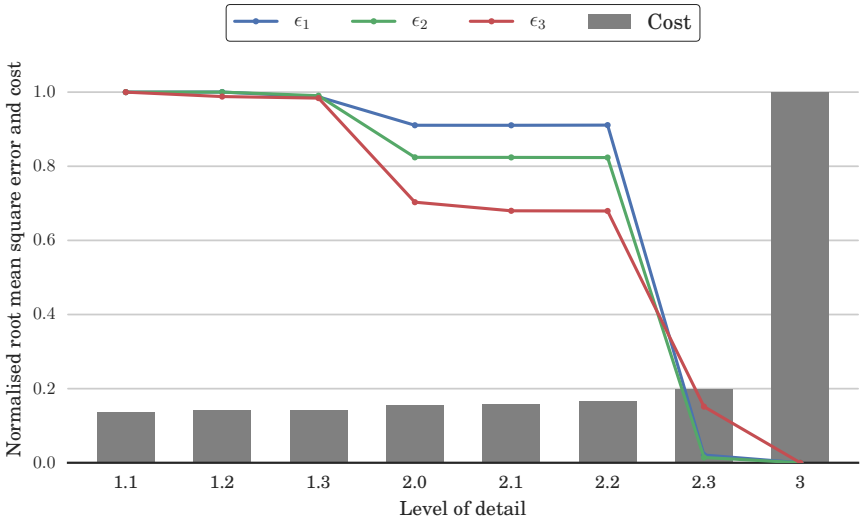


Figure 10.7: Errors and computation cost for each LOD. The metrics are normalised according to the least favourable result.

10.4.1 Findings

The main findings of the experiments, as outlined in Figure 10.7 and Table 10.1, suggest that the relative errors between most LODs are small, and the improvements of each LOD are not significant. Furthermore, we point out other findings:

- The improvement of LOD2 over LOD1 is almost negligible if considering the shadow as a whole (only a 3% reduction in the area error).
- Overall, a finer LOD does bring a more accurate result. However, that is not always the case for each building. The improvement depends on the configuration of the analysed area. As an example, Figure 10.9 illustrates the distribution of ϵ_2, ϵ_3 errors for LOD 2.2. It reveals that for many buildings the error is negligible (e.g. in that LOD for 19% buildings there is no error ϵ_2; for LOD1.1 that value is at 10%). A more detailed inspection revealed that this applies to buildings with flat roofs and no roof superstructures. If such buildings dominate in an area to be analysed, the acquisition of finer LODs is probably not beneficial.
- Modelling dormers (LOD2.2 and 2.3) and other roof details has a negligible

influence on the quality of the prediction. This can be explained by the fact that dormers and chimneys are not present in all buildings, and they make a difference only during a limited time of the day (see the example in Figure 10.4).

- Different data (types of buildings) and different settings (day, time, location) result in a different behaviour and magnitude of the error, indicating that related experiments should be diverse. Figure 10.8 shows the variation of the magnitude of errors as a function of time during one day.
- Figure 10.8 also indicates that while in absolute terms the ϵ_1 and ϵ_2 errors increase with the actual size of the shadow (close to sunrise and sunset), their relative counterparts decrease. Furthermore, the behaviour of ϵ_3 is different.
- LOD3 contains openings, which have no influence on the shadows (unless in the special case of the sun rays passing through two windows, but this triviality was not taken into account). The improvement over LOD2 is mostly caused by overhangs and other smaller details, which on the other hand are probably not appreciated by the use cases that require shadow estimation as an input. Furthermore, an LOD3 model entails a substantial cost of acquisition and processing, which also has to be taken into account.

10.4.2 Evaluation of error metrics

The results of a seemingly simple analysis of determining the accuracy of a shadow estimation suggest that errors can be approached from different perspectives as there are different ways to quantify not only the outcome but also the arising errors. Figure 10.9 illustrates the relationship between our second and third error metrics, and the distribution for each. We have computed the correlation between the errors to investigate their relation:

$$r_{|\epsilon_1|,\epsilon_2} = 0.967 \qquad r_{|\epsilon_1|,\epsilon_3} = 0.099 \qquad r_{\epsilon_2,\epsilon_3} = 0.085.$$

An interesting outcome is that there is a low degree of correlation between the area error metrics and the Hausdorff distance. As visible in the scatter plot, there are several cases in which the magnitude of the area error metrics (ϵ_1 and ϵ_2) are small (<0.1 m^2), with a significant value of ϵ_3. This is most evident in the results of LOD2.3 where the first two error metrics are low, and the Hausdorff distance is not negligible. Manual inspection revealed that these deviations are caused by small LOD3-only details such as chimneys, which render small shadow area differences, but since they protrude, ϵ_3 is of noticeable value (e.g. see the chimney in Figure 10.5).

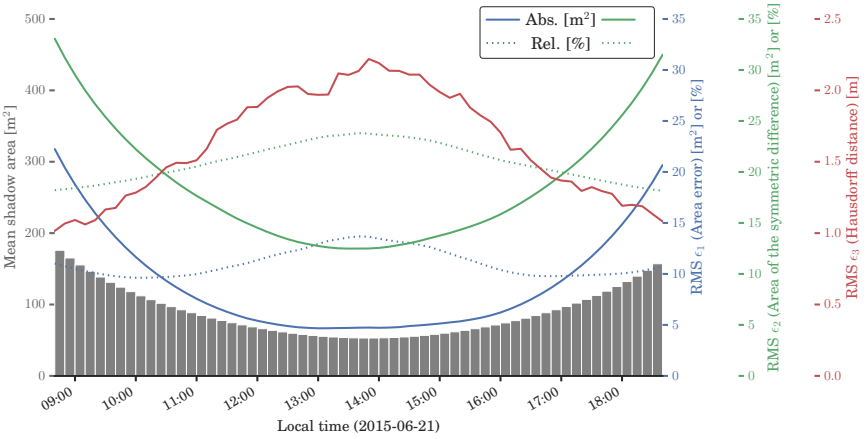


Figure 10.8: Behaviour of the three error metrics. The metrics exhibit differ behaviour with respect to the position of the sun, and thus—the size of the shadow in the ground. The values refer to the experiments involving LOD2.2.

We are ambivalent on the use of the Hausdorff distance for this purpose. Besides the advantage that ϵ_3 revealed some discrepancies that were not detected by the first two, it helped to put the area errors in perspective, i.e. some large area errors were in fact caused by practically insignificant deviations (e.g. long and narrow strips of shadows). The disadvantage is that the Hausdorff distance is not a stable error metric (see Figure 10.8), and it is sensitive to computational and geometric errors, e.g. caused by floating point errors and slivers.

For the first two error metrics we have computed also a relative counterpart, as the relation to the total size of the shadow. This helped to understand the true magnitude of the error (e.g. that the RMS of ϵ_2 for LOD1.2 is 16% of the shadow size).

10.5 DISCUSSION ON THE INFLUENCE OF THE LOD

So far we have presented a study on the influence of the LOD of a 3D building model on the quality on a 3D GIS spatial analysis. We benchmark the relative influence of several LODs for estimating the shadow, and conclude that investing in a dataset at a fine LOD is not always a good idea. For example, for 50% of buildings modelled at the coarse LOD1.1, the error (ϵ_1) is within 10.2% of the size of the shadow (percentile rank of a score), which practically does not have significant influence for some use cases. For areas with a higher share of buildings with a flat roof this fact would be

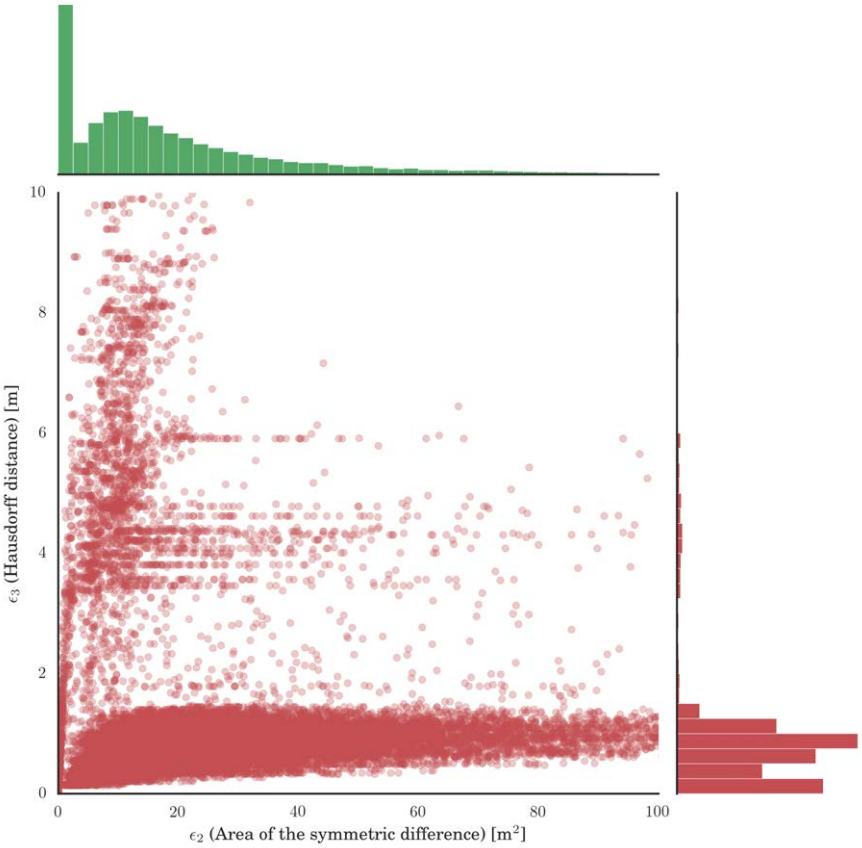


Figure 10.9: Scatter plot showing the relationship between the error metrics ϵ_2, ϵ_3 and their distribution (for LOD2.2). The histogram on the right shows the distribution of ϵ_3, while the one on top shows the distribution of ϵ_2. The latter shows that for a substantial share of samples the error ϵ_2 is insignificant.

even more substantiated.

Therefore, we challenge the universally accepted assumption that finer LODs inherently bring more accurate results in spatial analyses, and we argue that such analyses are important to understand the impact of LOD on a specific use case.

Shadows do not have a unique metric, a fact that is also valid for many other spatial analyses. We use three error metrics: area error (ϵ_1), area of the symmetric

difference (ϵ_2), and the Hausdorff distance (ϵ_3), which exhibit different observations. In our study, we determine the influence of the resolution of the models on raw shadows as standalone concepts. While we find that the LOD has a variable influence, these smaller improvements may not always benefit a use case. Actually, it may damage it: while the improvement is negligible, the acquisition and processing costs could be substantially higher. This depends on the *weight* a shadow has as an input in a use case. For example, in geo-visualisation a more accurate shadow probably does not make any difference, while in some other such as predicting the yield of photovoltaic panels it might be tangible.

10.6 INVOLVING GEOMETRIC REFERENCES AND OTHER USE CASES

Hitherto the experiments have involved estimating the results of one spatial analysis, and the concept of geometric references (discussed in Chapter 5) has not been taken into account. It is *a priori* obvious that utilising models with different geometric references may result in substantial deviations in spatial analyses sensitive to the geometry, such as computing the volume of buildings and estimating their shadows. On the other hand, in some spatial analyses, such as the estimation of flooding, different geometric variants might not have a significant impact. Hence, besides investigating numerically the impact of LODs, it is important to research the differences between models of different geometric references when employed in a spatial analysis, and to obtain insights how GRs affect the outcome of a particular spatial analysis.

Two additional use cases are included, and data in multiple geometric references, as defined in Chapter 5, have been generated.

Analysis 2: Gross volume of buildings Estimating the volume of a building has gained substantial attention in 3D GIS [657]. Nowadays it is essential in use cases such as energy demand estimation [78, 114, 120, 888], determination of property taxes [129], estimation of the population in a given area (Chapter 9), urban planning [373], material flow modelling and quantification of development densities [713], in the volumetric visibility analysis of urban environments [889], estimating the stock of materials in the building sector [738], and waste management [733].

Similarly as shadows, volume is also subject to multiple interpretations. For example, it can refer to the net internal volume, to the volume of the bounding box, the volume of the shell, and to the gross volume. In these experiments the gross volume is taken into account (Figure 6.7). We have computed the volume of building solids (in m^3) with the Feature Manipulation Engine (FME[39]), automated by an iterating

[39] https://www.safe.com/fme

Python script.

Analysis 3: Area of the building envelope 3D city models are suitable to calculate the area of the exposed building shell. The area of the building envelope is calculated as the sum of areas that comprise the shell[40] of a building. We have implemented a software prototype that calculates the area from CityGML models. The information of the building envelope provides valuable input in several applications, and 3D city models are frequently used for this purpose. For example, in assessing the cost of energy-efficient retrofitting of a building [116, 124], estimating the loss of energy in households [78, 83, 118], estimating indoor thermal comfort [105], predicting cooling requirements [119], thermal simulations involving computational fluid dynamics [322, 325], analysing the urban heat island effect [890], and in urban design evaluation [192].

10.6.1 Experiments and discussion

In the experiments with shadows, 400 buildings have been generated. This quantity appears to warrant a diversity in the buildings. For each building dozens of different sun positions have been used, resulting in dozens of different results (shadows) for the same building. The two new spatial analyses are different—they end in a single result. In order to retain a comparable number of samples, for these two spatial analyses we generate an extended dataset of 10000 buildings. The buildings, as it is the case with the dataset for shadows, have been generated in 78 representations resulting in 780k CityGML models (21 GRs for each of the LOD1.1–3; 3 GRs for each of LOD2.0–2; 2 GRs for LOD2.3; and a single representation for each of LOD3.0, LOD3.1, LOD3.2 and LOD3.3). A visual excerpt of the generated models is illustrated earlier in Figure 6.5.

The full results of the three experiments are provided numerically in Table 10.2 and graphically in Figure 10.10. In the following sections we compare the distribution of errors for a better understanding of the deviations, and discuss the results. In order to directly compare the differences between spatial analyses, in the table we present the relative errors.

Results of experiment 1 (shadow) The results of the first experiment (using the error metric ϵ_1) are given in Figure 10.11 (only representative LODs are given due to redundancy). They also suggest that different geometric references have a substantial influence on this spatial analysis. Buildings with the footprint modelled at its

[40] Here it is important to note that the measure of the area of the building envelope is not clear in literature. For example, it is not clear whether the roof overhangs are included in the measurements. In our experiments we include roof overhangs.

Table 10.2: The relative RMSEs, expressed in percents, of the three analyses by the LOD and geometric references. The coloured markers indicate the magnitude of the error.

	Shadow			Volume			Envelope		
Representation	F0	Fd	F1	F0	Fd	F1	F0	Fd	F1
LOD1.1 H0	21	19	17	18	25	33	18	15	14
H1	19	16	16	14	25	34	15	12	13
H2	15	13	14	6	25	36	10	9	12
H3	15	14	16	4	28	41	7	9	14
H4	16	16	20	7	32	46	7	11	17
H5	22	25	30	19	44	59	12	19	25
H6	23	26	31	20	45	60	13	19	26
LOD1.2 H0	21	19	17	18	25	33	18	15	14
H1	19	16	16	13	25	34	14	12	13
H2	15	13	14	6	25	36	10	9	12
H3	15	14	16	4	28	41	7	9	14
H4	16	16	20	8	32	46	7	11	17
H5	22	25	30	19	44	60	12	19	26
H6	23	26	31	20	45	60	13	20	26
LOD1.3 H0	22	19	17	18	25	32	18	15	14
H1	19	16	16	14	25	34	15	12	13
H2	15	13	14	6	25	36	10	9	12
H3	15	13	16	3	28	40	7	8	14
H4	16	16	20	7	32	46	7	11	17
H5	22	25	30	19	44	59	12	18	25
H6	23	25	31	20	45	60	12	19	26
LOD2.0	14	12	13	0	26	38	7	8	12
LOD2.1	14	12	13	0	26	38	7	8	12
LOD2.2	14	10	13	0	26	38	7	9	12
LOD2.3	1	10		0	26		0	9	
LOD3.0			13			38			12
LOD3.1	0			0			0		
LOD3.2	0			0			0		

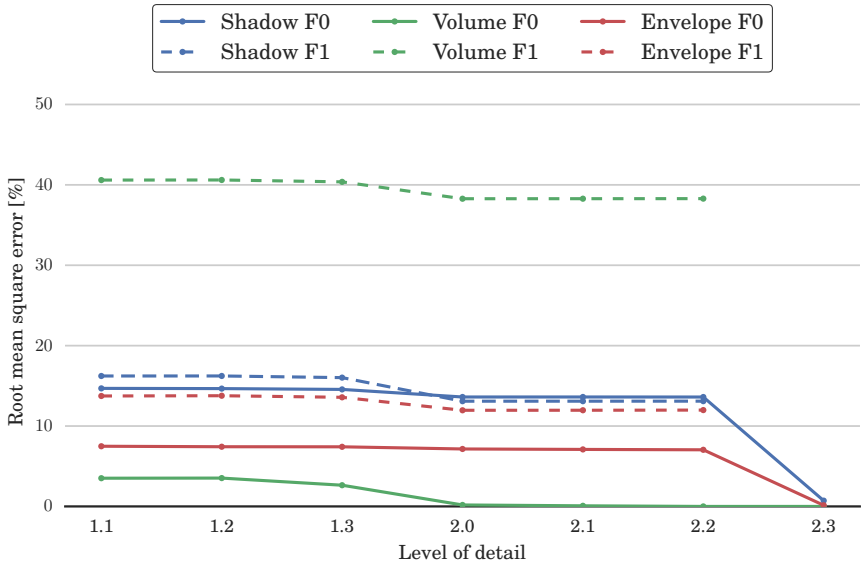


Figure 10.10: Geometric references have a higher impact than the LOD. Comparison of F0 and F1 versus increasing the LOD.

actual location (F0) generally provide a more accurate analysis. The range of errors in LOD1 is from 13% to 31%, while in LOD2 the errors range from 1% to 14%.

Results of experiment 2 (volume) The distribution of errors in the second experiment is given in Figure 10.12. We notice that the results are pronouncedly different from the first experiment, affirming that it is important to run these experiments for different spatial analyses. In LOD1, the errors range from 9% to 56%, and in LOD2 from 10% to 38%. Again, there is a substantial number of samples for which the error is small (peaks at zero), due to a number of building configurations that do not differ significantly among different geometric references.

The experiments suggest that the optimal representation appear to include models with a horizontal reference F0, due to the more truthful representation of the building body. Furthermore, the experiments indicate that LOD1 models may be fairly accurate in the computation of the volume if their top surface is modelled at the half of the roof structure.

In this analysis, however, LOD2.3 models do not appear to provide an advantage over coarser LOD2.(0–2) models, since the roof overhangs are not included in the

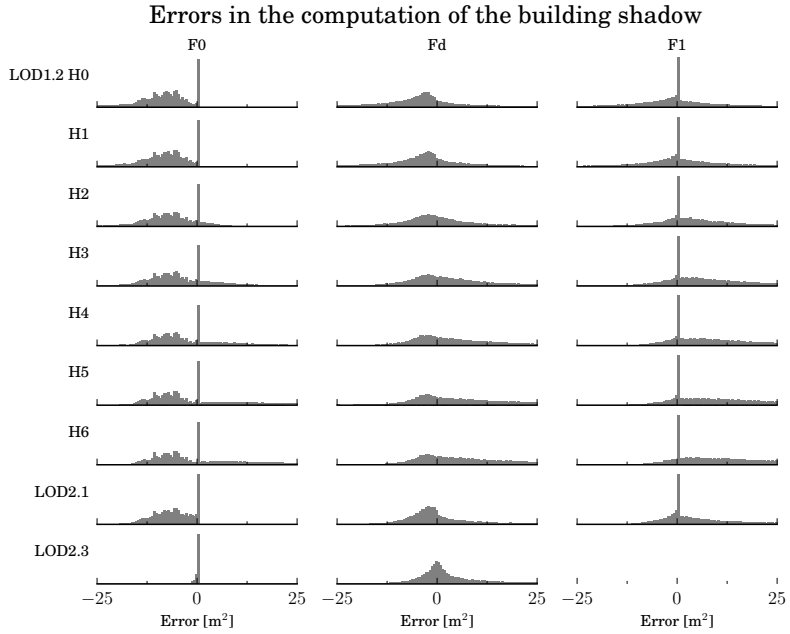


Figure 10.11: Results of the first experiment involving the estimation of the area of the shadow cast on the ground.

computation of the volume, and therefore their presence provides no advantage here.

Results of experiment 3 (envelope) Figure 10.13 reveals the distribution of error of estimating the area of the building envelope. We observe from the results that there is a substantial difference between geometric references, and that the errors between references within the same LOD are not shifted—each distribution is unique, and it does not correspond to any known probability distribution. Some of the errors are gross (e.g. an error of 26% in case of LOD1-H6-F1), rendering models of this reference unusable for this purpose.

The peaks at 0 are from buildings with flat roofs without roof overhangs, where the geometric references do not have a significant influence. Within LOD1 the RMSE ranges from 7% to 26%, while in LOD2 from < 1% to 12%.

Our experiments reveal that for LOD1 the most suitable geometric reference is the combination of F0 and H4, and for LOD2 a model with the footprint at its ac-

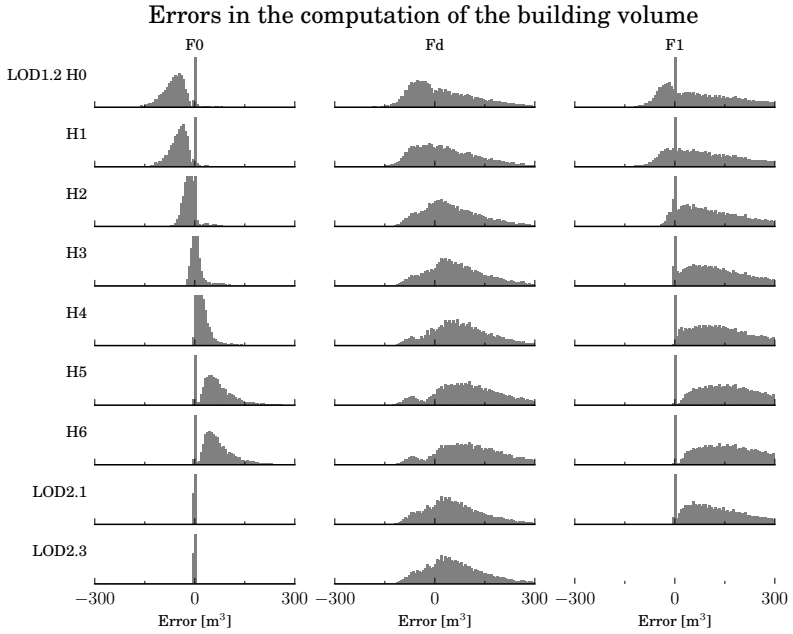


Figure 10.12: Results of the second experiment involving the computation of the building volume.

tual location (F0). The Fd reference appears to be somewhat advantageous over F1. A paradoxical observation is that some combinations of references indicate that a coarser LOD can be more accurate than a finer LOD (e.g. LOD1-H3-F0 has a smaller RMSE than LOD2.1-F1).

LOD2.3 models with the footprint at F0 appear to provide an advantage over coarser LOD2 models with the same reference, due to the more factual representation of the roof.

10.7 CONCLUSIONS

This chapter presents a comprehensive study on the influence of level of detail and geometric references on the accuracy of spatial analyses. In addition, the chapter reveals obstacles in analyses such as this one, and the implementation that was introduced solves such problems and overcomes the shortcomings of related work. A

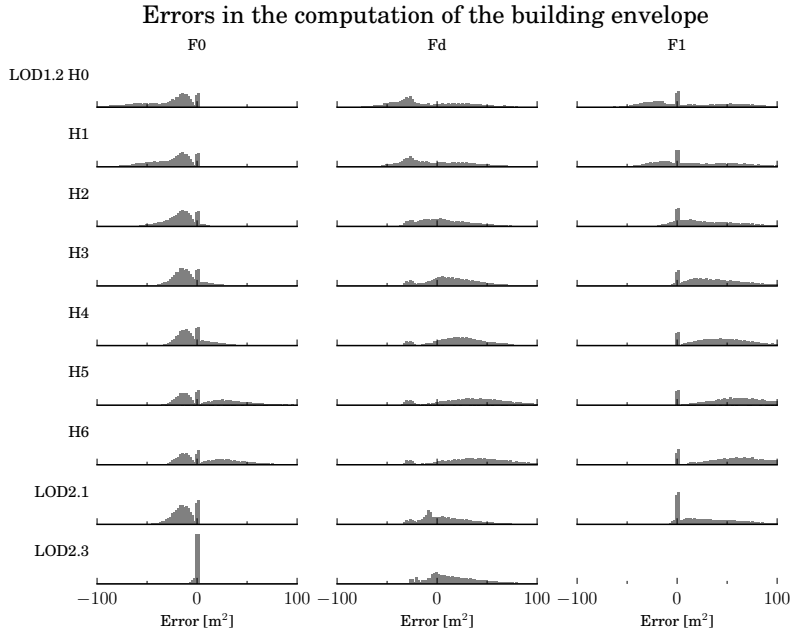


Figure 10.13: Results of the third experiment involving the computation of the area of the building envelope.

strength of this work is that three spatial analyses have been considered, and procedurally generated data that is consistent and free of errors has been used as input. For example, procedural modelling ensures that the data is generated strictly according to the LOD specification that is analysed.

The chapter provides valuable insight in the research of determining the impact of different representations when used for a spatial analysis. First, it has demonstrated that as it is the case with measuring detail, measuring benefit and errors is difficult. There are multiple ways to express the outcome of a spatial analysis and to assess its error. Comparing the results between multiple spatial analyses may also be a challenge due to different scales and units of measurements. It is not always easy to normalise these outcomes to use percentage errors. For example, population estimation and rendering shadows derives non-negative results (number of people and cast area), so percentage errors make sense. However, estimating the results that can be negative or do not have a comparable scale of measurement (e.g. water level, and angle of view) prevents a comparison.

Second, the impact of the error of deriving this 'crude' result of a spatial analysis eventually depends on the intended use. The error of 10% when deriving the volume of a building does not have the same impact when the data is used for population estimation or for taxation. In population estimation the results are aggregated per neighbourhood, hence they may be diminished, which is not the case for taxation (recall the analysis in Chapter 7.6).

Third, implementing experiments such as this one require custom solutions, as existing 3D GIS software does not enable feeding a spatial analysis in repeated iterations, and obtaining results in an automated way. For that reason, the spatial analyses have been implemented mostly from scratch (including the population estimation analysis presented in Chapter 9). An example is the spatial analysis of shadow estimation. Earlier in this thesis (Figure 1.2) a shadow study carried out with an open-source software has been exhibited. However, the software outputs shadows only visually as rasters and makes it difficult to automate inputs and outputs, a must for experiments such as ours.

The experiments indicate that variable behaviour of the LOD deriving mixed conclusions. On one hand, some spatial analyses require a minimum LOD, for example, the volume estimation is not possible with LOD0. On the other hand, while using volumetric data provides a large benefit over 2D data, there is no significant difference between LOD1 and LOD2. The final results vary among buildings—ultimately they depend on the architecture and morphology of a building. In certain circumstances (e.g. basic block buildings) a fine LOD does not bring any difference. In others (e.g. the complex buildings in Figure 7.12), modelling them in a fine LOD is crucial. That is, LOD1 may be sufficient for one building but poor for another. Relating this conclusion to the computation of building volume in Chapter 9, it is unlikely that using LOD2 instead of LOD1 would bring a significant benefit for population estimation.

Furthermore, while the LODs refined in Chapter 4 are different from each other (not only visually but also in terms of acquisition, amount of geometry, and memory), these differences are not significant when the LODs are used for a spatial analysis. This means that the cost invested in obtaining a 3D city model may not result into proportional benefits.

In a very different manner, our experiments reveal that models of different geometric references have a significantly larger effect than the LOD, so the GR is more important to consider than the LOD. For example, in the second spatial analysis (volume), using a coarse model LOD1.2 with the reference H3/F0 results in an error of 4%, while using a more detail model LOD2.2 but with the reference F1 results in a large error of 38%.

This finding may also imply that there is no such a thing as a general purpose 3D city model, since each use case prefers a model with a particular geometric reference.

Therefore, when acquiring a 3D city model, the choice of the geometric reference should be driven by the intended use of the models. Furthermore, when dealing with multiple applications, one should accept uncertainty and the fact that the model will likely not be equally suitable for all applications. These findings suggest that it is important to carry out such experiments for each spatial analysis to understand the different behaviour of the specifications, but several can be built upon these, so could be reused to quickly consider a new use case.

The study also suggests that different representations induce different influence in different spatial analyses. For example, the range of errors is different between spatial analyses (cf. maximum error of 60% and 26%, in the second and third experiment, respectively), hence different spatial analyses exhibit different sensitivity by using different GRs.

The work therefore proves the importance of considering the geometric reference when acquiring and utilising 3D city models. The relative differences between the results of spatial operations utilising models of the same LOD but of different geometric references may be gross.

Recent developments in using 3D data in energy modelling and in similar domains [891, 892] suggest the need for the integration of multiple spatial analyses. These spatial analyses may not benefit from a particular specification, as each may prefer its particular LOD/GR combination. Producing data in different LOD/GR combinations and further research on linking multiple representations would come handy here, because it would enable switching between representations that are most optimal for a particular spatial analysis, and enable the use of data of different LOD/GR combinations within the same application.

The most important conclusions of this chapter are:

1. We have found that because each spatial analysis has different requirements there is no optimal LOD or GR. Our approach can be used to determine the most suitable GR for a specific spatial analysis.

2. An interesting observation is that an LOD1 with a specific GR may yield more accurate results than an LOD2 for some spatial analyses.

3. The results of the three spatial analyses indicate that the effect of the GRs strongly depends on the configuration of the building. For example, models of buildings with flat roofs and no roof overhangs are invariant across multiple LODs and GRs. This is in contrast with buildings with a more complex configuration, such as the ones that contain non-flat roofs, and protrusions such as balconies, garages, and alcoves. We would expect that in other geographic areas, e.g. those with larger buildings and of different shape, the errors would be of different magnitudes. For future work it would be interesting to

investigate the relationship between the input distribution and the systematic error of the spatial analyses.

4. To some extent, the Fd reference (offset walls) appears to be advantageous over F1, however, it does more harm than good for buildings without overhangs or with overhangs that are shorter than the distance d. This is also visible in the histograms in the absence of peaks at 0.

5. The distribution of errors is not shifted between GRs: it is unique for each GR. This is caused by differences in the configurations of buildings.

Finally, our experiments have indicated that LOD2.3—the enhanced version of the LOD2—which contains explicitly modelled roof overhangs, may bring an improvement in accuracy and performance over the 'usual' simple LOD2 instances LOD2.0–2. Such models are not complex to acquire if terrestrial and airborne observations are available. Hence, we encourage practitioners to consider this model in their production workflows as it appears to be an optimal choice balancing the cost of procurement and the quality of the results.

For future work it would be beneficial to investigate the significance and influence of the LOD on the outcome of an application that uses estimated shadows as input. For example, the error in the prediction of the duration a wall is shaded during the day, or the error in the estimation of the loss of the solar potential in kWh/year due to shadows (but also payback time of solar panels [893]). Shadow can also translate in money (property value lost caused by shading), indicating that there may be several different ways to quantify the results of a spatial analysis.

CHAPTER 11

Sensitivity of LOD to positional errors

This chapter is based on my papers [894] and [895]:

> Biljecki F, Heuvelink GBM, Ledoux H, Stoter J (2015): Propagation of positional error in 3D GIS: estimation of the solar irradiation of building roofs. *International Journal of Geographical Information Science*, 29(12): 2269–2294. doi: 10.1080/13658816.2015.1073292

> Biljecki F, Ledoux H, Stoter J (2014): Error propagation in the computation of volumes in 3D city models with the Monte Carlo method. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, II-2: 31–39. doi: 10.5194/isprsannals-II-2-31-2014

The results presented in Chapter 10 are obtained with synthetic data without errors to investigate the relative differences between LODs. Reality is different: when data is acquired, it is burdened by errors intrinsic to acquisition techniques. In this chapter we investigate how positional errors influence the data modelled in different LODs. While error propagation in GIS is a topic that has received a lot of attention, it has not been researched with 3D GIS data. Hence we extend error propagation to 3D city models. A new use case is used for the purpose of this chapter: estimation of solar irradiation of roofs, supported by a software prototype that I have developed during this research. The experiments rely on Monte Carlo method in which procedurally generated 3D city models are repeatedly disturbed with simulated positional errors in parametric form and with varying XY/Z levels to reflect actual acquisition outcomes, a novelty in the subject of error propagation. The experiments indicate that that LOD1 is less affected by positional errors than LOD2 and LOD3. Furthermore, the chapter draws attention to the finding that procedurally generated data appears to be quite suited for error propagation experiments, and that the planar and vertical uncertainty have a different influence on the estimations.

11.1 INTRODUCTION

The acquisition and utilisation of geographic information go together with errors. Error propagation in GIS refers both to the unavoidable erroneous nature of processing and utilising GIS datasets, and to a field of research that investigates the propagation of errors in the input to a GIS operation to the errors in its output [407, 896].

Previous chapters have not taken into account errors induced by the acquisition of data (realisation of the specification). The goal of this chapter is twofold: (1) to investigate how positional errors affect the result of a 3D spatial analysis; and (2) to determine whether the behaviour is similar among different LODs.

Two common methods to determine the propagation of errors through GIS operations are the Taylor series method and the Monte Carlo method [897]. The former is analytical and establishes mathematical functions that portray the uncertainty propagation, while the latter is a numerical simulation method. Many GIS analyses are too complex to be solved analytically, hence the Monte Carlo method is frequently used [898].

While error propagation is a mature and thoroughly studied subject in geoinformation science, research has essentially been limited to 2D situations. Furthermore, work is focused towards raster data and attribute uncertainty, and research on vector data and positional uncertainty are less common. Extending the error propagation problem from 2D to 3D is more complex than one may expect, which motivated us to research this topic in 3D.

The foundations of error propagation are valid for any dimension, however, in the context of 3D city models there are several factors to take into account:

(1) *Level of detail.* The models may be derived in multiple forms distinguished by LOD. An arising research question is whether different LODs have an influence on the propagation of errors (e.g. if a 3D model at coarse LOD is less affected by errors in comparison of a dataset at a finer LOD).

(2) *Combination of multiple acquisition techniques and accuracy levels.* In contrast to 2D data, a significant number of 3D city models is produced with a combination of multiple acquisition techniques, such as photogrammetry, lidar, and extrusion from 2D data (see Section 2.2 for an overview). Since a model may be constructed using multiple techniques with each having different accuracy levels, this results in different positional accuracy for the x, y coordinates on the one hand, and the z coordinate on the other hand. The augmentation method presented in Chapter 7 is a good example of such difference, because the footprints have been derived from a very accurate source, while the heights have been estimated from the number of floors and other attributes.

(3) *Geometry and measurements*. The construction of 3D city models generally involves many more measurements than the 2D acquisition of the same real-world feature (e.g. height of the eaves of a roof, and position of a window on a wall).

(4) *Applications*. 3D city models are nowadays used for applications not possible with 2D data, such as shadow estimation (see Section 2.3 for an overview, and Chapter 10 for a specific case). Furthermore, they are used in applications that are considerably more complex than when used with 2D data, such as estimating noise pollution [259] and estimating of the energy demand of buildings [114].

The objective of this chapter is to generalise the error propagation analysis to 3D GIS operations, and to investigate how the Monte Carlo method can be used to obtain insight into error propagation in spatial analyses in 3D city modelling in multiple LODs. In this analysis the two spatial analyses introduced in the previous chapter are continued to be used: estimating the volume and envelope area of buildings. The shadow analysis is not used, because due to extensive computational load the experiments cannot be performed (explained later). Instead we introduce a new spatial analysis: the estimation of the solar irradiation on roofs of buildings, for the aim of assessing the suitability of the installation of photovoltaic (PV) modules for producing electric energy. This application is a widely used example of the use of 3D city models and it requires 3D GIS operations. For this study, we conducted an analysis for the city of Delft with a software prototype implemented during the PhD research, but the software and method are sufficiently generic to be applicable to other locations in the world. This use case is more interesting than calculating the volume and envelope area of buildings, because it involves multiple spatial operations and it has a more complex methodology. Hence we focus mostly on this use case, but we come back to the other spatial analyses later. Furthermore, we explain the methodology of this spatial analysis in detail. By giving a detailed description we aim for a better understanding of the process. Here we also bridge the literature gap because, despite its popularity, this 3D application has not been thoroughly described methodologically.

In this research, we focus on the uncertainties of the geometry of the model, and investigate how positional accuracy influences the quality of a 3D spatial analysis, and whether this behaviour is consistent across different LODs. Several different types of errors can be introduced during the acquisition, which propagate through a spatial analysis in different ways, depending on the context. For example, the standard ISO 19157 on geographic data quality defines several types of errors, e.g. completeness, topological, positional, thematic (attribute), and temporal errors [628].

Positional accuracy is a prominent spatial data quality element, and a principal descriptor in the metadata of a geo-dataset. In most analyses the error in the position affects the outcome of a GIS operation, hence we focus on it.

11.2 RELATED WORK AND RESEARCH OPPORTUNITIES

The subject of error propagation is researched in several disciplines, such as physics [899], medical imaging [900] and chemistry [901], and it has a solid mathematical and experimental foundation that can be applied to other disciplines. It is related to GIS: geographical observations describe phenomena with spatial, temporal, and thematic components that are all acquired with uncertainty [902]. Besides measurement errors, uncertainty may be caused by processing, generalisation, and several other factors [903]. Hence, understanding the propagation of errors in GIS is important. For example, it is important to set expectations when obtaining and utilising datasets for a specific purpose. One may define a maximum acceptable uncertainty of the result of an operation, and by performing the analysis of the propagation of uncertainty within the operation may reversely determine the maximum allowed uncertainty in the input data. An example is the usage of 3D city models in estimating the visibility between two points in space. This technique is used by radio engineers and telecom companies to estimate the radio signal coverage [312], and for security purposes as in determining the optimal surveillance camera placement [200, 201]. While the former may be successfully accomplished with a rough and not overly accurate 3D city model, the latter relies on street-scale data where a comparatively small error may result in significant errors in the output.

The topic of error propagation has been researched and documented in numerous publications that go back several decades. Heuvelink et al. [904] performed a quantitative analysis of gridded 2D data in a raster geographical information systems with two use cases, Arbia et al. [905] modelled the error propagation of overlay operations in raster GIS, and van Oort et al. [906] researched the propagation of positional uncertainty of vertices in polygons to the computation of its area. Shi et al. [907] researched the propagation of the error of the interpolation in DEMs, de Bruin et al. [908] investigated the uncertainty of the input geographic data to planning costs in agriculture, and Shi et al. [909] describes the propagation of buffer-related errors and completeness. Furthermore, Heuvelink and Burrough [910] investigated error propagation in cartographic modelling using Boolean logic and continuous classification. The documented analyses include the propagation of spatial errors and attribute errors, with the latter being more frequently represented. For example, Veregin [911] investigated the propagation of thematic errors.

In this chapter we focus on positional error, which has been the subject of several

error propagation analyses. For example, Goulden et al. [912] investigate the propagation of positional error in point clouds to the calculation of various topographic attributes, such as slope, aspect, and watershed area. Griffith et al. [913] and Zandbergen et al. [914] observe the influence of positional error in geocoded addresses on various administrative use cases such as the assignment of houses to census blocks, and allocation of students to their nearest public schools. Positional errors are omnipresent in GIS and they have been extensively discussed in the literature, hence they do not require a lengthier introduction.

To the best of our knowledge, there is no work published in the international scientific literature on uncertainty propagation in 3D spatial analysis. Shi [915] presented a generic approach for modelling positional uncertainty in multi-dimensional objects, but did not address the propagation of uncertainty.

Propagated errors are defined as the discrepancies between performing identical operations on the true and on the observed, error-contaminated data layers. Error propagation modelling is the formal process of representing the transformations in data quality that occur through GIS operations on data layers [905]. The error propagation problem can be formulated mathematically as follows [896]. Let $U(\cdot)$ be the output of a GIS operation $g(\cdot)$ on m inputs $A_i(\cdot)$:

$$U(\cdot) = g(A_1(\cdot), \ldots, A_m(\cdot)). \tag{11.1}$$

The operation $g(\cdot)$ may represent virtually any GIS operation, such as stream network calculation from a DEM or electromagnetic field modelling from mobile phone base stations [314, 916]. The objective of the error propagation analysis is to determine the error in the output $U(\cdot)$, given the operation $g(\cdot)$ and the errors in the inputs $A(\cdot)$. The error or uncertainty is usually expressed with the variance of $U(\cdot)$.

The Taylor series method, one of the two methods to assess the error propagation in GIS, requires the mathematical derivative of the operation $g(\cdot)$. However, this method is not suited for complex operations since the derivative may not exist or be difficult to compute analytically. In contrast, the Monte Carlo method is made to work by running the assessment model repeatedly with random disturbances introduced into the uncertain inputs, where the degree of disturbance is in accordance with its uncertainty. When used in error propagation, it identifies the relationship between the input error distribution and that of the model outputs. Thus, it allows to determine whether the input error is amplified or is suppressed [917]. Following the previously introduced notation, the reasoning of the Monte Carlo method is to compute the result of $g(A_1, \ldots, A_m)$ repeatedly, with input values $A_i, i = 1, \ldots, m$ that are randomly sampled from their joint distribution [896]. The spread in the sample of m model outputs obtained in this way captures the uncertainty in the output.

Examples of the Monte Carlo method employed in the analysis of error propagation in GIS include an application in flood management [918], a GIS-based assessment of seismic risk [917], potential slope failures [919], natural resource analysis [920], interpolation of DEMs [921], analysis of highway maintenance [922], and evaluation of the accuracy of agricultural land valuation using land use and soil information [923].

11.3 USE CASE: ANNUAL SOLAR IRRADIATION WITH SOLAR3DCITY

The estimation of the solar irradiation at a location is one of the most prominent use cases in 3D GIS (see the recent literature review of Freitas et al. [104] or the overview in Section 2.3). The aim of this use case is to quantify *how sunny* a location is. This is done primarily for terrain [80, 924] and for building roofs, mainly to assess the suitability and economic return of the installation of solar panels [90, 92]. The latter application is gaining significant interest in the 3D city modelling research community and industry.

It is useful to introduce some terminology. Solar irradiance describes the *instantaneous* rate of energy that is being delivered to a surface (power per unit area), usually expressed in W/m^2 (cf. Figure 2.4c). It varies depending on the location of the surface, time and date, atmospheric conditions, and other factors. The solar irradiation is the total amount of solar energy that has been collected on a surface area within a given time, i.e. solar irradiance *integrated* over time. It is also known as insolation, and it is typically expressed in kWh/m^2/year [925].

In case of determining the suitability of a photovoltaic installation on a surface, such as a building roof, the normalised solar irradiation value may be coupled with the area of the surface resulting in the solar irradiation of a surface in kWh/year. Contemporary solar panels are able to utilise only a fraction of this energy (about 15-20% according to a 2016 paper [926]), but the exact amount also depends on additional factors, such as the ambient temperature. When estimating the amount of energy that can potentially be captured, we refer to the solar or photovoltaic potential [88]. Because solar potential depends on specific technical panel settings, and because the technology is continuously improving, in this research we focus on the perennial magnitude of the solar irradiation.

The use of GIS data and 3D city models in this use case is crucial. First, solar irradiation differs at different locations on Earth, i.e. due to different day lengths and the position of the sun (see Figure 10.3 in the previous chapter). Second, solar irradiance may significantly differ between building roofs in the same area, depending on the orientation of the roof: a more favourable angle of a surface to the sun means a better exposure and more solar energy. Two roof surfaces of the same size

at the same location but of different orientations and inclinations (azimuth and tilt), may drastically differ in their solar potential [85, 927]. This will also become evident from our experiments. Third, the larger the surface of a roof, the more solar energy is available, hence the information about the surface area of the roof is important as well.

11.3.1 General overview of the 3D use case

The solutions significantly vary from small to large-scale applications, each yielding a different spatial resolution of results. For example, from a neighbourhood level with a value for each building [928], to a national analysis for an average value for each municipality [929] and global analyses as rasters of pixels of large size [930–932]. Perpendicularly, the complexity of the solutions vary from relatively straightforward (clear-sky models) to complex solutions accounting for meteorological conditions, vegetation, and shadowing [97, 98, 239, 241, 242].

This study deals with large-scale analysis (i.e. on buildings), which may be performed on different sets of data, e.g. voxels [96, 933], lidar point clouds [99–102], and the ones with building data, e.g. derived with a combination of lidar and GIS data [86]. While most of the work dealing with the solar potential of buildings is focused on rooftops, some extend the solar potential to vertical façades [94].

In addition, supplementary information about buildings, such as the building type and number of inhabitants, may be used in order to relate the potential gain to the actual energy usage of the household [117, 934]. That is a reason why semantic 3D city models, such as CityGML data [10, 408], are being increasingly used for this use case [90, 116, 591]. Performing the analysis with semantic 3D city models has further advantages over the aforementioned forms, e.g. the roof surfaces can be easily identified for the analysis.

11.3.2 Methodology for the estimation of the solar irradiation with 3D city models

Solar radiation can be decomposed into three components [84, 95], and these are analysed separately in the estimations. The radiation that is not reflected or scattered in the atmosphere and reaches the surface directly is known as *direct radiation*. The radiation scattered from atmospheric particles and clouds is called *diffuse radiation*. The part of the radiation that is reflected from the ground onto an inclined (roof) surface is denoted as *reflected radiation*. The three components of radiation together form the *global radiation*. Accounting for all three components is important, because in certain settings each may significantly differ [935].

The estimation of the direct and reflected irradiation and their adjustment for the tilted and oriented surface are straightforward [936]. The estimation of the diffuse

solar irradiance is more complex, but there are several empirical models that can do this, and their comparison has been the subject of several research papers [937–939].

Hence, the solar irradiation I is estimated by integrating the solar irradiance E over a period of time:

$$I = \sum_i E_i \cdot \Delta t_i, \tag{11.2}$$

where Δt_i is the duration of the i^{th} time interval.

Since the solar irradiance depends on the geographic location, from a GIS point of view, the solar irradiation I of a surface at a location $[\varphi, \lambda, h]$, with tilt τ, azimuth α, and area A may be seen as a function:

$$I = f([\varphi, \lambda, h], \tau, \alpha, A, [w]), \tag{11.3}$$

where φ, λ, and h indicate the latitude, longitude, and elevation of the surface, and $[w]$ denotes a series of other components such as meteorological conditions.

Since the model f is complex, the propagation of errors can best be solved numerically with the Monte Carlo method, involving a large number of building models and disturbances of simulated input errors. Considering the large number of simulations, in our work we do not take into account shading effects since these entail computationally expensive algorithms.

11.3.3 Design and implementation of Solar3Dcity

Because there is a lack of open-source software package for the use case, and also because as in Chapter 10 we need a highly customised solution, we designed and implemented our own software prototype Solar3Dcity. The software supports data stored according to the OGC standard CityGML, and can handle large datasets, which is a requirement for a Monte Carlo simulation. Another motivation for building our own software is that it enables creating a custom format for the exchange of the results in the way that is suitable for our workflow.

Figure 11.1 illustrates the general workflow of the software, which consists of deriving four geographic parameters for each building; these are essential for the solar potential analysis: area of the roof, tilt, azimuth, and the geographic location of the surface. The latter value is important in order to compute the solar ephemeris for a location on Earth and to retrieve the historical irradiance data, but note that it will not vary between buildings on a local basis, as the values do not change within the spatial extent of a typical city. Hence it is important to emphasise that this value is not affected by positional errors, unless they are caused by gross errors, such as wrong coordinate reference system, which is beyond the scope of this work.

Because a roof can consist of more than one planar surface, the annual solar irradiation is calculated separately for each surface, and their values are summed up to

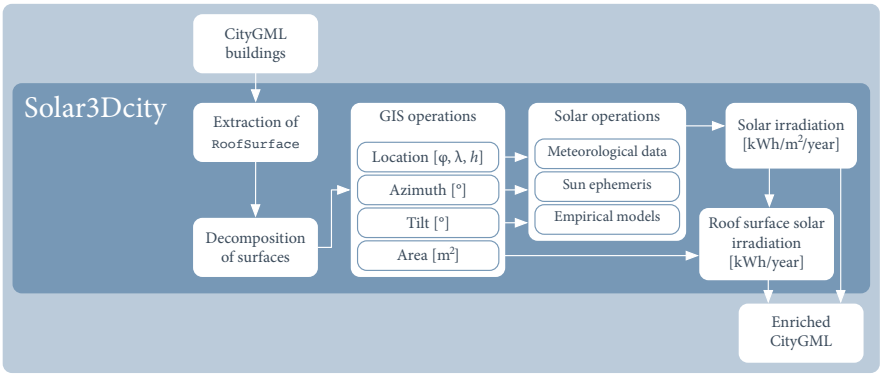


Figure 11.1: Software architecture of Solar3Dcity, the experimental prototype developed in this research. The implementation reads semantic 3D city models in the CityGML format, extracts the surfaces that represent the roof, and enriches them with the value of the annual solar irradiation and with the total value for all roof surfaces of a building.

obtain a single value of the total solar irradiation in kWh/year for a building under the assumption that all roof surfaces are suitable for the installation of the photovoltaic panels. However, it is not economically feasible to install solar panels in all cases [83, 93], e.g. on a small roof area that does not get a good deal of sunlight, hence we define the filtered total building solar irradiation, which is the total irradiation excluding roof surfaces that are lower than 2 m^2 and have a solar irradiation smaller than 850 kWh/m^2/year. Note that introducing this filter is also interesting in the analysis of the propagation of errors, because the determination of whether a surface passes this filter is prone to input errors, and it can potentially result in Type I (false positives) and Type II (false negatives) errors (i.e. rejecting a surface that is in reality suitable, and vice-versa).

Before the computations, each surface is geometrically validated (e.g. non-planar surfaces and self-intersecting boundaries), ruling out invalid cases occurring in real-world data and making this implementation robust. As in the software developed in the previous chapter, the solar ephemeris are from PyEphem/XEphem, which have been proven suitable for solar radiation studies [940]. The computations use the empirical anisotropic model developed by Perez et al. [941], which was implemented in the Solpy library [942]. The historical solar irradiance data, which are important to reflect the actual climatic conditions and to approximate a typical meteorological year [943], are taken from the nearest meteorological station (in our case from the IWEC dataset—International Weather For Energy Calculations [944]). The annual global solar irradiation is then calculated for each roof by integrating the hourly ir-

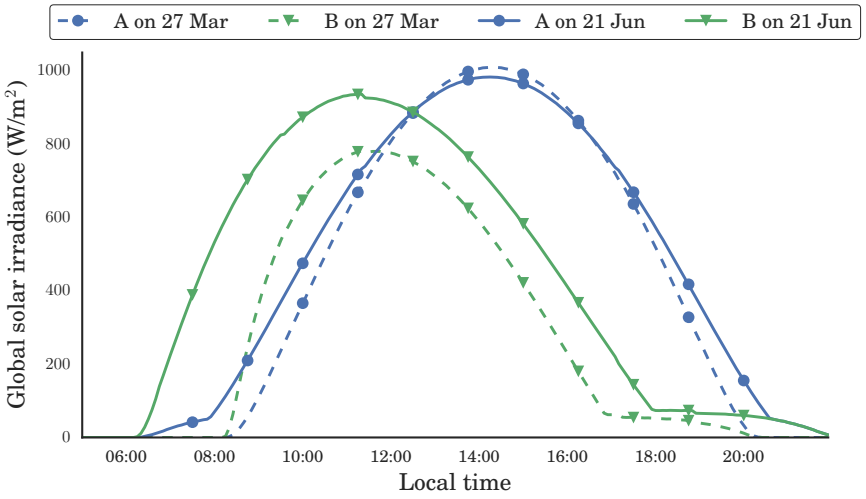


Figure 11.2: Clear-sky global solar irradiance for different scenarios. Surface A is facing south and it is tilted by 40 degrees, while surface B is inclined by the same angle but it is oriented towards the east. In order to obtain the annual solar irradiation of each combination (tilt, azimuth) of a surface, the irradiance needs to be summed up for each hour for each day. Moreover, it is important to note that these data are location-dependent.

radiance values of the entire year (similar to the approach of Mardaljevic and Rylatt [240]). Figure 11.2 illustrates why it is necessary to do so—it reveals the solar irradiance for two different surfaces at two different days. It is evident that there are lots unique combinations that have to be calculated for each day of the year.

To remove redundancy and to speed up computations, we precomputed the normalised function $ir = f([\varphi, \lambda, h], \tau, \alpha)$ for the location of the study area (i.e. the city of Delft in this study). Once the orientation of a surface is estimated from the 3D city model, the software samples the irradiation value from the function, rather than doing the computations all over again. Figure 11.3 illustrates the values of such a function, i.e. the solar irradiation in kWh/m^2/year for every azimuth/tilt combination. This function, also known as *tilt-orientation-factors* (TOF), is also a product itself since it helps finding the optimal tilt and orientation for a location, which is the main aim of several location-based studies [930, 945, 946]. The last component, i.e. the area of the surface, is thence multiplied with this normalised irradiation value to obtain the irradiation of the roof surface.

After the irradiation is computed, the value of the estimated annual solar irradi-

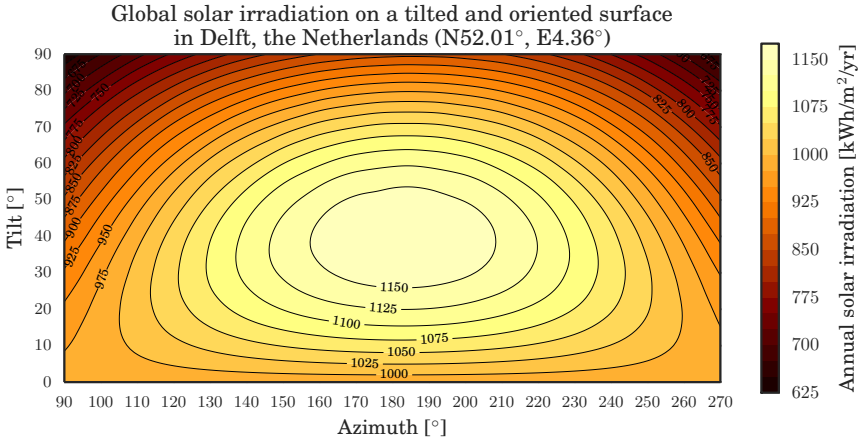


Figure 11.3: The annual global solar irradiation for Delft as a function of tilt and azimuth, as estimated by our software Solar3Dcity. The centre of the plot is facing south, and it is bounded by east and west facing surfaces, meaning that this is only half of the complete function. Values for surfaces facing north are further decreasing. The plot indicates significant variation of the solar irradiation between differently tilted and oriented surfaces, suggesting the importance of 3D city models and GIS operations in this application.

ation per building and its roof surfaces are stored in the CityGML file. An example of the visualisation of such enriched CityGML file is illustrated in Figure 11.4.

11.4 METHODOLOGY AND IMPLEMENTATION

11.4.1 Overview

The uncertainty propagation method involves four main steps (Figure 11.5):

1. Generate a set of buildings in a parametric and spatio-semantic form (3D city models) with procedural modelling. We again use procedural models mainly because of the advantages discussed in the Chapter 10, but also because it is easier to disturb these with simulated acquisition errors. Synthetic data have been used for similar purposes, e.g. for analysing bathymetric models [947], uncertainty in land cover maps [863], and assessing classification performance [948].

2. Perturb the simulated *base* dataset generated in the first step repeatedly with random disturbances sampled from the probability distribution representing

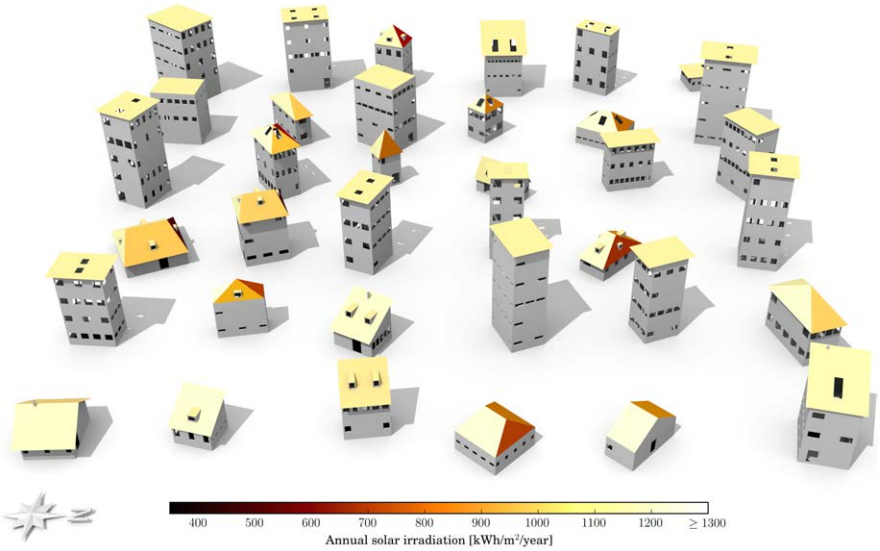


Figure 11.4: Annual solar irradiation of building roofs. The roofs in this example CityGML dataset are coloured based on how much solar energy they receive in one year per square metre (normalised annual insolation). The orientation of the north is given in the bottom left part of the figure. To put these values in perspective for the experiments: a rooftop of a building with an optimally inclined shed roof of an area of 100 m^2 and without roof windows receives around 117 MWh of solar energy per year. Not to be confused with Figure 2.4c, which illustrates the solar irradiance (amount of energy at a specific time and date).

the geometrical errors. For this we implemented an error engine that builds on top of the procedural modelling engine by *injecting* errors in the process.

3. Analyse both the ground truth and erroneous models with Solar3Dcity to estimate the solar irradiation of roofs.

4. Compute the errors in the solar estimations, and calculate common uncertainty measures such as the Mean Error (ME) and Root Mean Square Error (RMSE).

11.4.2 Considered levels of detail

Because different geometric LODs contain different geometric content, it is interesting to investigate how acquisition errors propagate depending on the LOD. The

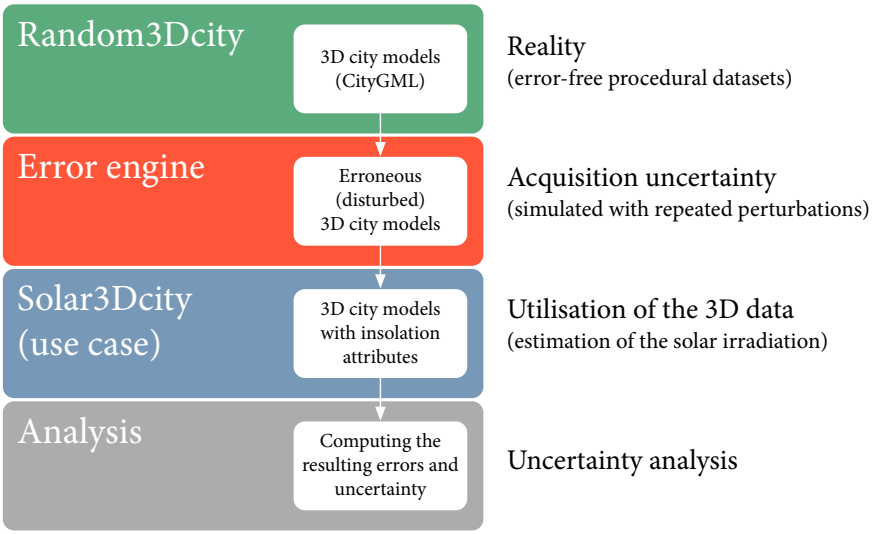


Figure 11.5: Workflow of the uncertainty propagation analysis. Each part is a separate software module developed in this PhD research, and data flows from one to the other. The first and third modules are detailed in Figures 6.1 and 11.1, respectively.

experiments have been carried out with all LODs and geometric references. However, to keep things clear only a few LODs are described in the results. While the different LODs have been discussed at multiple points in this dissertation, it is important to examine them again in the context of this use case.

LOD1 is a coarse model in which regardless of the roof shape, the top surfaces have the same orientation and tilt (they are flat). Hence their value in this use case is questionable as they contains a gross systematic error. While they provide a good approximation of a building with a flat roof, in reality, when solar panels are installed on flat roofs they are not laid flat [949], as they are optimally oriented according to TOF models [243] (as the one in Figure 11.3). This is a different practice from that used in case of non-flat roofs, where the solar panels are laid on the roof, essentially retaining the same orientation as the underlying roof surface. Nevertheless, in our analysis we include LOD1 models, to investigate whether the uncertainty propagation differs between different LODs.

LOD2 is a model with simple roof shapes, and it is most commonly used in solar potential analyses worldwide. Conceptually, LOD2 models provide the exact tilt and azimuth of the roof, and they are relatively straightforward to acquire (e.g. see Haala and Kada [578] and Musialski et al. [502]). However, as discussed in Chap-

ter 4, LOD2 models may not always include dormers and roof windows, hence they tend to have a systematically larger roof area than in reality. For this reason in the experiments we include two variants of LOD2: LOD2.1, which has 'clean' roof structures, and LOD2.2, which includes dormers and similar roof superstructures. Here it is important to note that having an LOD2.3 in this use case does not provide added value because the roof surface is equal to the one of LOD2.2.

As a detailed architectural model, LOD3 contains roof windows and other smaller roof features such as chimneys. LOD3 models have not been much used for solar potential analyses (I am aware only of a few instances, e.g. the recent papers of Peronato et al. [560] and Li and Liu [950] using LOD3 in an micro-analysis including one house). For LOD3 models, Monte Carlo simulations have not been performed with values of σ greater than 0.3 m, because acquisition techniques are usually more accurate in the case of LOD3 than in the case of LODs 1 and 2.

11.4.3 Procedural models and their perturbation

As in the previous chapter, the experiments use procedurally generated data. Procedural models have not been used before in uncertainty propagation. In the continuation of this chapter it will become evident that they are in fact quite useful for error propagation.

As explained in Chapter 6, our engine first generates a real-world feature (e.g. a building) F_i by deriving a set of n parameters p_i that unambiguously define it:

$$F_i = \{p_i^1, p_i^2, \dots p_i^n\}. \tag{11.4}$$

One example of the parameters is the height from the lowest point to the eaves of the building F_i: $p_i^h = 6.54$ m. However, in reality, we do not know the true parameter values because these are subject to observation error. Hence, we represent the true parameter value by a random variable that is modelled as the sum of the parameter estimate and a zero-mean, normally distributed stochastic error ϵ, resulting in the uncertain parameter $\widetilde{p^j} = p^j + \epsilon^j$.

Therefore, between the first (parametric) and the second (3D city model realisation) step (see again the software architecture in Figure 6.1) we insert a stochastic engine that *degrades* the parameters by sampling values from a normal probability distribution function with standard deviation σ:

$$\widetilde{p_i^j} = \mathcal{N}(p_i^j, \sigma^2). \tag{11.5}$$

Simulating acquisition errors with an uncertainty engine is in line with error propagation analysis methods in 2D GIS (e.g. see Brown and Heuvelink [951], Heuvelink et al. [952], Ben-Haim et al. [953], Xue et al. [954]). In the engine we use the normal

distribution because it is commonly used in geomatics [955–957], and because it has a theoretical underpinning through the Central Limit Theorem from statistics.

The above results in an erroneous version of the building $\widetilde{F}_i = \{\widetilde{p}_i^1, \widetilde{p}_i^2, \dots \widetilde{p}_i^n\}$, e.g. for the aforementioned height example we might get $\widetilde{p}_i^h = 6.71$ m when using $\sigma = 0.2$ m. The engine can generate an infinity of erroneous versions of the building by repeating the perturbation method, each time sampling a new from the probability distribution of the parameter error.

Disturbing the parameters rather than the coordinates of the geometry of the generated 3D city model has several advantages: it is more straightforward, it diminishes the chances of creating degenerate geometries (e.g. non-planar surfaces and invalid solids), and it conforms to the real-world practices of acquisition of 3D city models. Most software packages nowadays facilitate modelling of buildings by measuring the distances of buildings' lengths at right angles [958], especially in automatic workflows [33]. Furthermore, as the noise is added to the parameters, which represent the lengths of the edges, the right angles in the geometry are then preserved. Finally, there is also a data specification justification. Consider that a project requires the acquisition of buildings as LOD1 block models with a flat top surface, which regardless of the acquisition errors are then always block models. If in our approach the coordinates of the geometry were disturbed, the noise in the geometry would result in non-flat top surfaces breaking the specification, which is different from the reality.

We follow the assumption of uncorrelated errors in coordinates, as 3D city models are often acquired in different acquisition campaigns (e.g. footprints are acquired with a geodetic survey, while the elevation of the building is acquired with airborne laser scanning). However, we acknowledge that correlated errors may to some extent influence the outcome of the analysis, as demonstrated by Navratil and Achatschitz [959].

11.4.4 Varying uncertainty in 2D/3D and multiple accuracy classes

3D city models may be derived with different approaches involving diverse technologies, each with different capabilities when it comes to the accuracy. Hence it is important to investigate different magnitudes of positional error (standard deviation σ). Taking into consideration multiple accuracy classes also helps in understanding the impact that increasing the error of the input has on the error of the output.

As hinted at in multiple parts of the thesis, in 3D GIS, the accuracy $\sigma_{x,y}$ of lateral coordinates (X, Y) and the accuracy σ_z of vertical coordinates (Z) in the geometry are often different, hence another point that we consider is the varying level of accuracy in the planar and vertical coordinates ($\sigma_{x,y} \neq \sigma_z$). This varying accuracy is due to two reasons. First, 3D city models are often constructed by combining multiple

acquisition techniques and with the integration of data from different sources. This is often the case for LOD2 as well. The varying accuracies are especially emphasised in 3D city models derived with extrusion, and in cases such as in Chapter 7 where the 3D model has been derived from very accurate footprints, while the heights have been estimated from attributes resulting in incomparable accuracy levels[41]. Second, within the same acquisition technique, the accuracy levels may vary. This is an inherent property of 3D acquisition techniques such as laser scanning [540, 877, 912]. For example, the specification of the lidar system ALS70-CM from Leica [960] states 'The system produces data after post processing with a lateral placement accuracy of 5–38 cm and vertical placement accuracy of 7–16 cm (one standard deviation) [...]'. Hence, varying uncertainty levels are one of the important characteristics of 3D modelling to consider for uncertainty propagation.

With the exception of satellite platforms [21, 38], researchers regularly report submeter accuracy of 3D acquisition techniques [537, 961–964]. Therefore, for $\sigma_{x,y}$ and σ_z we define a series of 11 accuracy classes (ranging from 0.0 to 1.0 m with 0.1 m increments), and permute them resulting in 121 combinations, e.g. $\sigma_{x,y}$ = 0.3 m / σ_z = 0.6 m. The *0th* accuracy class ($\sigma_{x,y} = \sigma_z = 0$) has been considered as well in order to ensure that there are no errors in our implementation of simulating uncertainty.

Monte Carlo simulation requires that the model is run repeatedly with different perturbations [917], commonly in a large number m. Hence, in the workflow each building in each perturbation d is disturbed according to 121 different uncertainty classes, resulting in an erroneous building $\widetilde{F}_i^{j,k}, j = 1, \ldots, 121; k = 1, \ldots, m$. After the perturbations of the parameters, the simulated 3D city models are input in the spatial analysis (see Figure 11.6 for an example). Variability between solar irradiation resulting from the perturbation characterises how uncertainty about a building's geometry propagates to its solar irradiation.

Figure 11.7 hints at the reason why perturbations have to be executed in a large number of instances: the plot reveals that with the increasing number of perturbations the results improve and become more stable. However, that comes at a cost of longer simulations, and that differs between use cases.

Therefore, it is important to find an optimal balance between the number of iterations and duration of experiments. However, other factors should also be balanced in the equation, considerably affecting the load of the experiments since they multiply: number of considered LODs × number of considered geometric references × number of buildings × number of accuracy classes ($\sigma_{x,y}$, σ_z) × number of perturbations m (Monte Carlo simulations).

[41] While this case is common, also the opposite is possible. For example, a case in which the heights have been derived from lidar, but the footprints from an old coarse scale topographic map.

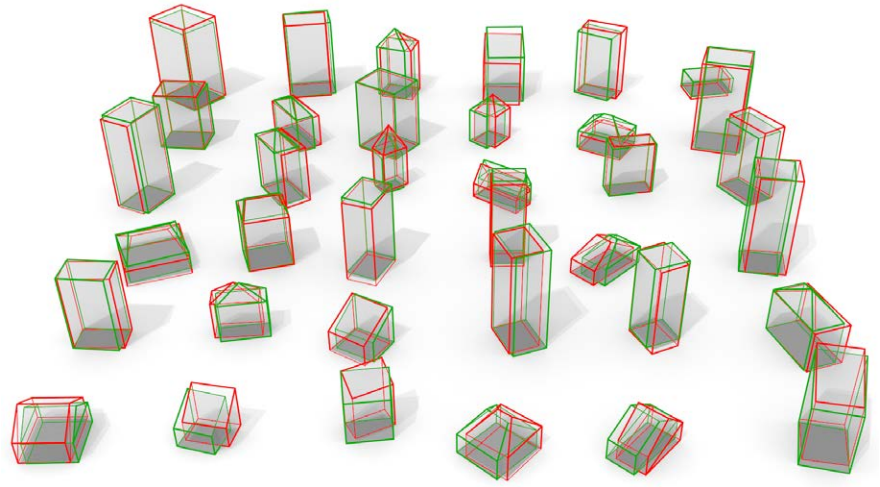


Figure 11.6: Composite of the geometry of two LOD2.1 models of the buildings shown in Figure 11.4. The green model is the ground truth, while the red model has been disturbed with errors using $\sigma_{x,y} = \sigma_z = 0.4$ m. Note that this is only one of the considered 121 accuracy classes, indicating the outcome of only one iterations in one LOD, meaning that in our experiments there are millions of such datasets.

This quantity of factors that each considerably increases the amount of data suggests that compromises have to be made to reduce the simulations to a reasonable amount of time. For example, keeping the number of 10000 buildings as used in the previous chapter would require years of simulations.

We have decided to pursue the following choices. First, the number of buildings was reduced to 100. It was found out that having 100 buildings instead of 400 or 10000 as used in the experiments in Chapter 10 still retains a diverse architecture and different cases without significantly compromising the results. Furthermore, while the procedural modelling engine produces clean data without topological errors, disturbing the geometry may induce topological errors, causing iterations to be repeated until a dataset without errors is generated. Our implementation includes a built-in validator according to international standards in GIS [633], therefore simulations that had topological errors were discarded[42] to avoid the introduction of inconsistencies other than positional errors. Hence another advantage of reducing the

[42] On a related note, it is interesting to mention that such faults, while unwanted in this research, turned out to be a very useful byproduct for a completely different purpose: as test datasets for the OGC CityGML Quality Interoperability Experiment [655].

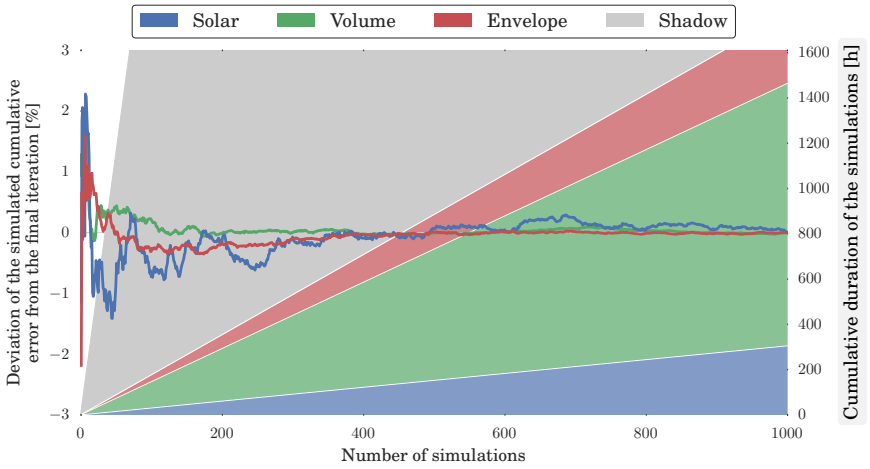


Figure 11.7: Performance of the Monte Carlo simulations. The estimated error (left axis) converges as the number of iterations increases, at the expense of longer simulation runtimes (right axis), potentially resulting in months of experiments to obtain satisfying results. The plot reveals that not all spatial analysis exhibit the same behaviour. The overly long run times disqualify the shadow spatial analysis (Chapter 10) from being used in these experiments.

number of buildings is to reduce the likelihood of topological errors leading to more robust simulations. Second, based on the plot in Figure 11.7 we have concluded that 1000 perturbations are sufficient because it appears that the results become stable at that point. An exception is the solar irradiation analysis because it appears that it requires more runs. However, this was not a nuisance because this spatial analysis is quicker than the one of computing the volume of buildings.

With these choices, the simulations took about 3 months, and resulted in the generation of 1.8 billion CityGML buildings. Each of these has been subject of three spatial analyses. It is clear that not many spatial analyses are possible with such a huge amount of data, and that not many of them allow the automatic interaction with the results. For example, experimenting with the estimation of noise is not possible (recall Figure 2.4b showing the analysis that took several hours for just one small spatial extent). On a related note, Figure 11.7 also reveals why experiments involving the shadow analysis cannot be performed—the analysis is too complex. The plot shows the duration for one sun position, indicating that the shadow analysis is on about two orders of magnitude slower than the other three spatial analyses if multiple sun positions are considered.

11.4.5 Errors and measures of uncertainty

We computed the RMSE of the outcome of the considered three spatial analyses. In this section we focus on the solar spatial analyses, hence we give the insight into the intermediary results of the spatial analysis—we computed the following error measures for each accuracy class and each LOD:

- RMSE of the estimated tilt, azimuth, and area of the roof surfaces.
- RMSE of the building rooftop annual irradiation in kWh/year.
- RMSE of the filtered building rooftop annual irradiation in kWh/year, i.e. those surfaces that fulfil the minimum thresholds of the normalised annual irradiation and area (Section 11.3.3).
- The share of false positives and false negatives in determining the feasibility of installing a solar panel (in %). This analysis is Boolean (a surface is feasible or not), hence we are computing the share of errors in the total number of estimations.

To put these values in perspective, for each RMSE, also the relative RMSE in percentage has been computed, as in Chapter 10. As ground truth we take the undisturbed model ($\sigma = 0$) of the building in that LOD because we are interested how do positional errors relatively affect an LOD. Furthermore, for all cases the distribution of errors is examined.

11.5 EXPERIMENTS, RESULTS, AND DISCUSSION

Due to the large number of results, we focus on the most important results only, and only on four LODs: LOD1, LOD2.1, LOD2.2 and LOD3. Figure 11.8 illustrates the propagation of positional error to the error in the estimation of the total building irradiation and to the filtered (feasible) value. Because of limited space we show this relationship for the 11 classes where $\sigma_{x,y} = \sigma_z$. The plot indicates that the relationship between the uncertainty in the input and uncertainty in the output is linear, and that the results for each LOD are practically equal, with the small exception of LOD1. The slightly smaller error of the LOD1 can be explained by the absence of errors in the tilt and azimuth (the top surface is always flat, as imposed by the model specification), hence only the uncertainty in the area has an influence.

Because the uncertainty propagation in LOD2.1, LOD2.2 and LOD3 is equivalent, from here we will present the results of one of these LODs. Figure 11.10 illustrates the propagation of positional uncertainty to the uncertainty of the three components required for computing the solar irradiation: azimuth, tilt, and area. The figure also

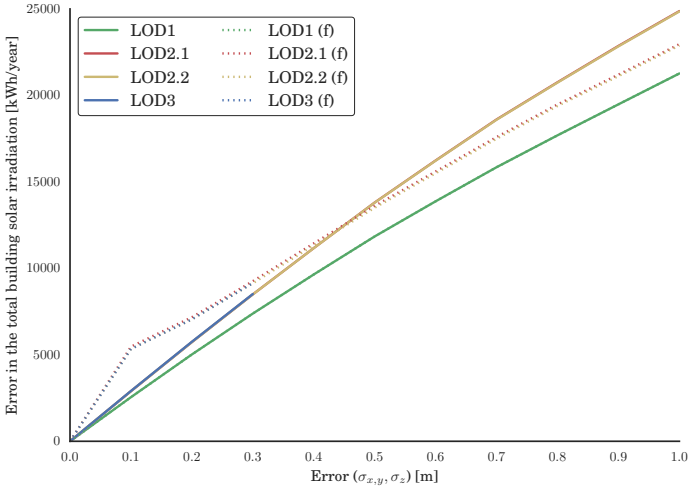


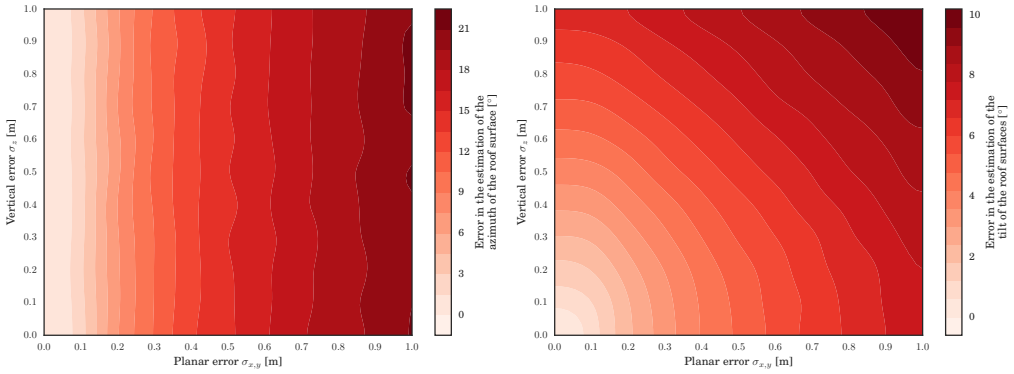
Figure 11.8: Propagation of positional error in case $\sigma_{x,y} = \sigma_z$ to the total (filtered) building irradiation. The first value is the sum of the insolation of all roof surfaces, and the second one is its filtered value, indicated by (f), where only feasible surfaces are taken into account. Note that the values for LOD2.1, LOD2.2, and LOD3 are equal.

reveals the RMSE of the normalised irradiation [kWh/m^2/year] for each uncertainty class. Note that in computing the value of the normalised solar irradiation, the area has not been used (which was used for the total building irradiation [kWh/year], shown later).

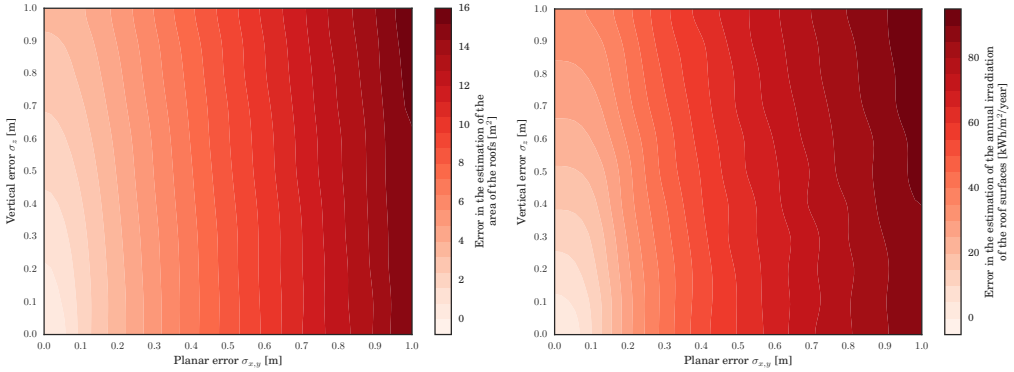
Since the values of the uncertainty in the solar irradiation may be of interest to practitioners, we provide them in Table 11.1. The plots reveal a different degree of the influence of the planar and vertical positional uncertainty. For example, the estimation of the orientation of the roof is not affected by the vertical uncertainty. Obviously, the exact results depend on the analysed setting (i.e. location, empirical model, etc.), but we believe that the general conclusions presented here are consistent for most, if not all scenarios.

Figure 11.11 and Table 11.2 provide the uncertainty in the estimation of the total solar irradiation of roof surfaces of buildings for all accuracy classes. The results reveal that at smaller planar uncertainties, the vertical positional uncertainty influences the total irradiation uncertainty, while for greater planar uncertainties its influence is negligible.

Figure 11.12 indicates the Type I and Type II errors for the estimation of the feasibility of installing a solar panel. The plot indicates a different pattern in the influence



(a) Error in the azimuth (σ_α). **(b)** Error in the tilt (σ_τ).



(c) Error in the area (σ_A). **(d)** Error in the norm. irr. (σ_{ir}).

Figure 11.10: Results of the error propagation experiments on the uncertainty of the three components of the estimation of the irradiation (azimuth, tilt, area), and the normalised irradiation. The plots indicate different sensitivity of particular geometric operations to planar and vertical error.

of planar and vertical errors. Furthermore, the plot signifies the impact of a particular degree of geometric accuracy of data on decision making about the suitability of solar panel installation.

Finally, Figure 11.13 shows the distribution of errors for various settings (LODs and uncertainties). It appears that regardless of the setting, the errors are distributed according to a symmetric probability density function. However, none of the distributions is normally distributed. This was tested with a normality test of D'Agostino [965] that combines skew and kurtosis to produce an omnibus test of normality, and

Table 11.1: Standard deviation of the estimation of the normalised irradiation of roof surfaces of buildings [kWh/m^2/year]. The error is estimated for the 121 accuracy classes. The bracketed values indicate the relative RMSE in percents.

σ_z [m]	$\sigma_{x,y}$ [m] 0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
0.0	0 (0)	10 (1)	28 (4)	42 (6)	53 (8)	60 (9)	67 (10)	74 (11)	80 (12)	84 (13)	90 (13)
0.1	4 (1)	11 (1)	28 (4)	42 (6)	53 (8)	60 (9)	67 (10)	73 (11)	80 (12)	85 (13)	90 (13)
0.2	8 (1)	13 (2)	29 (4)	42 (6)	52 (8)	61 (9)	68 (10)	74 (11)	79 (12)	84 (13)	89 (13)
0.3	12 (2)	15 (2)	31 (5)	43 (7)	54 (8)	62 (9)	68 (10)	74 (11)	80 (12)	85 (13)	90 (14)
0.4	16 (2)	18 (3)	33 (5)	45 (7)	55 (8)	62 (9)	69 (10)	76 (11)	80 (12)	87 (13)	90 (14)
0.5	19 (3)	22 (3)	34 (5)	47 (7)	55 (8)	64 (10)	69 (11)	75 (11)	82 (12)	87 (13)	91 (14)
0.6	23 (3)	25 (4)	36 (5)	47 (7)	57 (9)	64 (10)	70 (11)	77 (12)	82 (12)	87 (13)	91 (14)
0.7	26 (4)	28 (4)	39 (6)	49 (7)	58 (9)	66 (10)	72 (11)	77 (12)	83 (13)	88 (13)	93 (14)
0.8	29 (4)	31 (5)	40 (6)	51 (8)	59 (9)	66 (10)	73 (11)	79 (12)	84 (13)	88 (13)	93 (14)
0.9	32 (5)	33 (5)	42 (6)	53 (8)	60 (9)	67 (10)	74 (11)	79 (12)	84 (13)	89 (14)	94 (14)
1.0	34 (5)	36 (5)	44 (7)	54 (8)	62 (9)	68 (10)	74 (11)	80 (12)	85 (13)	90 (14)	94 (14)

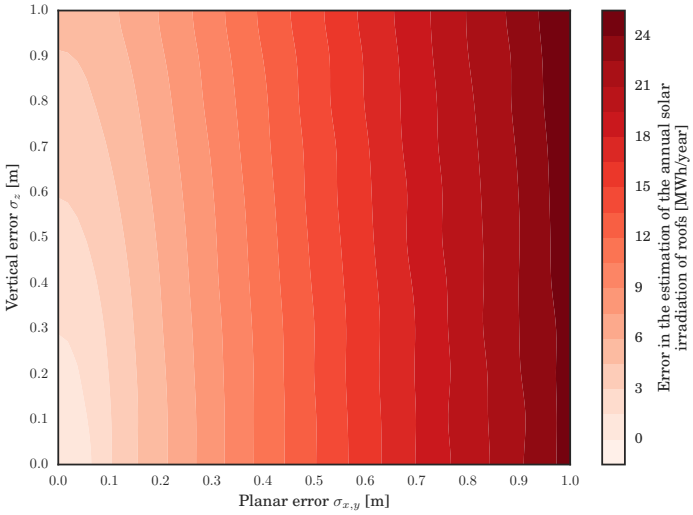


Figure 11.11: Error in the estimation of the total irradiation of a building (σ_1).

by attempting to fit a function (also shown in the Figure 11.13).

From the results, the following conclusions can be drawn:

- The propagation of errors is similar for LOD2 and LOD3. For LOD1 it is

Table 11.2: Standard deviation of the estimation of the total global solar irradiation of roof surfaces of buildings [MWh/year], estimated for 121 accuracy classes. The bracketed values indicate the relative RMSE in percents.

σ_z [m]	$\sigma_{x,y}$ [m] 0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
0.0	0.0 (0)	2.8 (5)	5.6 (10)	8.3 (15)	11.0 (21)	13.4 (25)	15.8 (30)	18.1 (34)	20.2 (37)	22.3 (41)	24.2 (44)
0.1	0.5 (1)	2.9 (5)	5.6 (11)	8.3 (16)	10.9 (20)	13.4 (25)	15.8 (30)	18.0 (34)	20.3 (38)	22.2 (41)	24.2 (44)
0.2	1.0 (2)	3.1 (6)	5.7 (11)	8.4 (16)	11.0 (21)	13.5 (25)	15.9 (30)	18.1 (34)	20.2 (38)	22.4 (41)	24.2 (44)
0.3	1.6 (3)	3.3 (6)	5.9 (11)	8.5 (16)	11.0 (21)	13.5 (25)	15.9 (30)	18.1 (34)	20.4 (38)	22.5 (41)	24.2 (44)
0.4	2.1 (3)	3.5 (6)	6.0 (11)	8.7 (16)	11.2 (21)	13.6 (25)	16.0 (30)	18.2 (34)	20.4 (38)	22.5 (41)	24.4 (44)
0.5	2.6 (4)	3.9 (7)	6.2 (11)	8.8 (16)	11.3 (21)	13.8 (26)	16.1 (30)	18.4 (34)	20.5 (38)	22.5 (41)	24.4 (45)
0.6	3.1 (5)	4.2 (7)	6.5 (12)	8.9 (16)	11.5 (21)	13.9 (26)	16.2 (30)	18.4 (34)	20.5 (38)	22.5 (41)	24.5 (44)
0.7	3.5 (6)	4.6 (8)	6.7 (12)	9.1 (17)	11.6 (22)	14.0 (26)	16.3 (30)	18.6 (34)	20.6 (38)	22.7 (41)	24.5 (45)
0.8	4.0 (6)	5.0 (8)	7.0 (12)	9.4 (17)	11.8 (22)	14.2 (26)	16.6 (31)	18.7 (34)	20.8 (38)	22.8 (42)	24.6 (45)
0.9	4.4 (7)	5.3 (9)	7.3 (13)	9.6 (17)	12.0 (22)	14.3 (27)	16.6 (31)	18.8 (35)	21.0 (38)	22.9 (42)	24.8 (45)
1.0	4.9 (8)	5.7 (10)	7.6 (13)	9.9 (18)	12.3 (23)	14.5 (27)	16.8 (31)	18.9 (35)	21.1 (39)	23.1 (42)	24.9 (45)

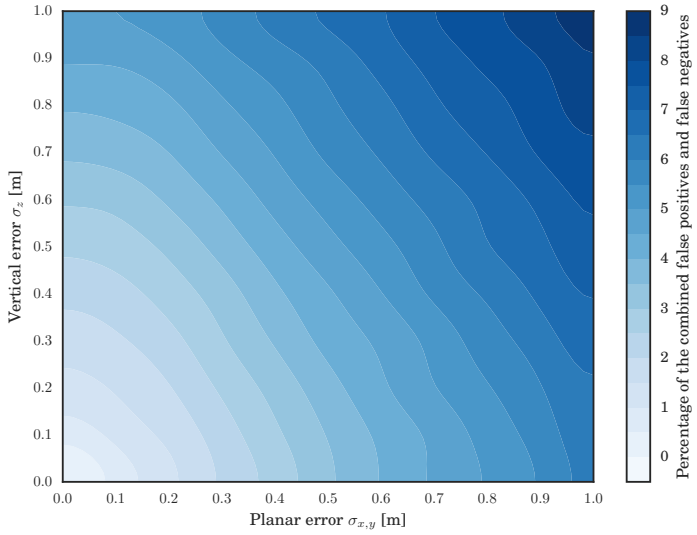


Figure 11.12: Likelihood of leading to a wrong decision. The plots illustrate the summed false positives (Type I errors) and false negatives (Type II errors) in the determination of the feasibility of a photovoltaic installation on a roof surface.

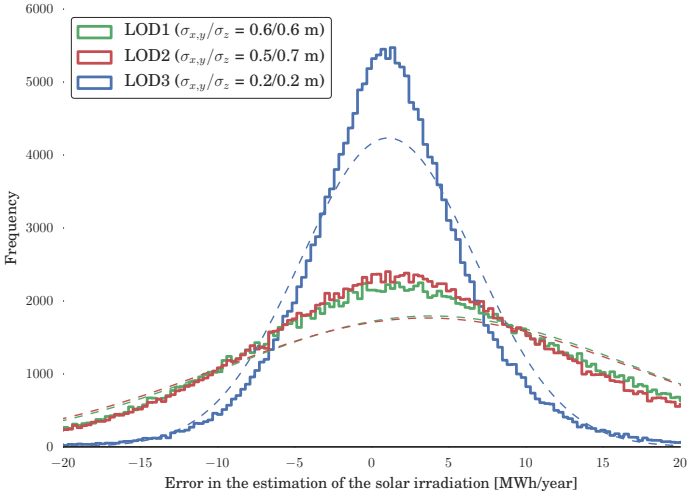


Figure 11.13: Distribution of errors in the estimation of the annual solar irradiation. The errors of three configurations (LOD and uncertainty class $\sigma_{x,y}/\sigma_z$) are illustrated, with the attempted fit of the normal distribution for each (dashed in the corresponding colour).

somewhat smaller, but the usability of these results is doubtful because LOD1 imposes flat roofs and involves a large model error that is not taken into account in the error propagation analysis (i.e. the error is calculated within each LOD, as the model with $\sigma = 0$ is taken as ground truth). Hence the conclusion is that data in different LODs is differently affected by positional error when used in a spatial analysis, but the difference is subtle and the bigger picture of errors should be taken into account (this if the topic of Chapter 12).

- LOD3 models are equally prone to errors as LOD2 models. Hence, it may not make sense to acquire a high LOD if the acquisition method is not very accurate. This finding is further examined in the next chapter.
- The planar and vertical accuracy have a different influence on the output uncertainty, hence it is important to consider these separately.
- The mean error deviates from zero in most cases, suggesting a systematic error in solar irradiation. This can be explained from the non-linear relationship between uncertain parameters and solar irradiation.
- The errors in the estimation of the feasibility of installing a solar panel on

a roof surface may reach approximately 10% at σ = 1.0 m (i.e. 10% of roof surfaces are deemed feasible for the installation of solar panels while in reality they are not, and vice-versa), indicating that in many occasions 3D models acquired in this and coarser range of uncertainty, may not be suitable for this use case.

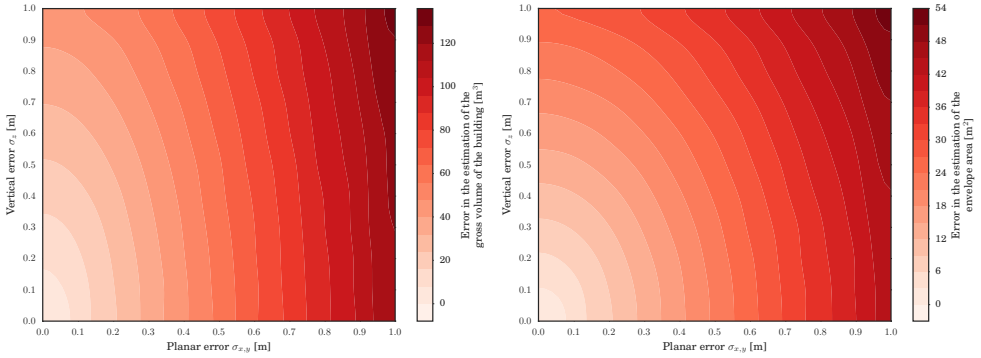
- Relatively, the propagation of error in the estimation of the normalised solar irradiation is smaller than in the estimation of the total building rooftop solar irradiation (normalised irradiation multiplied with the roof surface area), because the calculation of area is more sensitive to the uncertainty in the input data.

Now that we have covered the main results, we can focus on two spatial analyses from the previous chapter: estimating the gross volume and the envelope area of buildings. In summary, the main results are:

- The propagation of positional error also increases linearly in the considered LODs.
- The LODs are similarly prone to errors (interestingly, in these cases LOD1 is slightly more sensitive to positional errors than LOD2/3).
- The influence of planar and vertical in these two analyses is somewhat different than in the estimation of solar irradiation (Figure 11.15). This indicates the different sensitivity of planar and vertical coordinates for each spatial analysis.

11.6 CONCLUSIONS

In this chapter we have performed a 3D GIS error propagation analysis involving 3D city models and their application in three spatial analyses. The analysis involved the use of procedurally generated 3D city models, and it took into account multiple LODs in multiple accuracy classes with varying dimensional accuracy, all of which are novelties in the research field. The uncertainty propagation analysis enabled us to calculate the uncertainty in the estimated values on the specified uncertainty in the input data, and it gave us detailed insights, especially about the solar irradiation estimation. For example, the uncertainty analysis demonstrated that on average a 0.4/0.3 m positional error causes an uncertainty of about 8700 kWh/year in the estimation of the total annual insolation of the roof of a building, which is approximately a relative deviation of 16%. Another important insight is that the GIS components required for estimating the solar irradiation of a roof surface of a building (azimuth,



(a) Uncertainty in the gross volume.

(b) Uncertainty in the envelope area.

Figure 11.15: Results of the error propagation experiments on the uncertainty of the other two use cases. The plots reveal different sensitivity to planar and vertical errors, and call attention to run such experiments for each spatial analysis, as they may differ in propagation of error.

tilt, area), indicate a different sensitivity to the input uncertainty, and that the planar and vertical uncertainty have a different influence on the estimations.

The obtained results are of importance to practitioners that rely on 3D city models for the estimation of the solar irradiation of buildings (and other use cases as well). They need this type of information to be able to decide whether installing solar panels is economically attractive. The uncertainty quantification extends such an analysis by allowing us to take the risk of making wrong decisions into account. These observations may also be important to practitioners dealing with related applications of estimating the solar irradiation on a building, such as in urban planning [107, 224], for research into the thermal comfort of buildings [105, 106], indoor illuminance to determine daylight autonomy [110, 966], crisis management [108], and in the valuation of real estate [109].

Instead of using real-world data, again we used procedurally generated (synthetic) 3D city models, which proved suitable for this purpose. This chapter presents additional benefits of using procedurally generated data in 3D GIS. Besides of being an unlimited source of diverse data in multiple LODs, procedurally generated data are easier to disturb with positional errors simulated with a stochastic engine that is integrated in the architecture of the procedural modelling engine.

It is relevant to note that our instance takes advantage of the availability of LOD3, since in the computation of the area, it accounts for roof windows, roof overhangs, dormers and chimneys, which decrease or increase the area and configuration of

the roof that is usable for photovoltaic panel installations. Having the information about roof windows is important, but it is frequently not possible due to lack of highly-detailed data and this information is in practice supplemented manually with surveys on site [116].

The error propagation task is further impeded by the fact that software support is limited. In analyses such as this several building models are disturbed in hundreds of simulations, resulting in a large abundance of datasets that have to be analysed. Thus the capability to automatically and repeatedly load the data and analyse the results is essential. Moreover, the large number of model runs entails an increased computational cost, which can be substantial in some spatial analyses.

While we have focused on the three spatial analyses, a good deal of the developed work is applicable to other 3D use cases. However, in practice, many spatial analyses are not feasible for these experiments because they may entail very long simulation times.

For estimating the solar irradiation of roofs, I have implemented an experimental software prototype called Solar3Dcity. To make our method generally applicable, Solar3Dcity can be used for any location on Earth. Furthermore, this tool is released open-source for free public use, leading to being used by other researchers [967].

For future work it would be beneficial to store the obtained solar energy values as information in 3D city models, according to the concept of *dynamizers* introduced by Chaturvedi and Kolbe [968].

In the next chapter we will combine the insights obtained in this chapter and Chapter 10 in order to put in perspective both types of error for the three considered spatial analyses.

CHAPTER 12

Combining LOD and positional errors

This chapter is based on my paper [969]:

> Biljecki F, Heuvelink GBM, Ledoux H, Stoter J (2017): The effect of acquisition error and level of detail on the accuracy of spatial analyses. *Cartography and Geographic Information Science*, advance online publication. doi: 10.1080/15230406.2017.1279986

Chapter 10 analysed the influence of the LOD of the 3D city model on the results of a spatial analysis, while Chapter 11 examined the impact of positional errors within each LOD. This chapter revisits the previous two chapters by combining them, investigating the relative contribution of each error to put them in perspective. Comparing the two types of errors is important for different reasons. For example, it may help practitioners deciding whether it is more beneficial to invest into a more detailed dataset rather than a more accurate one. The chapter presents an error framework that separates the two types of errors, and it realises it by simulating both types of errors simultaneously. The experiments, performed on the same three spatial analyses as in the previous chapter, demonstrate the magnitude of the two types of error, and show how they individually and jointly propagate to the output of the three spatial analyses considered in the previous chapter. The most notable result is that in most cases the positional error has a significantly higher impact than the LOD. As a consequence, it is suggested that it is pointless to acquire geoinformation at a fine LOD if the acquisition method is not accurate, and instead it is advised focusing on the accuracy of the data. A contribution of this work in the error propagation field is that a multiple error propagation analysis is performed.

12.1 INTRODUCTION

The previous two chapters have investigated the influence of the LOD and positional errors on the result of a spatial analysis. These different qualities affect spatial analyses in distinct ways, and investigating the propagation of a specific type of error (e.g. thematic error) has been extensively researched in geographical information science. However, mixed error propagation studies, which analyse the joint propagation of multiple error types, are virtually non-existing. Error propagation analyses commonly focus on one type of error and on one spatial analysis, and they are never carried out at multiple scales. This prevents the understanding of the relation, magnitude, and relative contribution of each type of error.

In this chapter we again focus on the errors induced by (1) different LODs and by (2) positional errors incurred by the acquisition. However, we run joint experiments to isolate and quantify them, and to investigate whether the benefit provided by spatial data at finer LOD is still valid in cases of significant acquisition errors.

Understanding the relationship between detail and acquisition error is important for stakeholders in GIScience in order to put the two quality characteristics into perspective. For example, the presented approach provides practitioners and scientists a way of determining whether it is worth increasing the accuracy of the dataset, or rather its LOD, when designing the specification of a dataset to be acquired, so that the produced data will be suitable for a specific purpose (e.g. 'What should the minimum accuracy and level of detail available in the data be, so that these are usable for accurately calculating the volume of buildings?'). Likewise, it is relevant to set expectations about the capabilities of a certain dataset: this involves determining whether a dataset is adequately detailed and accurate enough to derive sufficiently reliable results in a spatial analysis. For example, a user can avoid ordering the acquisition of an expensive and overly detailed dataset, which in a certain spatial analysis brings only a minuscule benefit when compared with a less detailed and less costly alternative.

The type of questions that we address in this chapter are usual considerations for GIS users:

- Given two distinct datasets covering the same area (from multiple sources), where one is less detailed but more accurate than the other, which is the better choice for a particular spatial analysis?
- At what LOD and at what accuracy should a 3D model be acquired to be usable for a particular spatial analysis? Understanding this aspect would aid data producers in designing a specification that bears in mind the intended use of the data.

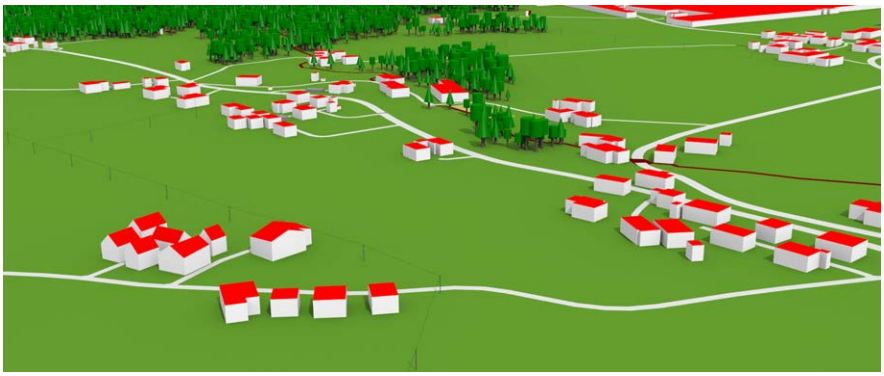


Figure 12.1: 3D building models of heterogeneous lineage within the same dataset. Example of a 3D model of a village in Austria obtained with crowdsourcing. In reality, almost all buildings in the area have pitched roofs, however, they are modelled only in a few buildings on the left (LOD2). The rest of the buildings are represented by simple block models (LOD1).[43]

- Is it beneficial to acquire a dataset at a fine LOD if the acquisition technique has poor accuracy? Understanding this aspect may prevent wasting effort to produce a dataset that is detailed and it is perhaps visually pleasing, but is ultimately not acceptable for a particular spatial analysis because of poor accuracy. This reasoning was also described by Burrough and McDonnell [970]: 'The quality of GIS products is often judged by the visual appearance of the end-product [...]. Uncertainties and errors are intrinsic to spatial data and need to be addressed properly, not swept away under the carpet of fancy graphics displays.'.

This research is also relevant in regard to the increasing availability of datasets with heterogeneous quality [971]. An example of such a dataset is one based on old but accurate cadastral data, in which newer buildings have been supplemented with other acquisition techniques such as footprints digitised from aerial images. Such approaches may result in data of variable accuracy and differing LODs, a phenomenon inherent to volunteered geoinformation [509, 545, 672, 972, 973]. Because such datasets are becoming increasingly used for spatial analyses [668], it is worthwhile to investigate for which portions of the dataset caution should be exercised due to lower reliability than in other parts of the dataset. Figure 12.1 illustrates an example of such dataset.

Our experiments help in determining which LOD and accuracy are sufficiently

[43] Data courtesy of OpenStreetMap contributors and OSM2World.

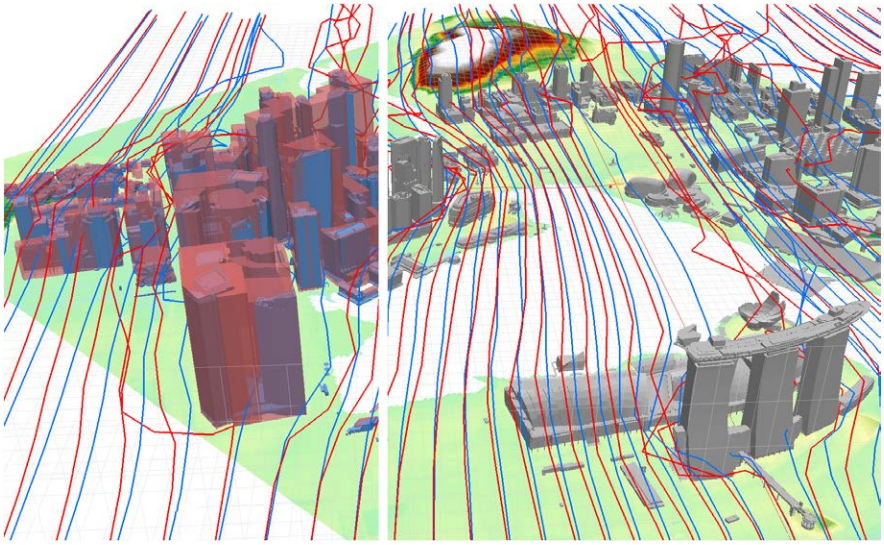


Figure 12.2: The results of two wind flow analyses of the Singapore central business district. One analysis is carried out on an accurate and detailed 3D city model illustrated here (LOD2.3; flow lines in blue), and another one with a crude block model (LOD1.2; flow lines in red). The left panel illustrate both LODs in corresponding colours, while the right panel reveals only the data at the finer LOD for clarity.[44]

acceptable for particular spatial analyses, e.g. which buildings in a given dataset can or cannot be used for a particular purpose.

Figure 12.2 illustrates the goal of this final chapter: to run the same spatial analysis (estimating wind flow) on two datasets with different LOD and positional accuracy. While the difference between the two LODs is obvious, it cannot be determined whether the improvement in the results should be attributed primarily to the progression of the LOD or to the increase in the accuracy of the data. Furthermore, due to the absence of ground truth data, it is not clear whether the improvement that a more detailed dataset brings is still advantageous, as it may still deviate considerably from the real-world.

In Section 12.2 we introduce a theoretical framework and an overview of related work. Section 12.3 presents the design of a method to decompose and quantify the two types of errors under consideration. As in Chapter 11, we select three spatial analyses (estimating the area of the building envelope, gross volume of a building,

[44] Data, analysis, and image courtesy of the Singapore Land Authority.

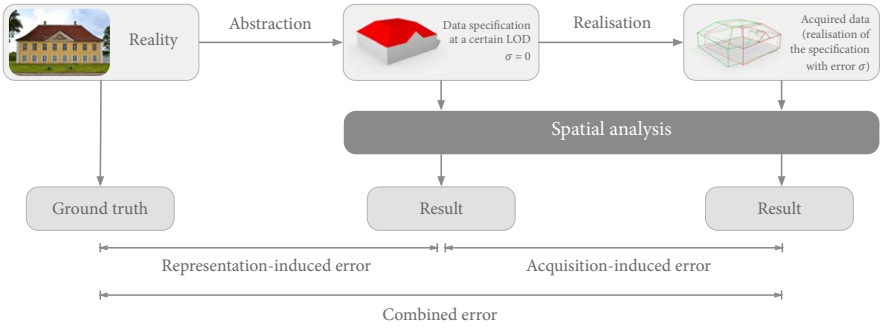


Figure 12.3: Longitudinal error decomposition as discussed in this research. Errors are induced at different stages of a typical GIS process, most prominently errors induced by abstraction and by realisation of the data. In reality it would include additional errors; see the overview of Lunetta et al. [974].

and solar irradiation of roofs) in order to investigate their different behaviours. We investigate whether these spatial analyses are more sensitive to positional error or to the reduction in the LOD. The results are presented in Section 12.4.

12.2 BACKGROUND: DECOMPOSITION OF ERRORS AND RELATED WORK

12.2.1 Decoupling errors

We decompose the errors induced in a typical GIS process into multiple components. Figure 12.3 illustrates our standpoint based on the work presented in the previous chapters: error is induced before any acquisition has even taken place because a specification is designed to capture a certain subset of reality at a certain LOD. For example, a data producer may decide to model buildings with simple roof shapes, without openings and finer details; such an LOD induces an error which we call *representation-induced error*.

The second step in the process is to *realise* the specification with data acquisition techniques. Due to the imperfection of measurements, several types of errors are induced in the process and the results of a spatial analysis are further degraded. When such errors propagate through a spatial analysis, we call this as *acquisition-induced error*. Here we focus on the positional error, as in Chapter 11.

We term the combination of the acquisition-induced and positional errors as *combined error*. The two types of errors have been introduced in Chapter 10 and Chap-

ter 11, hence they do not require an introduction and overview of related work.

12.2.2 Multiple error propagation analyses

In this section we focus on overviewing instances that are to an extent related to our research, which is rare because virtually all error propagation analyses focus on one type of error. In contrast, our analysis considers two types of errors. To the extent of our knowledge, we are aware of only a few analyses that investigate multiple types of errors, i.e. multiple error propagation analyses. Moreover, not all of these investigate the error propagation simultaneously, i.e. in most cases a separate analysis is made for each type of error.

Shi et al. [975], Couturier et al. [976], and Tayyebi et al. [977] investigate the combined effect of positional and thematic error in land cover maps. Rios and Renschler [978] mix probabilistic and fuzzy positional error models to expose the error in the detection of the contamination of groundwater. Lee et al. [979] examine the propagation of error in blood lead-level measurements of children and the locations of their residential addresses. Their analysis reveal the error in the aggregated results per census block.

An effort that is to some extent related to ours is the recent paper of Leao [980] examining the trade-off between spatial resolution and the quality of climate data. The analysis is performed on 2D raster data. A characteristic of the data is that due to interpolation the relationship between resolution and quality is not consistent, and the paper seeks to find the balance between the two.

12.2.3 Analysis-induced error

A spatial analysis per se is not faultless: no matter how accurate and detailed a dataset we have at our disposal, there will usually be error induced by the imperfection of the empirical models and other factors behind a spatial analysis. Usually these differences are due to external factors that are not influenced by geographic information. Here we do not focus on such an error, but we deem that this type of error has been overlooked in related work, and it is important to acknowledge its presence by dedicating a few paragraphs to it. A few examples follow:

1. Chapter 9 has indicated that based on the distribution of building stock population estimation can be conducted. However, different factors such as vacant buildings and variable apartment densities due to socio-economic aspects affect the accuracy estimates. These factors are typically not included in an analysis and hence invoke an analysis-induced error.

2. Chapter 11 has demonstrated how to estimate the solar irradiation of building rooftops for determining the suitability of installing photovoltaic panels. However, estimation models are empirically derived, and use other data that are prone to errors (e.g. cloud cover data). Besides the imperfection of the empirical models, each year is subject to different atmospheric conditions. Again, all of these factors are beyond the scope of the quality of geographic information and are typically ignored in a GIS analysis.

3. Geographic information may be used to predict the energy demand of households based on the morphology of a building, among other factors [121, 624, 981]. However, such predictions are sensitive to building occupation, energy consumption habits and lifestyles of occupants, and differences in insulations of homes; which are regularly not included in the modelling [982–984].

While spatial analyses have been extensively researched, surprisingly they are rarely validated using true data, most likely due to a variety of reasons. Foremost, the true value of a specific phenomena is frequently absent as the exact value is typically unobservable [985], or it is not feasible to acquire it, as large scale validation utilising more accurate data is expensive and laborious. For example, in order to validate the analysis-induced error of solar potential estimates it would be required to gauge the output of a myriad of solar panels or to place instruments (e.g. pyranometers) on many roofs [86, 862, 986]. Hence, when such scarce studies are available, they are limited to small sample sizes.

Another reason for the infrequency of such studies is that the output of a spatial analysis using real-world data already contains different types of errors (see again Figure 12.3). Determining the analysis-induced error is difficult because it may not be possible to isolate other errors from the error budget. For example, Brito et al. [987] assess the performance of using lidar data to predict the sky view factor, by comparing measurements with estimates derived with other methods. Here it is not possible to deduce whether the performance improves if the density of the point cloud (akin to scale or LOD) is augmented. Furthermore, there is no proof that the measurements that represent the ground truth are of an order of magnitude more accurate to warrant their role. Researchers working on other spatial analyses such as estimating the residential stock also cite this problem [129].

This topic is important to keep in mind when assessing the propagation of positional and LOD errors.

12.3 DATA AND METHOD

The method used in this research is the same as in the previous chapter: a 3D building model in multiple LODs is intentionally degraded with simulated acquisition

errors in repeated iterations (Monte Carlo simulation), and it is used in multiple spatial analyses. The results of the spatial analyses using the erroneous models are compared with the results of their error-free counterparts. Furthermore, to keep things straightforward and because of the small differences between LODs, as discussed in the previous two chapters, here we focus on three representations: LOD1.2, LOD2.2, and LOD3.3. For the geometric references we use the optimal GRs found in the experiments in Chapter 10.

The main difference with the previous chapters is that in this chapter we overview the bigger picture and take into account both errors. For the ground truth we take the results of LOD3.3. This representation marks the boundary of 3D GIS, as data of finer detail are considered to belong to the BIM arena. Moreover, LOD3.3 models are rarely used in spatial analyses and we are not aware of any LOD3.3 model produced on a large spatial scale due to excessive costs of acquisition. Hence LOD3.3 is a good choice as ground truth reference data.

A particularity of our three spatial analyses is that they are prone to positional errors that affect *deformable* objects (whose relative position can vary under uncertainty, e.g. the width of the modelled building may be smaller than it is in reality). However, the analyses considered here is not affected by errors related to positional error in *rigid* objects (e.g. displacement of a building by 20 meters due to processing errors does not alter its volume). For more on this topic the reader is referred to Heuvelink et al. [952].

An example of a part of the simulations is illustrated in Figure 12.4.

12.4 RESULTS AND DISCUSSION

Due to the large number of results we specifically focus on the most important findings. We present the results graphically in plots and tables, and describe them in the text. Furthermore, in order for a direct comparison of results between different spatial analyses, we present the errors in percentages of the true value, as in the previous chapters. Some of the results were already given in the previous two chapters[45] , but here we isolate them for a better overview.

[45] Chapter 10 gives the results for the spatial analyses when used with a dataset of 10 000 buildings. As discussed in Chapter 11, for the error propagation analysis the dataset has been reduced to 100 buildings to balance the complexity and duration of the simulations. For that reason the figures do not perfectly correspond, as there is a deviation in the results of less than 0.5%. This is also an interesting insight in the discussion of what is the minimum number of buildings to use in the experiments.

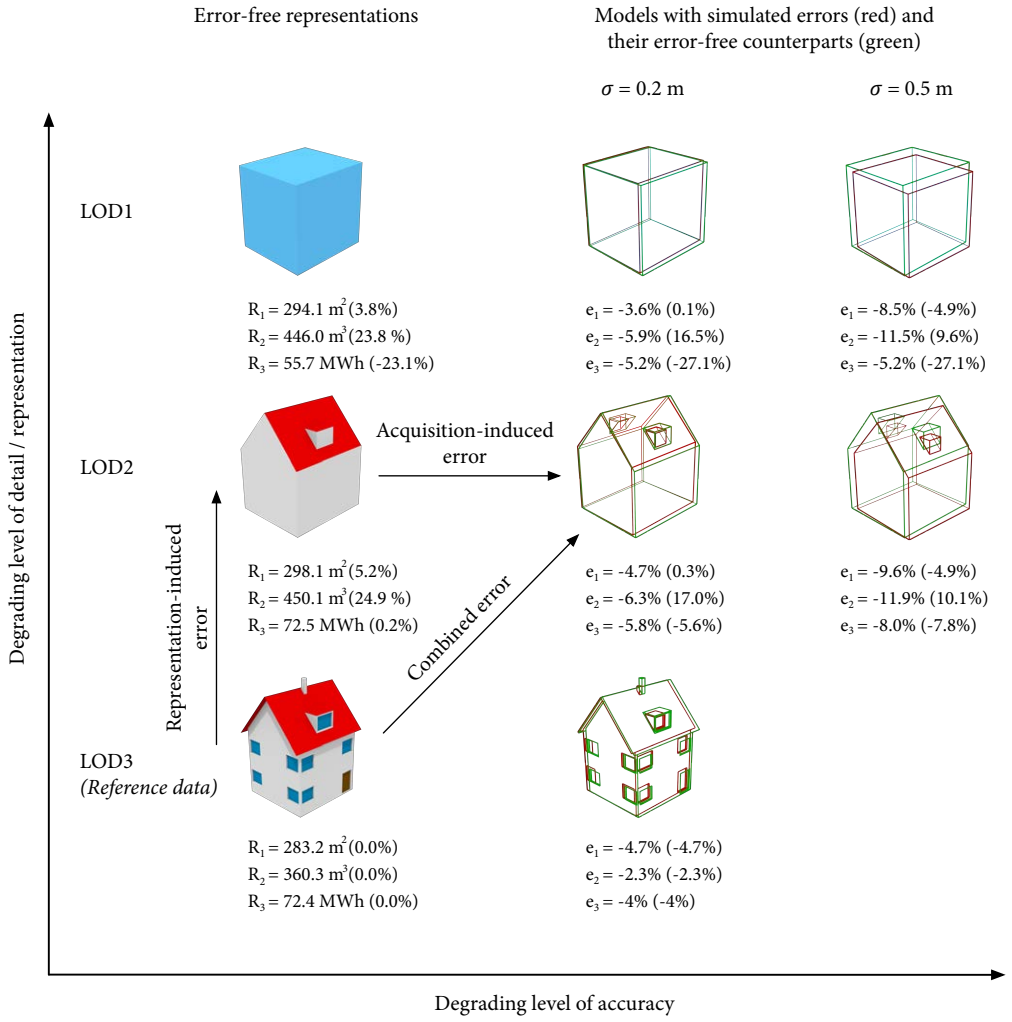


Figure 12.4: Illustration of a subset of the experiments and results of one simulation. This example shows a sample of one 3D building model disturbed according to two different levels of accuracy ($\sigma_{x,y,z}$ = 0.2 and 0.5 m). The results from the three considered spatial analyses are listed in the figure, and two types of errors are given for each case: the acquisition-induced error, and the combined-error (in parentheses). This particular case is interesting because in the first two spatial analyses the LOD1 model inherently results in less inaccurate results than the LOD2 model.

Table 12.1: Representation-induced errors in the three considered spatial analyses.

	RMSE in spatial analyses [%]		
Representation	Envelope	Volume	Solar
LOD1	7.9	3.7	14.2
LOD2	7.4	0.0	3.4
LOD3	0.0	0.0	0.0

12.4.1 Representation-induced errors

Table 12.1 reveals the magnitude of errors induced by the representation. As previously discussed, the findings indicate that LOD2 is a better choice than LOD1 in all three spatial analyses, as it resembles the abstracted phenomena in more detail. However, it is visible that in the first two spatial analyses the difference between the two is relatively small. Hence, it might not be justifiable to acquire a finer model. LOD2 may come at a significantly higher cost but for a marginal improvement.

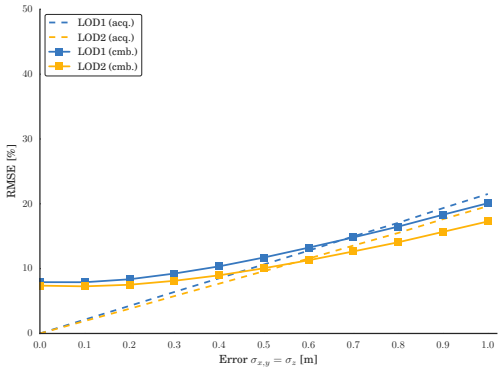
The results for the computation of volume are interesting: LOD2 conceptually derives the same volume as LOD3, because the details brought by LOD3 (e.g. windows, facade details, awnings, chimneys) do not bring any difference to the computation of volume. Hence, ignoring acquisition errors at this point, it appears that LOD3 does not bring any benefit over LOD2 when it comes to the computation of gross volume.

12.4.2 Acquisition-induced errors

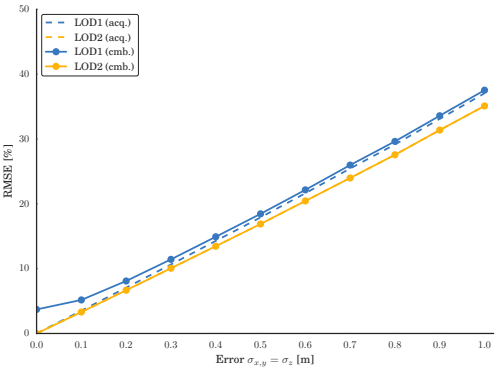
The results given in Section 12.4.1 indicate the magnitude of error if the models were (theoretically) acquired without acquisition errors. This section first considers the effect of positional errors in isolation, as in Chapter 11. These effects are visualised in the dashed lines of Figure 12.6, for each spatial analysis separately.

The results indicate that the error propagates linearly in all three analyses, but that it has a different impact on the final result. For example, the error induced by acquisition scenario $\sigma_{x,y}$ = 0.4 m / σ_z = 0.2 m for the three analyses, is 6.6%, 12.6%, and 20.8%, respectively.

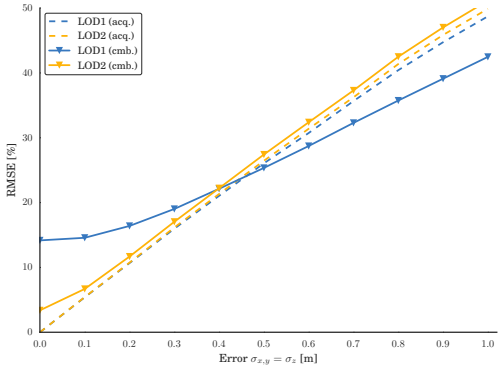
The propagation of positional error is similar between LODs in all three spatial analyses. However, notice that in the first experiment the error for LOD1 (dashed blue line) is larger than that of LOD2 (dashed yellow line), which is not the case for the third spatial analysis. This result suggests that positional error affects different LODs in different spatial analyses in different ways. Nevertheless, these differences are tiny.



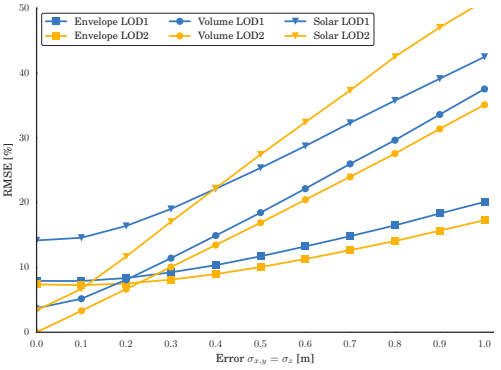
(a) Experiment 1: envelope.



(b) Experiment 2: volume.



(c) Experiment 3: solar.



(d) Combined errors of the 3 analyses.

Figure 12.6: The propagation of varying positional error per each spatial analysis. The results indicate different behaviour of each spatial analysis, preventing a generalised conclusion. The solar spatial analysis exhibits a paradox: initially the representation-induced error is very large since most of the roof shapes inherently significantly deviate from the real situation. However, at a larger positional uncertainty the combined error is notably smaller than with LOD2.

Table 12.2: Comparing RMSEs (%) of two datasets with opposite qualities. A fine detailed (LOD2) model acquired with poor accuracy, and a coarse model (LOD1) acquired with higher accuracy.

	LOD1 σ = 0.2 m			vs	LOD2 σ = 0.5 m		
Spatial analysis	R	A	C		R	A	C
Envelope	7.9	4.3	**8.4**		7.4	9.6	**10.1**
Volume	3.7	7.1	**8.1**		0.0	16.9	**16.9**
Solar	14.2	10.7	**16.4**		3.4	26.6	**27.5**

Key: R–Representation-induced error, A–Acquisition-induced error, C–Combined error.

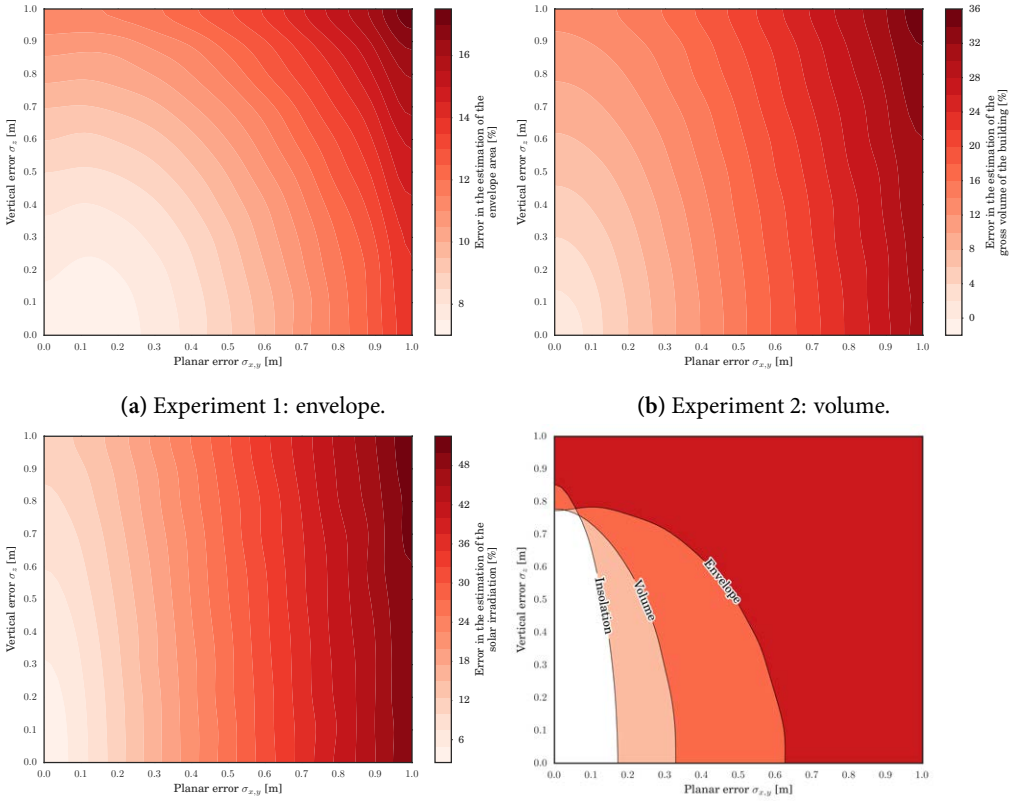
12.4.3 Two errors in combination

Figure 12.6 illustrates the combination of the two errors as solid lines. In addition, the plot in the bottom right shows the combined errors jointly for the three spatial analyses for comparison. The representation-induced error is also provided in the plots: this is the case where $\sigma = 0$.

The results suggest that the combined effect of the two error sources is not additive but is much more complex than that. The reason has to be studied in further research. The results also indicate that LOD2 is in most cases better than LOD1 by a thin margin, which means that despite the added positional error, the finer LOD2 still offers a slight benefit over the coarser LOD1.

However, the propagation of error in the third experiment gives unexpected results. LOD1 has an unfavourable starting point (the representation induces gross errors owing to flat rooftops), but eventually at σ = 0.5 m it surpasses the accuracy of the analysis with LOD2, probably owing to the more complex geometry of LOD2. A second unexpected result occurs in the first experiment (envelope area): the acquisition error (dashed line) is larger than the combined error (solid line), for both LOD1 and LOD2. These results indicate the presence of a systematic error.

Recall the dilemma discussed in the introduction in regard to using a dataset at a finer LOD but of lesser accuracy in contrast to the inverse situation. In our experiments an LOD1 acquired with σ = 0.2 m is a notably better choice than the finer LOD2 acquired with poorer accuracy (σ = 0.5 m); see Table 12.2 for comparison of all three considered spatial analyses.



(a) Experiment 1: envelope.

(b) Experiment 2: volume.

(c) Experiment 3: insolation.

(d) Joint plot of 10% RMSE threshold.

Figure 12.8: Contour plots indicating the influence of variable accuracy levels onto the error of the three spatial analyses for LOD2. Note that the colour ranges vary among plots. The bottom right plot reveals the differences in error propagation by illustrating the sensitivity of the spatial analyses at a certain threshold (value of 10% RMSE).

12.4.4 Influence of differing planar and vertical error

To retain the simplicity of the presentation, the errors so far have been considered with equal magnitudes ($\sigma_{x,y} = \sigma_z$). This section analyses the propagation of varying error magnitudes in the planar and vertical coordinates.

The behaviour for all three experiments is similar, both for LOD1 and LOD2. We therefore only present plots for the combined error in LOD2: see Figure 12.8.

It appears that the varying levels of planar and vertical accuracies have different impacts on the considered spatial analyses. The insights into the impact of planar and vertical accuracies, as outlined in Figure 12.8, may guide choosing the proper acquisition approach that warrants that the obtained 3D city model yields results with an error lower than a certain threshold. For all analyses, planar error has a larger effect than error in the vertical coordinates. However, the degree of such an influence differs, and this behaviour is mostly exhibited in the estimation of solar irradiation.

12.4.5 Influence of the building form

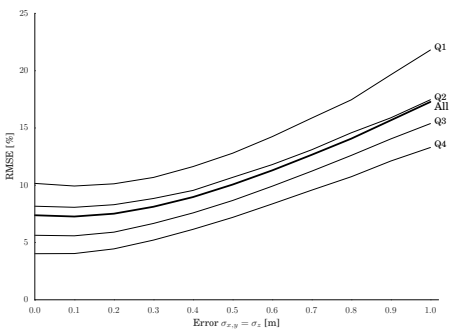
We noticed that errors (both in relative and absolute terms) substantially depend on the morphology of buildings. We therefore divided the buildings into four quartiles based on their volume. Figure 12.10 highlights the behaviour of errors for each quartile in each spatial analysis and different type of error. These plots clearly indicate that when it comes to relative errors, they are larger in smaller buildings. However, in absolute terms the behaviour of errors is opposite: the errors increase with the increase in the building size. The small exception in the third experiment (right plot in the bottom row; showing different order for Q3 and Q4) is caused by the varying degrees of insolation of roof surfaces. That is, rooftops with smaller areas may have a higher amount of solar irradiation than larger rooftops, so that rooftops of smaller buildings may have a larger absolute error in solar irradiation than rooftops of larger buildings.

These results imply that the outcome of analyses such as these also depend on the base dataset that is used, primarily because these are driven by the morphology of the buildings. Hence, the grammar of the procedural modelling engine has to be tuned according to the architecture that is intended to be analysed. In 2D this topic has been investigated by Berk and Ferlan [988], pointing out that the size and the shape of a parcel may characterise the propagation of error when calculating its area.

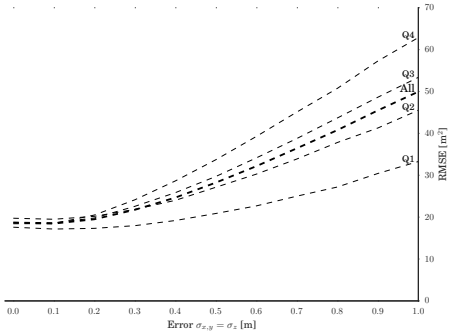
12.4.6 General discussion and key findings

A major finding of this chapter is that taking care of the accuracy of the data is more important than striving to produce data at a finer LOD, at least in the spatial analyses that we considered.

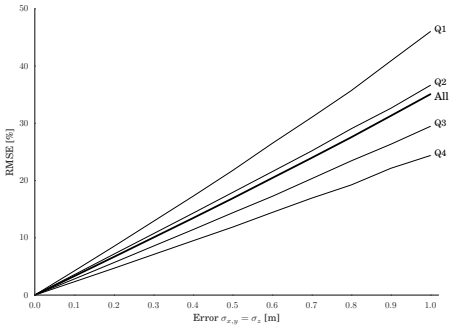
LOD1 and LOD2 are significantly different models—they are acquired with different approaches with the latter being more complex to produce. Despite such a distinction, when used in the first two spatial analyses (envelope area and gross volume) the difference in the performance of LOD1 and LOD2 is so small that it appears that in many cases it is not worth acquiring an LOD2. For example, when an LOD1



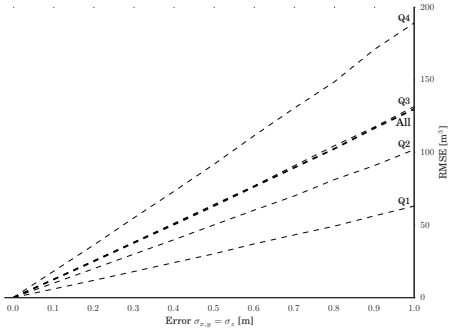
(**a**) Quartiles of exp. 1: relative errors.



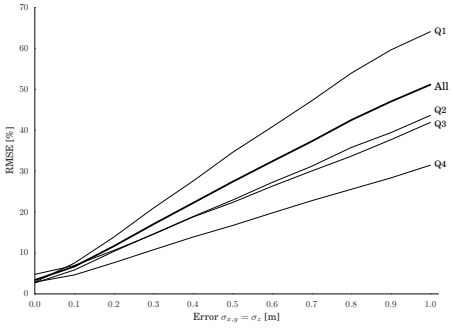
(**b**) Quartiles of exp. 1: absolute errors.



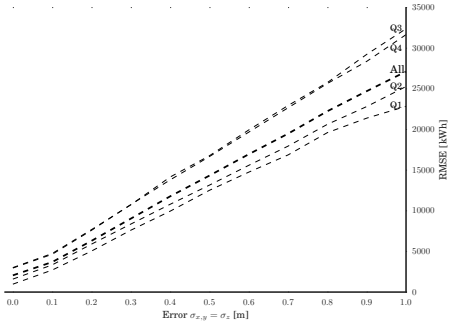
(**c**) Quartiles of the exp. 2: relative errors.



(**d**) Quartiles of the exp. 2: absolute errors.



(**e**) Quartiles of the exp. 3: relative errors.



(**f**) Quartiles of the exp. 3: absolute errors.

Figure 12.10: Dissecting the combined errors: building size influences the results. The plots on the left side indicate the errors in percentages, while the plots on the right express the errors in the units of measurements. In the plots on the right the quartiles are in the same order as on the ones on the left.

is used in the estimation of envelope area the RMSE is 7.9%, and LOD2 shaves off the error to 7.4%, which is practically negligible, an observation that was also discussed in Chapter 10. In such cases it may be more favourable to use the coarser LOD1–they are easier to acquire, they have a smaller storage footprint, and they are faster to process. Hence the increased costs for obtaining the finer LOD2 may not always be justified.

The leap between LOD2 and LOD3 is significantly larger than between LOD1 and LOD2. Hence it would be beneficial to strive towards the large-scale production of LOD3. However, such an advancement does not yet appear imminent: while many LOD3 models of limited spatial extent have been produced and used in various analyses, their production will remain expensive and wide coverage will not be feasible for some time.

Taking into account the realisation of the 3D city models, positional error has a substantial effect on the errors of the quantities estimated in this study (building envelope area, gross volume, and solar irradiation of rooftops). In these cases, positional error dominates over the error induced by a coarse LOD. A fairly small error of 0.2 m outweighs the benefit of a fine LOD, and our results indicate that in two of the three considered spatial analyses, an LOD1 acquired with σ = 0.2 m is a better choice than the finer LOD2 acquired with poorer accuracy (σ = 0.5 m). However, this is not the case for the solar irradiation of rooftops use case, in which LOD1 cannot be used due to its systematic shortcoming of having flat roofs. A paradox in this spatial analysis is that at poorer accuracies the error using LOD1 is smaller, due to the high sensitivity of solar irradiation estimation to positional errors.

The combined error cannot be easily decoupled into representation- and acquisition-induced error because they do not sum up. This is also obvious from the errors given in Figure 12.4, see for instance the case of LOD1 disturbed with σ = 0.5 m. Such a result indicates that there are interactions between the errors, as they are not additive.

These experiments provide insight into designing specifications of 3D city models while taking into account the intended spatial analyses; the increase of detail does conceptually bring some benefit, but in practice, when models are realised (and hence affected by the imperfection of measurement) the benefit is countervailed by acquisition errors. As a consequence, the representation-benefit between LOD1 and LOD2 becomes negligible.

The results also suggest that each spatial analysis has different behaviour. Hence, it is important to consider each spatial analysis separately in experiments. As the work also suggests that spatial analyses have a substantially different behaviour when compared to each other, data suitable for one spatial analysis may be of little value for another.

Real-world data offer little room for manoeuvre in experiments such as ours,

hence we suggest researchers in related work resorting to procedurally generated models as their benefit in such analyses is underestimated and unparalleled. By using a procedural approach we were able to obtain models that are burdened only with the errors we want (e.g. we did not have to worry that other errors such as completeness could have compromised our analysis), and by inducing specific errors in a simulation we could isolate the influence of different errors. On the other hand, it should be noted that synthetic data might not always properly represent real-world data.

A limitation of our study is that we do not address the inability of software to take advantage of finer LODs. For example, in theory there could be a substantial difference between using LOD1 and LOD2 for estimating noise pollution (e.g. sloped roofs henceforth available in LOD2 may bounce the noise in a different way resulting in substantially different predictions; see the work of Van Renterghem and Botteldooren [989], which demonstrates that roof shapes are an important factor to consider in noise pollution estimations). However, the software may not be capable of taking advantage of the more detailed geometry of the rooftops (e.g. it considers only the bounding box of a building), and will give the same results as for LOD1.

Looking into this matter is certainly our priority for future work, and for this, we plan to follow the approach of Ruiz Arias et al. [784]. In their analysis involving rasters of multiple resolutions, they compare the results from multiple software packages and conclude that some software solutions lead to larger error propagation. In our analysis we have dealt with volume and area computation, which should not differ between different software packages (when we programmed the two we compared the results from another software for validation). However, this is probably not the case for the solar irradiation use case, especially because due to computationally intensive calculations it was not possible to take into account shadowing effects. In the context of this chapter, the absence of shadowing is an example of a factor in the analysis-induced error.

In the context of analysis-induced errors it is also worthwhile to revisit the population estimation analysis presented in Chapter 9. The finest LOD that was used for these experiments is LOD1, and it has exhibited gross deviations in the estimation of population. These deviations appear not to be caused by the coarse LOD1, but rather by a gross analysis-induced error. Based on the experiments of calculating the volume, we can take the liberty to speculate that having an LOD2 for population estimation would make virtually no difference because it would improve the volume estimations (the main input for population estimation) only by a fraction.

Another aspect that we did not address, is that it is not always possible to separate the LOD and positional errors. Sometimes they are 'fused'; e.g. a complex building footprint may be simplified as a rectangle, and at the same time its geometry may be displaced due to generalisation rules. That is, because in coarser scales such

a rectangle would encompass the building (e.g. a bounding box) and the vertices would not correspond across multiple LODs [741]. In 3D GIS, LOD1 and LOD2 are usually realised using the same footprints (see Chapter 4), hence this does not affect a lot of our work and while we do not contend that we have a solution here, it is certainly important to acknowledge the occasionally blurry distinction between representation and acquisition errors.

Finally, to maintain a reasonable level of simplicity, this study did not take into account potential systematic errors such as projection errors [990, 991], which may—together with the errors caused by ignoring terrain elevation—affect the building footprint dimensions [988].

12.5 CONCLUSIONS

Level of detail and positional accuracy are arguably some of the main aspects that should be stored in the metadata of most GIS datasets. In this chapter we performed a combined (multiple) error propagation analysis that demonstrates how much error is induced on top of error caused by using different LODs. Our main contribution in the subject of error propagation is that we take into account simultaneously multiple types of errors, and we consider multiple spatial analyses. While errors may be induced at many different points in a typical GIS process [974, 992], we deem that acquisition- and representation-induced errors are the most prominent ones, hence we focused on them.

The main conclusions of this chapter are: (1) the positional error is in many cases significantly more dominant than representation error; (2) as a result of this, in a lot of instances there is no need to go for a high representation level (LOD3) because the added value will vanish due to acquisition error; (3) the two considered errors are not additive; (4) error propagation is case-specific, hence there is no general conclusion that can be drawn for all spatial analyses; and (5) when disturbed with larger positional errors a lower representation may give better results in a spatial analysis than a higher representation disturbed with the error of the same magnitude.

This chapter also suggests that LOD in 3D GIS and scale in 2D GIS are related but different concepts, as discussed in Chapter 3. Scale in 2D is mainly associated with accuracy and precision, with less detail on small scale maps, while for 3D that relation does not always hold. In 3D data of coarse detail at a high precision/accuracy level is common, regardless of the spatial scale.

Plans for future work are to investigate the behaviour of the propagated error in other spatial analyses, and to investigate the behaviour of correlated errors as these may significantly impact the error propagation [959]. Correlations between positional errors can be incorporated in a probabilistic model but will require station-

arity assumptions to limit the number of model parameters [952]. Furthermore, we intend to consider other types of acquisition error common in geoinformation, such as completeness, which next to positional accuracy is commonly cited as a principal acquisition-induced error [993, 994]. Finally, a continuation of the work would be to consider the bigger picture of errors and analyse their consequences by associating them to a meaningful application. For example, the impact of a 10% error in the estimation of volume depends on the intended use of the derived information. If the volume was used to determine property tax, an error of 10% would not be acceptable. However, if it was used to estimate the volumetric building stock of a complete neighbourhood for heating demand estimation or for population estimation, on a large scale such an error might be less severe as the errors would probably cancel out in the sum (cf. the experiments in Section 7.6).

CHAPTER 13

Conclusions and future prospects

13.1 KEY FINDINGS AND CONTRIBUTIONS

The concept of level of detail is deeply ingrained in 3D geographical information science. It characterises and influences all aspects of a typical geo-data life cycle: from contracting and acquisition, to application and maintenance. Nevertheless, it has not been given as much attention as it deserves, and in this research I have sought to bridge that gap. In this thesis I have studied and developed several different aspects of LOD in 3D GIS, e.g. a theoretical framework formalising the concept, a consistent specification of LODs, the generation of data in different LODs, and the utilisation of models coupled with other factors such as positional uncertainty. The main research question is answered by subdiving it into multiple research questions. In this section I answer these 10 research questions, with references to the chapters and related contributions.

The developed work moves the field of 3D GIS forward in multiple ways. For example, the work presents the introduction of error propagation research in 3D GIS, and the domain of data capture may benefit from the improved understanding and enhanced specification of the LOD concept.

Q1: What is already known about the concept of LOD in 3D city modelling?

I have confirmed that there are different understandings of what the LOD is and what it encompasses. My thesis underlines that despite their shared name, the LOD concepts in GIS and computer graphics are different (Chapter 3). However, the LOD concept in 3D GIS has not been investigated to a great extent, and the available knowledge is fragmented across papers often with conflicting terminology and statements. For example, I have found that the LOD term is often used erroneously (e.g. interchangeably with accuracy and quality), and even fundamental topics such as discussing acquisition from the point of view of LOD have not been much considered (Chapter 2 bridges this gap by introducing a taxonomy of 3D acquisition techniques).

Q2: How can we formalise the concept of LOD in 3D city modelling?

Through my experience with practitioners and researchers (e.g. owing to my involvement in the OGC CityGML Standard Working Group, EuroSDR 3D Special Interest Group, Singapore Land Authority, and discussions with a myriad of stakeholders from several countries) I have found that users have disparate views on the LOD and different standards on expressing the content of data. A possible reason for this is that the LOD is a subjective topic. Unlike scale in cartography and resolution in rasters, the LODs in 3D GIS are not linear and their quantification is not rational:

there are no scale values such as 1:10k and 1:20k in maps, and their ordering is also sometimes ambiguous and relative (see Figure 3.1 on page 43 as a case).

In Chapter 3 I have presented a theory on the LOD concept and an LOD framework. The framework underlines that the LOD concept is composed of six components, indicating how thoroughly a dataset has been acquired. It is based on the examination of dozens of tenders, data specifications, and portfolios of companies, resulting in the consolidation of the different views of 3D GIS stakeholders. Researchers can now use this framework to express the content of 3D geographic information in a more precise manner.

Q3: How can we design a consistent and unambiguous LOD specification?

I have found that, even within the same standard (CityGML 2.0), practitioners have different understandings of particular LODs. Furthermore, the current CityGML LOD specification contains an insufficient number of LODs. Because of these reasons, and owing to its prominence, the OGC standard CityGML has been the focus of this research, and served as a base for designing a consistent and unambiguous LOD specification.

The concepts developed in Chapter 3 have been used to design a consistent and unambiguous LOD specification. The specification that I have presented in Chapter 4 improves the five CityGML standard LODs by defining them in a more precise manner and by refining them into multiple LODs solving the ambiguity pointed out by practitioners and researchers. The 16 LODs (Figure 4.1) are a recommendation for stakeholders for a firmer definition of their products and requirements, as they reflect capabilities of acquisition techniques and diminish ambiguity within each CityGML standard LOD. The proposed LOD specification is based on examining real-world datasets, tenders, standards, company documentation, and discussions. While the formalised approach presented in Chapter 3 is strict and unambiguous, the improved specification was achieved as a balance between rigidity and flexibility, in order to foster adoption. The enhanced specification can be easily understood by non-experts and it has backwards compatibility with the standard CityGML LODs since it extends them. Thanks to the improved elements of the LOD concept, it is expected that misunderstandings between stakeholders will be diminished. In fact, the specification has already been *de facto* adopted by a few national mapping agencies (e.g. Poland, Sweden, Singapore), proving the finding that there are ambiguities in the current CityGML LOD concept. Shortly after its publication, the specification and the discussed metrics have already been featured in papers by other researchers [498, 995].

Currently, discussions are ongoing for the improvement of the LOD concept in the upcoming version of CityGML (version 3.0), and over the course of this research

I was involved in the working group on the CityGML 3.0 LOD concept. My research has drawn attention to the following important topics, which are under consideration for becoming integrated in the standard: (1) the research has led into a better definition of existing LODs; (2) regarding geometric references: it should be possible to store multiple representations of the same LOD (e.g. it should be possible to store LOD1.2 with the H3 and H6 height references simultaneously, following the results of experiments in Chapter 10); (3) the future CityGML LOD concept should enable the definition of custom LODs as flexible profiles, leading to the possible integration of my LOD specification developed in Chapter 4; (4) the ambiguity of indoor LODs has led to the abolishment of LOD4, and to the introduction of multiple LODs of indoor data coupled with different LODs of the exterior; and (5) it should be possible to define semantics in lower LODs, e.g. doors in footprints (LOD0).

For more on this topic the reader is referred to the latest proposal of the working group [411][46]. Furthermore, while indoor has not been in focus of this thesis, some insights from the theoretical part of this thesis were used in related research to define a new indoor LOD called LOD2+, see [129][47].

Finally, the accuracy aspect of 3D city models has been an integral part of this research, and the insights presented in this thesis may be used to accompany an LOD specification. For example, Section 7.3.4 analyses the CityGML LOD1 quality requirements, and argues that they are not precisely described. This may possibly lead to inferior data being considered as valid, causing unreliable results in spatial analyses.

Q4: Are there multiple valid variants of the specified LODs?

In Chapter 5 I have found that there may be differences in how the LODs are realised (geometric references). Different acquisition practices may lead to substantially different 3D city models, despite the fact that they are of the same LOD. In fact, it was found that there is at least a hundred different outcomes (combinations) of 3D city models of the same feature class.

The concept of geometric references has not been investigated before to a great extent. Hence research was performed and a comprehensive list of references in 3D has been found, which may be a component in the specification of 3D city models and may be added as metadata in future versions of CityGML.

[46] Löwner MO, Gröger G, Benner J, Biljecki F, Nagel C (2016): Proposal for a new LOD and multi-representation concept for CityGML. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, IV-2/W1: 3–12. doi: 10.5194/isprs-annals-IV-2-W1-3-2016

[47] Boeters R, Arroyo Ohori K, Biljecki F, Zlatanova S (2015): Automatically enhancing CityGML LOD2 models with a corresponding indoor geometry. *International Journal of Geographical Information Science*, 29(12): 2248–2268. doi: 10.1080/13658816.2015.1072201

Paradoxically, the definition of references of the basic LOD1 have proven to be more difficult than that of detailed LOD3 due to many different approaches to model it. For example, the top surface of LOD1 block models may be drawn at several different places. It appears that the more detailed a model is, the less details should be given about its geometric references (e.g. for LOD3 the footprint and rooftop are always at their actual position). Furthermore, as another paradox, in Chapter 10 I have found that data at a coarser LOD may yield better results than finer data, owing to a favourable geometric reference that is used in the acquisition. This means that the GR may have a higher influence than the LOD in spatial analyses.

Q5: How can we realise the developed LOD specification?

The LOD specification proposed in Chapter 4 is accompanied with a discussion on acquisition techniques. In Chapter 6 I have presented an approach to realise the specification with procedural modelling. I have developed an experimental procedural modelling engine Random3Dcity, which is the first engine with native CityGML support and the first publicly available CityGML data source that contains an extensive number of features modelled in multiple LODs and multiple geometric references. The procedural modelling engine, as is other the code behind this thesis (e.g. Solar3Dcity), was released as open source. Furthermore, sample datasets produced by the developed procedural engine have been released as open data, and have been been used by other researchers (Section 6.4).

I have found procedural modelling to be an excellent tool in research such as this. It has been the technique of choice to realise the developed specification, and it has been found essential in simulating acquisition errors. In fact, procedural modelling was the only feasible method for realising the LOD specification and for comparing the performance of different LODs in different spatial analyses. Among others, procedural modelling has the following advantages in 3D GIS: it is free and straightforward in generating a large amount of data according to predefined rules, and generating 'sterile' data in multiple LODs including highly detailed representations free of errors (e.g. positional and topological errors).

Q6: How can we increase the LOD of existing data?

I have found that augmentation of the LOD, as opposed to generalisation, is a narrow topic of research that requires further attention (see the overview in Section 2.2).

In Chapter 7 I have presented a new method for generating 3D city models. I have demonstrated that it is possible to predict the heights of buildings from non-elevation data, with a reasonable level of accuracy and to extrude building footprints obtaining volumetric 3D city models. This method, based on machine learning,

is intended for regions that lack elevation data, but are rich in building footprints. Using the method that I have developed the attributes in existing 2D datasets may be leveraged for creating 3D city models in LOD1 and using them for several analyses in which 2D is not sufficient.

Q7: How can we integrate multiple LODs of the same feature?

Datasets may come in multiple representations. In fact, in the future I foresee the increase of data delivered in multiple representations, motivated by reasons listed in Chapter 8. On the multi-LOD side, I have developed techniques to efficiently link such data, supporting future developments in multi-LOD aspects by facilitating their storage, maintenance, and update. Two techniques have been presented: one that detects corresponding geometries across multiple LODs and links them, and one that (theoretically) establishes links between LODs in a higher dimensional (4D) construction.

Q8: How does the LOD influence the quality of spatial analyses?

In the experimental portion of the work (Chapters 9 and 10) I have evaluated the performance of different LODs when employed in a spatial analysis. The impact of the LOD in spatial analyses comes with mixed conclusions. On one hand, some applications (e.g. visualisation) may strongly benefit from more detail, and some of the applications require a certain minimum LOD (e.g. estimation of the solar potential of roofs is not advised with LOD1 block models). On the other hand, experiments have indicated that the benefit provided by the added detail may be minuscule. The cost of producing a finer LOD may be much higher than the benefit they bring, but that ultimately depends on the application. For example, acquiring roof superstructures (LOD2.2) in comparison to basic roofs (LOD2.1) requires laborious photogrammetric mapping or advanced reconstruction techniques. Hence it was still important to refine these significant differences within the same CityGML 2.0 LOD, despite the subtle differences they bring in a spatial analysis.

Ultimately the conclusion is application-dependent, hence experiments have to be run for each application. This is actually good news for 3D generalisation: a dataset can be considerably simplified while not jeopardising the benefit it brings to an application.

Another finding here is that even though the application differences between volumetric LODs (e.g. LOD2.2 vs LOD2.3) may be subtle, the research has also revealed that a simple crude 3D model has many advantages over 2D geo-datasets (i.e. LOD1 vs LOD0). For example, the estimation of population (Chapter 9) is significantly more accurate when considering the vertical extent of buildings. This rea-

soning is also obvious from the fact that some applications are not possible with 2D, e.g. shadow estimation (Chapter 10 and Figure 1.2). Hence the jump from LOD0 to LOD1 is drastic, but the one from LOD1 to LOD2 (and within the standard CityGML 2.0 LODs) is less tangible, a fact that was also discussed in Chapter 12.

Here the contribution of my research is that it does such an analysis in a scientifically correct manner. Related work integrated multi-LOD data with different provenance, e.g. having different levels of accuracies. According to the experiments in Chapter 11 and in Chapter 12, different levels of accuracies have a large impact on the results of a spatial analyses, hence such analyses are severely biased. In my research I rely on procedural models, being generated from the same source and being devoid of errors. The differences between LODs have been evaluated in such a 'sterile' environment, providing unbiased results. I recommend for researchers to follow a similar approach when investigating the impact of the LOD on a spatial analysis.

A side finding of the research on the influence of the LOD on the quality of spatial analyses is that many applications of 3D city models are insufficiently documented, as research papers usually superficially describe the methodology and technical details behind the analysis, thus impeding replication, and making implementation scarce. Besides the comprehensive literature review on the application of 3D city models (Section 2.3), in this thesis a few use cases have been focused upon and their methodology is described in fine detail (population estimation in Section 9, shadow in Section 10, and solar in Section 11). Furthermore, a byproduct of this research is that the use case of population estimation is advanced: the research described in Chapter 9 presents the most comprehensive population estimation study carried out to date, where 3D buildings have been used, and it adds to the body of knowledge by comparing different approaches employed at different administrative levels.

Q9: How do acquisition errors influence LODs?

The thesis raises awareness that LOD is correlated with accuracy and quality, but they are distinct concepts that should not be mixed. Models at finer LODs tend to be more accurate, but one needs to be careful with this presumption as this thesis has indicated that the opposite case is common, and given the choice, it may mislead stakeholders into choosing a more detailed but less accurate dataset (see Chapter 12 for the discussion). Furthermore, sometimes associating the concept of accuracy with the LOD is meaningless because of features not existing in reality (e.g. design models, procedural models), which are not a result of reconstructing the real-world. During the research it was discovered that positional errors may be injected into the procedural workflow to simulate real-world errors and to generate models intentionally burdened with positional errors, an inescapable fact during the acquisition

of real-world data. The generated models were then used to investigate the propagation of uncertainty, by comparing the disturbed datasets with the error-free original. Procedural modelling is increasingly being used in GIS, but errors have never been within the scope of such a research. This research presents a contribution in understanding the propagation of error in 3D GIS and in leveraging procedural modelling. Finally, a contribution of my work is that the Monte Carlo method was employed for the first time in 3D GIS.

Q10: How can we distinguish errors induced by LOD specification and errors induced by the acquisition?

Distinguishing these two components is not possible with real-world data. In this research it is done by simulating the data with the desired characteristics. Procedural modelling offers the possibility to tune the positional error and the LOD we require in the analysis.

Another interesting finding that I have reached is that the quality of the realisation of the models (positional accuracy) may have a higher influence than the LOD. Chapter 12 proves that the benefit, that a finer LOD brings over a coarser LOD, virtually vanishes with positional errors. Such a conclusion may 'temper' practitioners and researchers who are usually hungry for data of finer detail, and rather motivate them to focus on the quality of the data and on the accuracy of the empirical models (described in Section 12.2.3).

13.2 FUTURE PROSPECTS

Trends in remote sensing and GIS, especially within the advancements of acquisition technologies, indicate the increase in the availability of 3D city models and their LODs. An example of a particularly interesting recent development is the use of spaceborne radar to derive 3D city models [996]. Advancements such as this one open new possibilities for future research.

There is room for additional work in each of the topics covered in the research, hence each of the chapters of this thesis contains ideas for future work. In this section a general roadmap for future work is presented. Most importantly, the thesis has developed transferable knowledge that can be applied in other domains and problems, such as dynamic geoinformation, 2D GIS, and point clouds, where there is room for investigating corresponding concepts.

13.2.1 Dynamic aspects

This thesis does not cover dynamic phenomena in 3D city models. Examples involve moving objects (e.g. elevators in indoor models), deformable objects and objects changing shape (e.g. waterbodies in a flood, and wall surfaces in blast simulations [997]), and features changing appearance (e.g. urban vegetation during different seasons).

Dynamic features in 3D city models are not yet found a lot in practice, but they are becoming increasingly researched, e.g. with the introduction of *dynamizers* by Chaturvedi and Kolbe [998]. The LOD concept is important in this topic, and the developed ideas (e.g. the framework presented in Chapter 3) are applicable to the dynamic aspects. In future work it would be beneficial to integrate the dynamic aspect as another metric in the developed framework. Furthermore, it would be beneficial to investigate the propagation of uncertainty in dynamic aspects, similar to the approach in Chapter 11.

13.2.2 LOD in the context of quality

Quality of 3D city models is becoming an increasingly important topic owing to the growing number of applications and users. With the exception of a few brief papers [999], work on this topic is sparse and it is mostly focused on the positional aspect [1000] and topology [60, 633].

Regarding the LOD concept in the context of quality is quite important, and my research presents a foundation for future work in this direction.

First, it is important to ensure that the LOD of the dataset corresponds to the one that was intended to be acquired or the one that is stated in the metadata (e.g. whether a dataset has been acquired according to the LOD2.2 requirements stated in Chapter 4). It is important to develop mechanism to check and express the quality of the correspondence of each of the metrics, between metadata (specification) and the realised data. For example, in change detection, which is becoming a topical subject owing to the increased frequency of surveys [1001, 1002], it is important to ensure that the updated model is acquired according to the correct LOD because details that are not present in the base dataset might be erroneously flagged as changed. Recent work indicates developments in the direction of quality control of some 3D aspects, e.g. there is work on checking the quality of textures [1003].

Second, the dataset may contain inconsistencies at the building level as some features may be mapped in lesser detail than others, resulting in an inconsistent LOD across a dataset. This stems not only from volunteered geoinformation (e.g. see Figure 12.1 as a case of heterogeneous LOD), but also in cases with the automatic reconstruction of roofs. In some approaches if the algorithm fails to reconstruct an

LOD2 building, it 'falls down' to LOD1 without noting this in the metadata.

Third, another quality aspect in this context is the integration of multi-LOD datasets: checking if the geometry has been re-used and if it is consistent (Chapter 8), minimising efforts for updating and maintenance and maximising storage efficiency. Some of these topics entail developing new quality indicators suited for 3D models, possibly extending present geo-data quality control standards, which have traditionally been confined to 2D.

Fourth, it is important to set quality requirements for each refined LOD. As briefly discussed in Section 7.3.4, CityGML 2.0 contains recommendations for the positional accuracy of each LOD. These values are occasionally found in literature [1004], but appear to be arbitrary and generous. It would be beneficial to define what makes an LOD valid from the point of view of acquisition quality (e.g. positional accuracy).

Finally, in the context of quality it is important to note that the LOD indirectly influences the topological and other aspects of the dataset. My recent paper on validation of publicly available CityGML data indicates that with the increase of LOD the likelihood of geometrical and topological errors increases [566][48].

13.2.3 Indoor LODs and integration of GIS and BIM

Recently there has been a surge in the applications of indoor geoinformation [1005, 1006], and research on their acquisition has intensified [1007–1010]. Indoor data have applications on their own (e.g. indoor navigation, daylight autonomy, and estimating glare [6, 925]), but they may also augment other applications with additional insight. For example, using 3D city models with indoor data it is possible to estimate which apartments will be affected by a flood [1011, 1012] (see Figure 2.4a).

My research did not investigate deeply into indoor data because 3D city models containing indoor data have so far been mostly theoretical and limited to instances with only a few buildings. Recent developments in automatic reconstruction of a basic interior from the building morphology and other characteristics [129, 130] hints at their increased availability in the near future. For these reasons, research on LODs for 3D indoor data is necessary, and my thesis could assist related efforts by using some of the developed concepts. For example, an LOD specification of indoor LODs may be introduced following the rationale presented in Chapter 4, and procedurally generated indoor data may be used to carry out experiments similar to the ones presented in Chapter 10.

[48] Biljecki F, Ledoux H, Du X, Stoter J, Soon KH, Khoo VHS (2016): The most common geometric and semantic errors in CityGML datasets. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, IV-2/W1: 13–22. doi: 10.5194/isprs-annals-IV-2-W1-13-2016

13.2.4 Cost and context-awareness

The concept of cost in computer graphics is significantly different from the one in GIS (see the discussion in Section 3.3). In GIS the cost is primarily the burden of acquisition and the budget. Because these are hard to quantify and are subject to temporal and geographical variations (e.g. costs go down due to automation), they have not been the subject of this work.

However, the topic of cost may be considered in the context of the efficiency of spatial analyses that are repeatedly carried out. The goal would be to analyse the costs of each LOD for a particular analysis and compare it to the benefit of each LOD, to enable dynamic switching to the optimal LOD. For example, if it is known that LOD1 brings slightly less accurate results than LOD2 when calculating the volume of buildings (as demonstrated in Chapter 10), but at a significantly faster turnaround, then the software could automatically switch to the LOD1 representation if both representation were available. The multi-LOD linking theory presented in Chapter 8 and the experiments performed in Chapter 10 may aid future research on this topic.

Furthermore, developments in 3D generalisation and the work presented in this thesis can open a new line of research: context-aware 3D generalisation driven by the intended use of the data. Such a research could aim at simplifying data to the LOD where a great deal of complexity is removed, while maintaining a comparable accuracy for a particular spatial analysis, akin to the process of LOD and simplification in computer graphics. In this context it is relevant to note the ongoing research of Besuievsky et al. [1013] as a related example. The researchers work towards generalising 3D city models to reduce the computational load when estimating the sky view factor, while preserving the accuracy of the spatial analysis. The simplification algorithm is customised, and it is driven by this particular spatial analysis. For example, it is aimed at keeping only the geometry that has the largest impact on the final calculation.

13.2.5 3D + LOD = 4D

Section 8.3 discussed the higher dimensional integration of 3D geometry and LOD. In relation to this idea, it would be beneficial to explore further the possibilities in their integration. Such an integration, coupled with the development of slicing mechanism that are context-aware (Section 13.2.4), may lead to the dynamic derivation of application-specific LODs.

13.2.6 Extended thematic classes

The production of 3D city models and this research have been focused on buildings and terrain, as the most prominent features of the urban environment. However,

national mapping agencies have recently been producing 3D city models focused on all thematic features. This is primarily due to three drivers:

1. Advancement in 3D acquisition techniques, such as developments in remote sensing dedicated to acquiring a specific thematic feature, e.g. power lines [1014], traffic signs [1015–1017], traffic lights [1018], roads [1019, 1020], railway [1021], and solitary vegetation features [25, 1022, 1023]. There is potential work here for defining the different LODs for other thematic classes.

2. Increase in spatial analyses that require non-building thematic features, e.g. vegetation for analysing the cooling effect of 3D vegetation [1024].

3. Work in the standardisation of data, e.g. developing an extension to CityGML to support underground and utilities [1025].

All the papers referred to in this section have been published within the past two years, indicating lively research on these topics. However, besides this research, the only other research on LODs of urban features is to define the LODs of terrain [1026], meaning that other thematic features have been neglected. While other thematic features such as roads do not allow for as much refinement as buildings, the basic definition of the five CityGML LODs does not foresee extensively other thematic classes (e.g. it is not clear what the LOD3 of roads is). The insights that I have presented in this thesis may support research into this direction. For example, besides developing the LOD specification (Chapter 4), geometric references (Chapter 5) for different thematic features may also be investigated.

BIBLIOGRAPHY

[1] Biljecki F, Ledoux H, Stoter J (2014): Redefining the Level of Detail for 3D models. *GIM International*, 28(11): 21–23.

[2] Gruber M, Pasko M, Leberl F (1995): Geometric versus Texture Detail in 3-D Models of Real World Buildings. *Automatic Extraction of Man-Made Objects from Aerial and Space Images*, Birkhäuser, Basel, pp. 189–198.

[3] Li Y, Ding D, Liu C, Wang C (2016): A pixel-based approach to estimation of solar energy potential on building roofs. *Energy and Buildings*, 129: 563–573.

[4] Benner J, Geiger A, Haefele KH, Isele J (2010): Interoperability of Geothermal Data Models. *Proceedings World Geothermal Congress*. Bali, Indonesia, pp. 1–9.

[5] Geiger A, Benner J, Haefele KH (2015): Generalization of 3D IFC Building Models. *3D Geoinformation Science*, Springer International Publishing, Dubai, UAE, pp. 19–35.

[6] Chen KW, Norford L (2016): Workflow for Generating 3D Urban Models from Open City Data for Performance-Based Urban Design. *Proceedings of the Asim 2016 IBPSA Asia Conference*. Jeju, Korea.

[7] Teo TA, Cho KH (2016): BIM-oriented indoor network model for indoor and outdoor combined route planning. *Advanced Engineering Informatics*, 30(3): 268–282.

[8] Goodchild MF (2011): Scale in GIS: An overview. *Geomorphology*, 130(1-2): 5–9.

[9] He S (2012): *Production and Visualization of Levels of Detail for 3D City Models*. Ph.D. thesis, École Centrale de Nantes, Nantes, France.

[10] OGC (2012): OGC City Geography Markup Language (CityGML) Encoding Standard 2.0.0.

[11] Billen R, Cutting-Decelle AF, Marina O, de Almeida JP, M C, Falquet G, Leduc T, Métral C, Moreau G, Perret J, Rabin G, San José R, Yatskiv I, Zlatanova S (2014): 3D City Models and urban information: Current issues and perspectives. *3D City Models and urban information: Current issues and perspectives – European COST Action TU0801*. EDP Sciences, Les Ulis, France, pp. 1–118.

[12] Zhu Q, Hu M, Zhang Y, Du Z (2009): Research and practice in three-dimensional city modeling. *Geo-spatial Information Science*, 12(1): 18–24.

[13] Döllner J, Baumann K (2006): Virtual 3D City Models as Foundation of Complex Urban Information Spaces. M Schrenk, ed., *11th international conference on Urban Planning and Spatial Development in the Information Society (REAL CORP)*. CORP – Competence Center of Urban and Regional Planning, pp. 107–112.

[14] Lancelle M, Fellner DW (2004): Current issues on 3D city models. *Proceedings of the 25th International Conference in Image and Vision Computing*. Akaroa, New Zealand, pp. 363–369.

[15] Nebiker S, Bleisch S, Christen M (2010): Rich point clouds in virtual globes – A new paradigm in city modeling? *Computers, Environment and Urban Systems*, 34(6): 508–517.

[16] Wallerman J, Bohlin J, Fransson JE, Lundberg K (2009): Forest data capture using optical 3D digital surface models from the C3 Technologies system. *ASPRS/MAPPS 2009 Specialty Conference, Digital mapping–From Elevation to Information*. San Antonio, Texas, United States, pp. 16–19.

[17] Gröger G, Plümer L (2011): Topology of surfaces modelling bridges and tunnels in 3D-GIS. *Computers, Environment and Urban Systems*, 35(3): 208–216.

[18] Jokela J (2016): *CityGML building model production from airborne laser scanning*. Master's thesis, Aalto University, Aalto, Finland.

[19] Jazayeri I, Rajabifard A, Kalantari M (2014): A geometric and semantic evaluation of 3D data sourcing methods for land and property information. *Land Use Policy*, 36: 219–230.

[20] Jazayeri I, Rajabifard A, Kalantari M (2014): 3D Data Sourcing for Land and Property Information. *GIM International*, 28(12): 25–27.

[21] Toth C, Jóźków G (2016): Remote sensing platforms and sensors: A survey. *ISPRS Journal of Photogrammetry and Remote Sensing*, 115: 22–36.

[22] Tack F, Buyuksalih G, Goossens R (2012): 3D building reconstruction based on given ground plan information and surface models extracted from spaceborne imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 67: 52–64.

[23] Colomina I, Molina P (2014): Unmanned aerial systems for photogrammetry and remote sensing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 92: 79–97.

[24] Nex F, Remondino F (2013): UAV for 3D mapping applications: a review. *Applied Geomatics*, 6(1): 1–15.

[25] Böhm J, Brédif M, Gierlinger T, Krämer M, Lindenberg R, Liu K, Michel F, Sirmacek B (2016): The IQmulus urban showcase: automatic tree classification and identification in huge mobile mapping point clouds. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLI-B3: 301–307.

[26] Kaartinen H, Hyyppä J, Kukko A, Jaakkola A, Hyyppä H (2012): Benchmarking the Performance of Mobile Laser Scanning Systems Using a Permanent Test Field. *Sensors*, 12(12): 12814–12835.

[27] Früh C, Zakhor A (2004): An Automated Method for Large-Scale, Ground-Based City Model Acquisition. *International Journal of Computer Vision*, 60(1): 5–24.

[28] Rosser J, Morley J, Smith G (2015): Modelling of Building Interiors with Mobile Phone Sensor Data. *ISPRS International Journal of Geo-Information*, 4(2): 989–1012.

[29] Sirmacek B, Lindenbergh R (2014): Accuracy assessment of building point clouds automatically generated from iphone images. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XL-5: 547–552.

[30] Hartmann W, Havlena M, Schindler K (2016): Towards complete, geo-referenced 3D models from crowd-sourced amateur images. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, III-3: 51–58.

[31] Stilla U, Soergel U, Thoennessen U (2003): Potential and limits of InSAR data for building reconstruction in built-up areas. *ISPRS Journal of Photogrammetry and Remote Sensing*, 58(1-2): 113–123.

[32] Shahzad M, Zhu XX (2015): Robust Reconstruction of Building Facades for Large Areas Using Spaceborne TomoSAR Point Clouds. *IEEE Transactions on Geoscience and Remote Sensing*, 53(2): 752–769.

[33] Fischer A, Kolbe TH, Lang F, Cremers AB, Förstner W, Plümer L, Steinhage V (1998): Extracting Buildings from Aerial Images Using Hierarchical Aggregation in 2D and 3D. *Computer Vision and Image Understanding*, 72(2): 185–203.

[34] Heo J, Jeong S, Park HK, Jung J, Han S, Hong S, Sohn HG (2013): Productive high-complexity 3D city modeling with point clouds collected from terrestrial LiDAR. *Computers, Environment and Urban Systems*, 41: 26–38.

[35] Pal Singh S, Jain K, Mandla VR (2013): Virtual 3D Campus Modeling by Using Close Range Photogrammetry. *American Journal of Civil Engineering and Architecture*, 1(6): 200–205.

[36] Agarwal S, Furukawa Y, Snavely N, Simon I (2011): Building rome in a day. *Communications of the ACM*, 54(10): 105–112.

[37] Somogyi A, Barsi A, Molnar B, Lovas T (2016): Crowdsourcing based 3D modeling. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLI-B5: 587–590.

[38] Duan L, Lafarge F (2016): Towards Large-Scale City Reconstruction from Satellites. *Spatial Information Theory. Cognitive and Computational Foundations of Geographic Information Science*, Springer International Publishing, Cham, pp. 89–104.

[39] Xiao J, Gerke M, Vosselman G (2012): Building extraction from oblique airborne imagery based on robust façade detection. *ISPRS Journal of Photogrammetry and Remote Sensing*, 68: 56–68.

[40] Kedzierski M, Fryskowska A (2014): Terrestrial and Aerial Laser Scanning Data Integration Using Wavelet Analysis for the Purpose of 3D Building Modeling. *Sensors*, 14(7): 12070–12092.

[41] Fruh C, Zakhor A (2003): Constructing 3D city models by merging aerial and ground views. *IEEE Computer Graphics and Applications*, 23(6): 52–61.

[42] Idris R, Latif ZA, Hamid JRA, Jaafar J, Ahmad MY (2014): Integrating airborne LiDAR dataset and photographic images towards the construction of 3D building model. *IOP Conference Series: Earth and Environmental Science. Proceedings of the 8th International Symposium of the Digital Earth (ISDE8)*, 18: 012049.

[43] Demir N, Baltsavias EP (2012): Automated modeling of 3D building roofs using image and LiDAR data. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, I-4: 35–40.

[44] Suveg I, Vosselman G (2004): Reconstruction of 3D building models from aerial images and maps. *ISPRS Journal of Photogrammetry and Remote Sensing*, 58(3-4): 202–224.

[45] Gimenez L, Robert S, Suard F, Zreik K (2016): Automatic reconstruction of 3D building models from scanned 2D floor plans. *Automation in Construction*, 63: 48–56.

[46] Lewis R, Séquin C (1998): Generation of 3D building models from 2D architectural plans. *Computer-Aided Design*, 30(10): 765–779.

[47] Gimenez L, Hippolyte JL, Robert S, Suard F, Zreik K (2015): Review: reconstruction of 3D building information models from 2D scanned plans. *Journal of Building Engineering*, 2: 24–35.

[48] Huang HC, Lo SM, Zhi GS, Yuen RKK (2008): Graph theory-based approach for automatic recognition of CAD data. *Engineering Applications of Artificial Intelligence*, 21(7): 1073–1079.

[49] Müller P, Wonka P, Haegler S, Ulmer A, van Gool L (2006): Procedural modeling of buildings. *ACM Transactions on Graphics*, 25(3): 614–623.

[50] Peeters A (2016): A GIS-based method for modeling urban-climate parameters using automated recognition of shadows cast by buildings. *Computers, Environment and Urban Systems*, 59: 107–115.

[51] Jiang-An W, Gang C, Pan L, Tao G (2016): 3D Modeling Of GNSS Vector Tracking Receiver in Urban Area. *International Journal of Hybrid Information Technology*, 9(7): 71–78.

[52] Häkkilä J, Posti M, Koskenranta O, Ventä-Olkkonen L (2013): Co-creating a digital 3D city with children. M Kranz, K Synnes, eds., *Proceedings of the 12th International Conference on Mobile and Ubiquitous Multimedia (MUM '13)*. ACM Press, Luleå, Sweden, pp. 1–2.

[53] Mao B, Harrie L (2016): Methodology for the Efficient Progressive Distribution and Visualization of 3D Building Objects. *ISPRS International Journal of Geo-Information*, 5(10): 185–23.

[54] Guercke R, Götzelmann T, Brenner C, Sester M (2011): Aggregation of LoD 1 building models as an optimization problem. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66(2): 209–222.

[55] Janečka K, Váša L (2017): Compression of 3D Geographical Objects at Various Level of Detail. *The Rise of Big Spatial Data*, Springer International Publishing, Cham, pp. 359–372.

[56] Corsini M, Barni M, Bartolini F, Caldelli R, Cappelini V, Piva A (2003): Towards 3D watermarking technology. *IEEE Region 8 EUROCON 2003. Computer as a Tool.* pp. 393–396.

[57] Uccheddu F, Corsini M, Barni M (2004): Wavelet-based blind watermarking of 3D models. *Proceedings of the 2004 workshop on Multimedia and security*. ACM Press, New York City, New York, USA, pp. 143–154.

[58] Wang K, Lavoué G, Denis F, Baskurt A (2008): A Comprehensive Survey on Three-Dimensional Mesh Watermarking. *IEEE Transactions on Multimedia*, 10(8): 1513–1527.

[59] Zhao J, Stoter J, Ledoux H, Zhu Q (2012): Repair and generalization of hand-made 3D building models. *Proceedings of the 15th ICA Workshop on Generalisation and Multiple Representation*. Istanbul, Turkey, pp. 1–10.

[60] Alam N, Wagner D, Wewetzer M, von Falkenhausen J, Coors V, Pries M (2014): Towards Automatic Validation and Healing of CityGML Models for Geometric and Semantic Consistency. *Innovations in 3D Geo-Information Sciences*, Springer International Publishing, pp. 77–91.

[61] Filippovska Y, Wichmann A, Kada M (2016): Space partitioning for privacy enabled 3D city models. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLII-2/W2: 17–22.

[62] Daly C (2006): Guidelines for assessing the suitability of spatial climate data sets. *International Journal of Climatology*, 26(6): 707–721.

[63] Ledoux H, Meijers M (2011): Topologically consistent 3D city models obtained by extrusion. *International Journal of Geographical Information Science*, 25(4): 557–574.

[64] Arroyo Ohori K, Ledoux H, Stoter J (2015): A dimension-independent extrusion algorithm using generalised maps. *International Journal of Geographical Information Science*, 29(7): 1166–1186.

[65] Prechtel N (2014): On strategies and automation in upgrading 2D to 3D landscape representations. *Cartography and Geographic Information Science*, 42(3): 244–258.

[66] Haala N, Brenner C, Karl-Heinrich A (1997): Generation of 3D city models from digital surface models and 2D GIS. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XXXII-3-4W2: 68–76.

[67] Aringer K, Roschlaub R (2014): Bavarian 3D Building Model and Update Concept Based on LiDAR, Image Matching and Cadastre Information. *Innovations in 3D Geo-Information Sciences*, Springer International Publishing, pp. 143–157.

[68] Thiele A, Wegner JD, Soergel U (2010): Building Reconstruction from Multi-aspect InSAR Data. U Soergel, ed., *Remote Sensing and Digital Image Processing*, Springer Netherlands, pp. 187–214.

[69] Janečka K, Kutal M (2014): The 3D model of the Small Fortress of the Terezin Memorial. *14th International Multidisciplinary Scientific GeoConference SGEM*. Sofia, Bulgaria, pp. 925–932.

[70] Liang J, Shen S, Gong J, Liu J, Zhang J (2017): Embedding user-generated content into oblique airborne photogrammetry-based 3D city model. *International Journal of Geographical Information Science*, 31(1): 1–16.

[71] Müller Arisona S, Zhong C, Huang X, Qin H (2013): Increasing detail of 3D models through combined photogrammetric and procedural modelling. *Geo-spatial Information Science*, 16(1): 45–53.

[72] Sugihara K, Murase T, Zhou X (2015): Automatic generation of 3D building models from building polygons on digital maps. *International Conference on 3D Imaging (IC3D)*.

[73] Sugihara K, Shen Z (2016): Automatic generation of 3D house models with solar photovoltaic generation for smart city. *3rd MEC International Conference on Big Data and Smart City (ICBDSC)*. Muscat, Oman, pp. 1–7.

[74] Edelsbrunner J, Krispel U, Havemann S, Sourin A, Fellner DW (2016): Constructive Roofs from Solid Building Primitives. *Transactions on Computational Science XXVI*, Springer, Berlin, Heidelberg, pp. 17–40.

[75] Ben-Joseph E, Ishii H, Underkoffler J, Piper B, Yeung L (2001): Urban simulation and the luminous planning table: bridging the gap between the digital and the tangible. *Journal of Planning Education and Research*, 21(2): 196–203.

[76] Biljecki F, Stoter J, Ledoux H, Zlatanova S, Çöltekin A (2015): Applications of 3D City Models: State of the Art Review. *ISPRS International Journal of Geo-Information*, 4(4): 2842–2889.

[77] Jacobson I (1992): *Object-oriented software engineering: a use case driven approach*. Addison-Wesley, Reading, Massachusetts.

[78] Kaden R, Kolbe TH (2014): Simulation-Based Total Energy Demand Estimation of Buildings using Semantic 3D City Models. *International Journal of 3-D Information Modeling*, 3(2): 35–53.

[79] Lu Z, Im J, Quackenbush L (2011): A Volumetric Approach to Population Estimation Using Lidar Remote Sensing. *Photogrammetric Engineering and Remote Sensing*, 77(11): 1145–1156.

[80] Kumar L, Skidmore AK, Knowles E (1997): Modelling topographic variation in solar radiation in a GIS environment. *International Journal of Geographical Information Science*, 11(5): 475–497.

[81] Fath K, Stengel J, Sprenger W, Wilson HR, Schultmann F, Kuhn TE (2015): A method for predicting the economic potential of (building-integrated) photovoltaics in urban areas based on hourly Radiance simulations. *Solar Energy*, 116: 357–370.

[82] Redweik P, Catita C, Brito M (2013): Solar energy potential on roofs and facades in an urban landscape. *Solar Energy*, 97: 332–341.

[83] Eicker U, Nouvel R, Duminil E, Coors V (2014): Assessing Passive and Active Solar Energy Resources in Cities Using 3D City Models. *Energy Procedia*, 57: 896–905.

[84] Šúri M, Hofierka J (2004): A New GIS-based Solar Radiation Model and Its Application to Photovoltaic Assessments. *Transactions in GIS*, 8(2): 175–190.

[85] Santos T, Gomes N, Freire S, Brito MC, Santos L, Tenedório JA (2014): Applications of solar mapping in the urban environment. *Applied Geography*, 51: 48–57.

[86] Jakubiec JA, Reinhart CF (2013): A method for predicting city-wide electricity gains from photovoltaic panels based on LiDAR and GIS data combined with hourly Daysim simulations. *Solar Energy*, 93: 127–143.

[87] Szabó S, Enyedi P, Horváth M, Kovács Z, Burai P, Csoknyai T, Szabó G (2016): Automated registration of potential locations for solar energy production with Light Detection And Ranging (LiDAR) and small format photogrammetry. *Journal of Cleaner Production*, 112: 3820–3829.

[88] Wiginton LK, Nguyen HT, Pearce JM (2010): Quantifying rooftop solar photovoltaic potential for regional renewable energy policy. *Computers, Environment and Urban Systems*, 34(4): 345–357.

[89] Peronato G, Rey E, Andersen M (2015): Sampling of building surfaces towards an early assessment of BIPV potential in urban contexts. *31st International PLEA Conference*. Bologna, Italy.

[90] Strzalka A, Alam N, Duminil E, Coors V, Eicker U (2012): Large scale integration of photovoltaics in cities. *Applied Energy*, 93: 413–421.

[91] Redweik P, Catita C, Brito MC (2011): 3D local scale solar radiation model based on urban LiDAR data. *ISPRS Workshop High-Resolution Earth Imaging for Geospatial Information*. Hannover, Germany, pp. 1–5.

[92] Li Z, Zhang Z, Davey K (2015): Estimating Geographical PV Potential Using LiDAR Data for Buildings in Downtown San Francisco. *Transactions in GIS*, 19(6): 930–963.

[93] Lukač N, Žlaus D, Seme S, Žalik B, Štumberger G (2013): Rating of roofs' surfaces regarding their solar potential and suitability for PV systems, based on LiDAR data. *Applied Energy*, 102: 803–812.

[94] Catita C, Redweik P, Pereira J, Brito MC (2014): Extending solar potential analysis in buildings to vertical facades. *Computers & Geosciences*, 66: 1–12.

[95] Liang J, Gong J, Li W, Ibrahim AN (2014): A visualization-oriented 3D method for efficient computation of urban solar radiation based on 3D–2D surface mapping. *International Journal of Geographical Information Science*, 28(4): 780–798.

[96] Hofierka J, Zlocha M (2012): A New 3-D Solar Radiation Model for 3-D City Models. *Transactions in GIS*, 16(5): 681–690.

[97] Hofierka J, Kaňuk J (2009): Assessment of photovoltaic potential in urban areas using open-source solar radiation tools. *Renewable Energy*, 34(10): 2206–2214.

[98] Liang J, Gong J, Zhou J, Zhou J, Ibrahim AN, Li M, Li M (2015): An open-source 3D solar radiation model integrated with a 3D Geographic Information System. *Environmental Modelling & Software*, 64: 94–101.

[99] Carneiro C, Golay F (2009): Solar Radiation over the Urban Texture: LIDAR Data and Image Processing Techniques for Environmental Analysis at City Scale. J Lee, S Zlatanova, eds., *3D Geo-Information Sciences*, Springer Berlin Heidelberg, pp. 319–340.

[100] Jochem A, Höfle B, Rutzinger M, Pfeifer N (2009): Automatic Roof Plane Detection and Analysis in Airborne Lidar Point Clouds for Solar Potential Assessment. *Sensors*, 9(7): 5241–5262.

[101] Gooding J, Crook R, Tomlin AS (2015): Modelling of roof geometries from low-resolution LiDAR data for city-scale solar energy applications using a neighbouring buildings method. *Applied Energy*, 148: 93–104.

[102] Yu B, Liu H, Wu J, Lin WM (2009): Investigating impacts of urban morphology on spatio-temporal variations of solar radiation with airborne LIDAR data and a solar flux model: a case study of downtown Houston. *International Journal of Remote Sensing*, 30(17): 4359–4385.

[103] Buffat R (2016): Feature-Aware Surface Interpolation of Rooftops Using Low-Density Lidar Data for Photovoltaic Applications. *Geospatial Data in a Changing World*, Springer International Publishing, Cham, pp. 337–350.

[104] Freitas S, Catita C, Redweik P, Brito MC (2015): Modelling solar potential in the urban environment: State-of-the-art review. *Renewable and Sustainable Energy Reviews*, 41: 915–931.

[105] Chwieduk DA (2009): Recommendation on modelling of solar energy incident on a building envelope. *Renewable Energy*, 34(3): 736–741.

[106] Nichol J, Wong MS (2005): Modeling urban environmental quality in a tropical city. *Landscape and Urban Planning*, 73(1): 49–58.

[107] Vermeulen T, Knopf-Lenoir C, Villon P, Beckers B (2015): Urban layout optimization framework to maximize direct solar irradiation. *Computers, Environment and Urban Systems*, 51: 1–12.

[108] Aarsen R, Janssen M, Ramkisoen M, Biljecki F, Quak W, Verbree E (2015): Installed base registration of decentralised solar panels with applications in crisis management. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XL-3/W3: 219–223.

[109] Helbich M, Jochem A, Mücke W, Höfle B (2013): Boosting the predictive accuracy of urban hedonic house price models through airborne laser scanning. *Computers, Environment and Urban Systems*, 39: 81–92.

[110] Saran S, Wate P, Srivastav SK, Krishna Murthy YVN (2015): CityGML at semantic level for urban energy conservation strategies. *Annals of GIS*, 21(1): 27–41.

[111] Robinson D, Stone A (2005): A simplified radiosity algorithm for general urban radiation exchange. *Building Services Engineering Research and Technology*, 26(4): 271–284.

[112] Carrión D, Lorenz A, Kolbe TH (2010): Estimation of the energetic rehabilitation state of buildings for the city of Berlin using a 3D city model represented in CityGML. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XXXVIII-4/W15: 31–35.

[113] Krüger A, Kolbe TH (2012): Building analysis for urban energy planning using key indicators on virtual 3D city models—The energy atlas of Berlin. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XXXIX-B2: 145–150.

[114] Strzalka A, Bogdahn J, Coors V, Eicker U (2011): 3D City modeling for urban scale heating energy demand forecasting. *HVAC&R Research*, 17(4): 526–539.

[115] Bahu JM, Koch A, Kremers E, Murshed SM (2013): Towards a 3D spatial urban energy modelling approach. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, II-2/W1: 33–41.

[116] Nouvel R, Schulte C, Eicker U, Pietruschka D, Coors V (2013): CityGML-based 3D city model for energy diagnostics and urban energy policy support. *Proceedings of BS2013: 13th Conference of International Building Performance Simulation Association*. Chambéry, France, pp. 218–225.

[117] Nouvel R, Zirak M, Dastageeri H, Coors V, Eicker U (2014): Urban Energy Analysis Based on 3D City Model for National Scale Applications. *Proceedings of the Fifth German-Austrian IBPSA Conference (BauSIM 2014)*. Aachen, Germany, pp. 83–90.

[118] Kaden R, Kolbe TH (2013): City-wide total energy demand estimation of buildings using semantic 3D city models and statistical data. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, II-2/W1: 163–171.

[119] Perez D, Kämpf JH, Scartezzini JL (2013): Urban Area Energy Flow Microsimulation for Planning Support: a Calibration and Verification Study. *International Journal On Advances in Systems and Measurements*, 6(3&4): 260–271.

[120] Bahu JM, Koch A, Kremers E, Murshed SM (2015): Towards a 3D Spatial Urban Energy Modelling Approach. *International Journal of 3-D Information Modeling*, 3(3): 1–16.

[121] Nouvel R, Mastrucci A, Leopold U, Baume O, Coors V, Eicker U (2015): Combining GIS-based statistical and engineering urban heat consumption models: Towards a new framework for multi-scale policy support. *Energy and Buildings*, 107: 204–212.

[122] Agugiaro G (2016): Energy planning tools and CityGML-based 3D virtual city models: experiences from Trento (Italy). *Applied Geomatics*, 8(1): 41–56.

[123] Robinson D, Campbell N, Gaiser W, Kabel K, Le-Mouel A, Morel N, Page J, Stankovic S, Stone A (2007): SUNtool – A new modelling paradigm for simulating and optimising urban sustainability. *Solar Energy*, 81(9): 1196–1211.

[124] Previtali M, Barazzetti L, Brumana R, Cuca B, Oreni D, Roncoroni F, Scaioni M (2014): Automatic façade modelling using point cloud data for energy-efficient retrofitting. *Applied Geomatics*, 6(2): 95–113.

[125] Tabrizi A, Sanguinetti P (2014): Case Study: Evaluation of Renewable Energy Strategies Using Building Information Modeling and Energy Simulation. *International Journal of 3-D Information Modeling*, 2(4): 25–37.

[126] Löwner MO, Sasse A, Hecker P (2010): Needs and potential of 3D city information and sensor fusion technologies for vehicle positioning in urban environments. *Advances in 3D Geoinformation*, Springer Berlin Heidelberg, pp. 143–156.

[127] Cappelle C, El Najjar ME, Charpillet F, Pomorski D (2011): Virtual 3D City Model for Navigation in Urban Areas. *Journal of Intelligent & Robotic Systems*, 66(3): 377–399.

[128] Coors V, Huch T, Kretschmer U (2000): Matching buildings: pose estimation in an urban environment. *Proceedings of the IEEE and ACM International Symposium on Augmented Reality (ISAR 2000)*. Munich, Germany, pp. 89–92.

[129] Boeters R, Arroyo Ohori K, Biljecki F, Zlatanova S (2015): Automatically enhancing CityGML LOD2 models with a corresponding indoor geometry. *International Journal of Geographical Information Science*, 29(12): 2248–2268.

[130] Shiravi S, Zhong M, Beykaei SA, Hunt JD, Abraham JE (2015): An assessment of the utility of LiDAR data in extracting base-year floorspace and a comparison with the census-based approach. *Environment and Planning B: Planning and Design*, 42(4): 708–729.

[131] Boeters R (2013): *Automatic enhancement of CityGML LoD2 models with interiors and its usability for net internal area determination*. Master's thesis, Delft University of Technology, Delft, the Netherlands.

[132] Henn A, Römer C, Gröger G, Plümer L (2012): Automatic classification of building types in 3D city models. *Geoinformatica*, 16(2): 281–306.

[133] Köninger A, Bartel S (1998): 3D-GIS for Urban Purposes. *Geoinformatica*, 2(1): 79–103.

[134] Ellul C, Altenbuchner J (2014): Investigating approaches to improving rendering performance of 3D city models on mobile devices. *Geo-spatial Information Science*, 17(2): 73–84.

[135] Rau JY, Cheng CK (2013): A cost-effective strategy for multi-scale photo-realistic building modeling and web-based 3-D GIS applications in real estate. *Computers, Environment and Urban Systems*, 38: 35–44.

[136] Pasewaldt S, Semmo A, Trapp M, Döllner J (2014): Multi-perspective 3D panoramas. *International Journal of Geographical Information Science*, 28(10): 2030–2051.

[137] Zhang L, Han C, Zhang L, Zhang X, Li J (2014): Web-based visualization of large 3D urban building models. *International Journal of Digital Earth*, 7(1): 53–67.

[138] Evans A, Romeo M, Bahrehmand A, Agenjo J, Blat J (2014): 3D graphics on the web: A survey. *Computers and Graphics*, 41: 43–61.

[139] Jochem R, Goetz M (2012): Towards Interactive 3D City Models on the Web. *International Journal of 3-D Information Modeling*, 1(3): 26–36.

[140] Mao B, Ban Y (2011): Online Visualization of 3D City Model Using CityGML and X3DOM. *Cartographica*, 46(2): 109–114.

[141] Coors V (2003): 3D-GIS in networking environments. *Computers, Environment and Urban Systems*, 27(4): 345–357.

[142] Zhu Q, Zhao J, Du Z, Zhang Y, Xu W, Xie X, Ding Y, Wang F, Wang T (2011): Towards Semantic 3D City Modeling and Visual Explorations. TH Kolbe, G König, C Nagel, eds., *Advances in 3D Geo-Information Sciences. Lecture Notes in Geoinformation and Cartography*, Springer Berlin Heidelberg, pp. 275–294.

[143] Wolff M, Asche H (2008): Geospatial Modelling of Urban Security: A Novel Approach with Virtual 3D City Models. *Computational Science and Its Applications – ICCSA 2008*, Springer, Berlin, Heidelberg, pp. 42–51.

[144] Wolff M, Asche H (2009): Towards Geovisual Analysis of Crime Scenes – A 3D Crime Mapping Approach. *Advances in GIScience*, Springer, Berlin, Heidelberg, pp. 429–448.

[145] Besuievsky G, Patow G (2013): Procedural modeling historical buildings for serious games. *Virtual Archeology Review*, 4(9): 160–166.

[146] Rüppel U, Schatz K (2011): Designing a BIM-based serious game for fire safety evacuation simulations. *Advanced Engineering Informatics*, 25(4): 600–611.

[147] Portalés C, Lerma JL, Navarro S (2010): Augmented reality and photogrammetry: A synergy to visualize physical and virtual city environments. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65(1): 134–142.

[148] Verbree E, van Maren G, Germs R, Jansen F, Kraak MJ (1999): Interaction in virtual world views-linking 3D GIS with VR. *International Journal of Geographical Information Science*, 13(4): 385–396.

[149] Germs R, van Maren G, Verbree E, Jansen FW (1999): A multi-view VR interface for 3D GIS. *Computers and Graphics*, 23(4): 497–506.

[150] van Maren G (2003): Key to Virtual Insight: A 3D GIS and Virtual Reality System. *Planning Support Systems in Practice*, Springer Berlin Heidelberg, pp. 193–204.

[151] Takase Y, Sho N, Sone A, Shimiya K (2003): Automatic generation of 3D city models and related applications. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XXXIV-5/W10: 5.

[152] Zhou Q, Zhang W (2004): A preliminary review on three-dimensional city model. *Geo-spatial Information Science*, 7(2): 79–88.

[153] Ghadirian P, Bishop ID (2008): Integration of augmented reality and GIS: A new approach to realistic landscape visualisation. *Landscape and Urban Planning*, 86(3-4): 226–232.

[154] Vlahakis V, Ioannidis N, Karigiannis J, Tsotros M, Gounaris M, Stricker D, Gleue T, Daehne P, Almeida L (2002): Archeoguide: An Augmented Reality Guide for Archaeological Sites. *IEEE Computer Graphics and Applications*, 22(5): 52–60.

[155] Richards-Rissetto H, Remondino F, Agugiaro G, Robertsson J, von Schwerin J, Girardi G (2012): Kinect and 3D GIS in archaeology. *18th International Conference on Virtual Systems and Multimedia (VSMM)*. IEEE, Milan, Italy, pp. 331–337.

[156] Zamyadi A, Pouliot J, Bédard Y (2013): A Three Step Procedure to Enrich Augmented Reality Games with CityGML 3D Semantic Modeling. *Progress and New Trends in 3D Geoinformation Sciences*, Springer Berlin Heidelberg, pp. 261–275.

[157] Moreno A, Segura Á, Zlatanova S, Posada J, García-Alonso A (2012): Benefit of the integration of semantic 3D models in a fire-fighting VR simulator. *Applied Geomatics*, 4(3): 143–153.

[158] Glander T, Döllner J (2009): Abstract representations for interactive visualization of virtual 3D city models. *Computers, Environment and Urban Systems*, 33(5): 375–387.

[159] Shiode N (2001): 3D urban models: Recent developments in the digital modelling of urban environments in three-dimensions. *GeoJournal*, 52(3): 263–269.

[160] Kemec S, Duzgun S, Zlatanova S, Dilmen DI, Yalciner AC (2010): Selecting 3D urban visualisation models for disaster management: Fethiye tsunami inundation case. *Proceedings of the 3rd International Conference on Cartography and GIS*. Nessebar, Bulgaria.

[161] Patel VM, Dholakia MB, Singh AP (2013): Tsunami Risk 3D Visualizations of Okha Coast, Gujarat (India). *International Journal of Engineering Science and Innovative Technology*, 2(2): 130–138.

[162] Manyoky M, Wissen Hayek U, Heutschi K, Pieren R, Grêt-Regamey A (2014): Developing a GIS-Based Visual-Acoustic 3D Simulation for Wind Farm Assessment. *ISPRS International Journal of Geo-Information*, 3(1): 29–48.

[163] Kwan MP (2000): Interactive geovisualization of activity-travel patterns using three-dimensional geographical information systems: a methodological exploration with a large data set. *Transportation Research Part C – Emerging Technologies*, 8(1): 185–203.

[164] Congote J, Moreno A, Kabongo L, Pérez JL, San José R, Ruiz O (2012): Web based hybrid volumetric visualisation of urban GIS data. T Leduc, G Moreau, R Billen, eds., *Usage, Usability, and Utility of 3D City Models – European COST Action TU0801*. EDP Sciences, Nantes, France, p. 03001.

[165] San José R, Pérez JL, González-Barras RM (2011): 3D Visualization of Air Quality Data. *Proceedings of the 11th International Conference "Reliability and Statistics in Transportation and Communication" (RelStat'11)*. Riga, Latvia, pp. 1–9.

[166] San José R, Pérez JL, González RM (2012): Advances in 3D visualization of air quality data. T Leduc, G Moreau, R Billen, eds., *Usage, Usability, and Utility of 3D City Models – European COST Action TU0801*. EDP Sciences, Nantes, France, p. 02002.

[167] Yan JK (1985): Advances in Computer-Generated Imagery for Flight Simulation. *IEEE Computer Graphics and Applications*, 5(8): 37–51.

[168] Besuievsky G, Patow G (2013): Customizable LoD for Procedural Architecture. *Computer Graphics Forum*, 32(8): 26–34.

[169] Smelik RM, Tutenel T, Bidarra R, Benes B (2014): A Survey on Procedural Modelling for Virtual Worlds. *Computer Graphics Forum*, 33(6): 31–50.

[170] Jobst M, Germanchis T (2007): The Employment of 3D in Cartography — An Overview. *Multimedia Cartography*, Springer Berlin Heidelberg, pp. 217–228.

[171] Oulasvirta A, Estlander S, Nurminen A (2009): Embodied interaction with a 3D versus 2D mobile map. *Personal and Ubiquitous Computing*, 13(4): 303–320.

[172] Schilling A, Coors V, Laakso K (2005): Dynamic 3D Maps for Mobile Tourism Applications. *Map-based Mobile Services*, Springer-Verlag, pp. 227–239.

[173] Rakkolainen I, Vainio T (2001): A 3D City Info for mobile users. *Computers and Graphics*, 25(4): 619–625.

[174] Nurminen A (2008): Mobile 3D City Maps. *IEEE Computer Graphics and Applications*, 28(4): 20–31.

[175] Musliman IA, Abdul-Rahman A, Coors V (2006): 3D Navigation for 3D-GIS — Initial Requirements. *Innovations in 3D Geo Information Systems*, Springer, Berlin, Heidelberg, pp. 259–268.

[176] Coltekin A, Lokka IE, Boer A (2015): The Utilization of Publicly Available Map Types by Non-experts – A Choice Experiment. *Proceedings of the 27th International Cartographic Conference (ICC2015)*.

[177] Bernasocchi M, Çöltekin A, Gruber S (2012): An Open Source Geovisual Analytics Toolbox for Multivariate Spatio-Temporal Data for Environmental Change Modeling. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, I-2: 123–128.

[178] Nedkov S (2012): *Knowledge-based optimisation of 3D city models for car navigation devices*. Master's thesis, Delft University of Technology, Delft, the Netherlands.

[179] Mao B, Ban Y, Harrie L (2014): Real-time visualization of 3D city models at street-level based on visual saliency. *Science China Earth Sciences*, 58(3): 448–461.

[180] Wu H, He Z, Gong J (2010): A virtual globe-based 3D visualization and interactive framework for public participation in urban planning processes. *Computers, Environment and Urban Systems*, 34(4): 291–298.

[181] Franić S, Bačić-Deprato I, Novaković I (2009): 3D model and a scale model of the City of Zagreb. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XXXVIII-2/W11: 1–7.

[182] Batty M, Chapman D, Evans S, Haklay M, Kueppers S, Shiode N, Smith A, Torrens PM (2000): Visualizing the city: communicating urban design to planners and decision-makers. Tech. Rep. Paper 26, London, United Kingdom.

[183] Döllner J, Kolbe TH, Liecke F, Sgouros T, Teichmann K (2006): The Virtual 3D City Model of Berlin – Managing, Integrating, and Communicating Complex Urban Information. *Proceedings of 25th Urban Data Management Symposium (UDMS 2006)*. Aalborg, Denmark, pp. 1–12.

[184] Buhur S, Ross L, Büyüksalih G, Baz I (2009): 3D City Modelling for Planning Activities, Case Study: Haydarpasa Train Station, Haydarpasa Port and Surrounding Backside Zones, Istanbul. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XXXVIII-1-4-7/W5: 1–6.

[185] City of Perth (2013): *Planning and development application. 3D model specification*.

[186] Adelaide City Council (2009): *Development information guide of the Adelaide 3D City Model: Frequently asked questions for architects, designers and developers*.

[187] Marina O, Masala E, Pensa S, Stavric M (2012): Interactive model of urban development in residential areas in Skopje. T Leduc, G Moreau, R Billen, eds., *Usage, Usability, and Utility of 3D City Models – European COST Action TU0801*. EDP Sciences, Nantes, France, pp. (02004)1–12.

[188] Koutsoudis A, Arnaoutoglou F, Chamzas C (2007): On 3D reconstruction of the old city of Xanthi. A minimum budget approach to virtual touring based on photogrammetry. *Journal of Cultural Heritage*, 8(1): 26–31.

[189] Novaković I (2011): 3D Model of Zagreb. *GIM International*, 25(1): 25–29.

[190] Ghawana T, Zlatanova S (2013): 3D printing for urban planning: A physical enhancement of spatial perspective. *Urban and Regional Data Management UDMS Annual 2013*, Urban and Regional Data Management: UDMS Annual 2013, London, pp. 211–224.

[191] Trapp M, Semmo A, Pokorski R, Herrmann CD, Döllner J, Eichhorn M, Heinzelmann M (2010): Communication of Digital Cultural Heritage in Public Spaces by the Example of Roman Cologne. *Digital Heritage*, Springer Berlin Heidelberg, pp. 262–276.

[192] Yang PP, Putra SY, Li W (2007): Viewsphere: a GIS-based 3D visibility analysis for urban design evaluation. *Environment and Planning B: Planning and Design*, 34(6): 971.

[193] Liu L, Zhang L, Ma J, Zhang L, Zhang X, Xiao Z, Yang L (2010): An improved line-of-sight method for visibility analysis in 3D complex landscapes. *Science China Information Sciences*, 53(11): 2185–2194.

[194] Lonergan C, Hedley N (2016): Unpacking isovists: a framework for 3D spatial visibility analysis. *Cartography and Geographic Information Science*, 43(2): 87–102.

[195] Peters R, Ledoux H, Biljecki F (2015): Visibility Analysis in a Point Cloud Based on the Medial Axis Transform. *Eurographics Workshop on Urban Data Modelling and Visualisation 2015*. Delft, Netherlands, pp. 7–12.

[196] Bartie P, Reitsma F, Kingham S, Mills S (2010): Advancing visibility modelling algorithms for urban environments. *Computers, Environment and Urban Systems*, 34(6): 518–531.

[197] Delikostidis I, Engel J, Retsios B, van Elzakker CPJM, Kraak MJ, Döllner J (2013): Increasing the Usability of Pedestrian Navigation Interfaces by means of Landmark Visibility Analysis. *Journal of Navigation*, 66(4): 523–537.

[198] Albrecht F, Moser J, Hijazi I (2013): Assessing façade visibility in 3D city models for city marketing. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XL-2/W2: 1–5.

[199] Rabban IE, Abdullah K, Ali ME, Cheema MA (2015): Visibility Color Map for a Fixed or Moving Target in Spatial Databases. *Advances in Spatial and Temporal Databases*, Springer International Publishing, Cham, pp. 197–215.

[200] Ying M, Jingjue J, Fulin B (2002): 3D-City Model supporting for CCTV monitoring system. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XXXIV: 1–4.

[201] Yaagoubi R, Yarmani M, Kamel A, Khemiri W (2015): HybVOR: A Voronoi-Based 3D GIS Approach for Camera Surveillance Network Placement. *ISPRS International Journal of Geo-Information*, 4(2): 754–782.

[202] Jung Moon S, Cheol Jeon M, Dam Eo Y, Bin Im S, Wook Park B (2013): Campus CCTV Allocation Simulation for Maximizing Monitoring Areas. *Advances in information Sciences and Service Sciences*, 5(7): 1192–1198.

[203] Afghantoloee A, Doodman S, Karimipour F, Mostafavi MA (2014): Coverage estimation of geosensor in 3D vector environments. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XL-2/W3: 1–6.

[204] Bassani M, Grasso N, Piras M (2015): 3D GIS based evaluation of the available sight distance to assess safety of urban roads. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XL-3/W3: 137–143.

[205] VanHorn JE, Mosurinjohn NA (2010): Urban 3D GIS Modeling of Terrorism Sniper Hazards. *Social Science Computer Review*, 28(4): 482–496.

[206] Yu SM, Han SS, Chai CH (2007): Modeling the value of view in high-rise apartments: a 3D GIS approach. *Environment and Planning B: Planning and Design*, 34(1): 139–153.

[207] Hamilton SE, Morgan A (2010): Integrating lidar, GIS and hedonic price modeling to measure amenity values in urban beach residential property markets. *Computers, Environment and Urban Systems*, 34(2): 133–141.

[208] Tomić H, Roić M, Mastelić Ivić S (2012): Use of 3D cadastral data for real estate mass valuation in the urban areas. P van Oosterom, ed., *Proceedings of the 3rd International Workshop on 3D Cadastres: Developments and Practices*. Shenzhen, China, pp. 73–86.

[209] Ellul C, Adjrad M, Groves P (2016): The impact of 3D data quality on improving GNSS performance using city models initial simulations. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, IV-2-W1: 171–178.

[210] Wang L, Groves PD, Ziebart MK (2013): GNSS Shadow Matching: Improving Urban Positioning Accuracy Using a 3D City Model with Optimized Visibility Scoring Scheme. *Navigation*, 60(3): 195–207.

[211] Hsu LT, Gu Y, Kamijo S (2016): 3D building model-based pedestrian positioning method using GPS/GLONASS/QZSS and its reliability calculation. *GPS Solutions*, 20(3): 413–428.

[212] Bradbury J, Ziebart M, Cross PA, Boulton P, Read A (2007): Code Multipath Modelling in the Urban Environment Using Large Virtual Reality City Models: Determining the Local Environment. *Journal of Navigation*, 60(1): 95–105.

[213] Wang L, Groves PD, Ziebart MK (2012): Multi-Constellation GNSS Performance Evaluation for Urban Canyons Using Large Virtual Reality City Models. *Journal of Navigation*, 65(3): 459–476.

[214] Kumar R, Petovello MG (2014): A Novel GNSS Positioning Technique for Improved Accuracy in Urban Canyon Scenarios using 3D City Model. *Proceedings of the ION GNSS+ 2014*. Tampa, FL, United States.

[215] Gang-jun L, Kefei Z, Falin W, Liam D, Retscher G (2009): Characterisation of current and future GNSS performance in urban canyons using a high quality 3-D urban model of Melbourne, Australia. *Journal of Applied Geodesy*, 3(1): 15–24.

[216] Kleijer F, Odijk D, Verbree E (2009): Prediction of GNSS Availability and Accuracy in Urban Environments Case Study Schiphol Airport. *Location Based Services and TeleCartography II*, Springer, Berlin, Heidelberg, pp. 387–406.

[217] Bétaille D, Peyret F, Ortiz M, Miquel S, Fontenay L (2013): A New Modeling Based on Urban Trenches to Improve GNSS Positioning Quality of Service in Cities. *IEEE Intelligent Transportation Systems Magazine*, 5(3): 59–70.

[218] Groves PD, Jiang Z (2013): Height Aiding, C/N 0 Weighting and Consistency Checking for GNSS NLOS and Multipath Mitigation in Urban Areas. *Journal of Navigation*, 66(5): 653–669.

[219] Wada Y, Hsu LT, Gu Y, Kamijo S (2017): Optimization of 3D building models by GPS measurements. *GPS Solutions*, 21(1): 65–78.

[220] Piñana-Díaz C, Toledo-Moreo R, Toledo-Moreo F, Skarmeta A (2012): A Two-Layers Based Approach of an Enhanced-Mapfor Urban Positioning Support. *Sensors*, 12(11): 14508–14524.

[221] Fisher-Gewirtzman D (2012): 3D models as a platform for urban analysis and studies on human perception of space. T Leduc, G Moreau, R Billen, eds., *Usage, Usability, and Utility of 3D City Models – European COST Action TU0801*. EDP Sciences, Nantes, France, p. 01001.

[222] Fisher-Gewirtzman D, Natapov A (2014): Different approaches of visibility analyses applied on hilly urban environment. *Survey Review*, 46(338): 366–382.

[223] Yasumoto S, Jones AP, Nakaya T, Yano K (2011): The use of a virtual city model for assessing equity in access to views. *Computers, Environment and Urban Systems*, 35(6): 464–473.

[224] Yasumoto S, Jones A, Yano K, Nakaya T (2012): Virtual city models for assessing environmental equity of access to sunlight: a case study of Kyoto, Japan. *International Journal of Geographical Information Science*, 26(1): 1–13.

[225] Wilson E (2014): 3D Vision: BC Assessment's Cool New Tools. URL http://www.rew.ca/.

[226] Johnson GT, Watson ID (1984): The Determination of View-Factors in Urban Canyons. *Journal of Climate and Applied Meteorology*, 23(2): 329–335.

[227] Brasebin M, Perret J, Mustière S, Weber C (2012): Measuring the impact of 3D data geometric modeling on spatial analysis: Illustration with Skyview factor. T Leduc, G Moreau, R Billen, eds., *Usage, Usability, and Utility of 3D City Models – European COST Action TU0801*. EDP Sciences, Nantes, France, pp. (02001)1–16.

[228] Chen L, Ng E, An X, Ren C, Lee M, Wang U, He Z (2012): Sky view factor analysis of street canyons and its implications for daytime intra-urban air temperature differentials in high-rise, high-density urban areas of Hong Kong: a GIS-based simulation approach. *International Journal of Climatology*, 32(1): 121–136.

[229] Gál T, Lindberg F, Unger J (2009): Computing continuous sky view factors using 3D urban raster and vector databases: comparison and application to urban climate. *Theoretical and Applied Climatology*, 95(1): 111–123.

[230] Hwang RL, Lin TP, Matzarakis A (2011): Seasonal effects of urban street shading on long-term outdoor thermal comfort. *Building and Environment*, 46(4): 863–870.

[231] Besuievsky G, Barroso S, Beckers B, Patow G (2014): A Configurable LoD for Procedural Urban Models intended for Daylight Simulation. G Besuievsky, V Tourre, eds., *Eurographics Workshop on Urban Data Modelling and Visualisation*. The Eurographics Association, Strasbourg, France, pp. 19–24.

[232] Unger J (2009): Connection between urban heat island and sky view factor approximated by a software tool on a 3D urban database. *International Journal of Environment and Pollution*, 36(1/2/3): 59.

[233] Hämmerle M, Gál T, Unger J, Matzarakis A (2011): Comparison of models calculating the sky view factor used for urban climate investigations. *Theoretical and Applied Climatology*, 105(3-4): 521–527.

[234] Muñoz D, Beckers B, Besuievsky G, Patow G (2015): Far-LoD: Level of Detail for Massive Sky View Factor Calculations in Large Cities. *Eurographics Workshop on Urban Data Modelling and Visualisation*. pp. 1–6.

[235] Herbert G, Chen X (2015): A comparison of usefulness of 2D and 3D representations of urban planning. *Cartography and Geographic Information Science*, 42(1): 22–32.

[236] Den Haag (2011): Voorstel van het college inzake beleid dakopbouwen. Gemeente Den Haag. Dienst Stedelijke Ontwikkeling. RIS 180461.

[237] City of Mississauga (2011): Standards for shadow studies. Mississauga Planning and Building Department, Development and Design Division. Resolution no. 0266-2011. URL http://www6.mississauga.ca/onlinemaps/planbldg/UrbanDesign/ShadowStudiesFinal_Feb2012.pdf.

[238] Alam N, Coors V, Zlatanova S, Oosterom PJM (2012): Shadow effect on photovoltaic potentiality analysis using 3D city models. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XXXIX-B8: 209–214.

[239] Alam N, Coors V, Zlatanova S (2013): Detecting shadow for direct radiation using CityGML models for photovoltaic potentiality analysis. C Ellul, S Zlatanova, M Rumor, R Laurini, eds., *Urban and Regional Data Management*, CRC Press, London, UK, pp. 191–196.

[240] Mardaljevic J, Rylatt M (2003): Irradiation mapping of complex urban environments: an image-based approach. *Energy and Buildings*, 35(1): 27–35.

[241] Nguyen HT, Pearce JM (2012): Incorporating shading losses in solar photovoltaic potential assessment at the municipal scale. *Solar Energy*, 86(5): 1245–1260.

[242] Tooke TR, Coops NC, Voogt JA, Meitner MJ (2011): Tree structure influences on rooftop-received solar radiation. *Landscape and Urban Planning*, 102(2): 73–81.

[243] Eicker U, Monien D, Duminil E, Nouvel R (2015): Energy performance assessment in urban planning competitions. *Applied Energy*, 155: 323–333.

[244] Yezioro A, Shaviv E (1994): Shading: A design tool for analyzing mutual shading between buildings. *Solar Energy*, 52(1): 27–37.

[245] Morello E, Ratti C (2009): Sunscapes: 'Solar envelopes' and the analysis of urban DEMs. *Computers, Environment and Urban Systems*, 33(1): 26–34.

[246] Knowles RL (2003): The solar envelope: its meaning for energy and buildings. *Energy and Buildings*, 35(1): 15–25.

[247] Lange E, Hehl-Lange S (2005): Combining a participatory planning approach with a virtual landscape model for the siting of wind turbines. *Journal of Environmental Planning and Management*, 48(6): 833–852.

[248] Coors V, Holweg D, Matthias E, Petzold B (2013): Broschüre "3D-Stadtmodelle". In GeoForum, Kommission 3D-Stadtmodelle der Deutschen Gesellschaft für Kartographie (DGfK), and Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation (DGPF).

[249] de Kluijver H, Stoter J (2003): Noise mapping and GIS: optimising quality and efficiency of noise effect studies. *Computers, Environment and Urban Systems*, 27(1): 85–102.

[250] Kurakula V, Kuffer M (2008): 3D noise modeling for urban environmental planning and management. M Schrenk, V Popovich, D Engelke, P Elisei, eds., *13th International Conference on Urban Planning in Information and Knowledge Society (REAL CORP 2008)*. Vienna, Austria, pp. 517–523.

[251] Pamanikabud P, Tansatcha M (2009): Geoinformatic prediction of motorway noise on buildings in 3D GIS. *Transportation Research Part D: Transport and Environment*, 14(5): 367–372.

[252] Lu L, Becker T, Löwner MO (2016): 3D Complete Traffic Noise Analysis Based on CityGML. *Advances in 3D Geoinformation*, Springer International Publishing, Cham, pp. 265–283.

[253] Ranjbar HR, Gharagozlou AR, Nejad ARV (2012): 3D Analysis and Investigation of Traffic Noise Impact from Hemmat Highway Located in Tehran on Buildings and Surrounding Areas. *Journal of Geographic Information System*, 4(4): 322–334.

[254] Law CW, Lee CK, Lui ASw, Yeung MKl, Lam Kc (2011): Advancement of three-dimensional noise mapping in Hong Kong. *Applied Acoustics*, 72(8): 534–543.

[255] EC (2002): Directive 2002/49/EC of the European Parliament and of the Council. *Official Journal of the European Communities*, (L189): 12–25.

[256] Butler D (2004): Noise management: Sound and vision. *Nature*, 427(6974): 480–481.

[257] Czerwinski A, Kolbe TH, Plümer L, Elke SM (2006): Interoperability and accuracy requirements for EU environmental noise mapping. H Kremers, V Tikunov, eds., *International Conference on GIS and Sustainable Development (InterCarto – InterGIS 12)*. Berlin, Germany, pp. 182–194.

[258] Kubiak J, Ławniczak R (2016): The propagation of noise in a built-up area (on the example of a housing estate in Poznań). *Journal of Maps*, 12(2): 231–236.

[259] Stoter J, de Kluijver H, Kurakula V (2008): 3D noise mapping in urban areas. *International Journal of Geographical Information Science*, 22(8): 907–924.

[260] Law CW, Lee CK, Tai MK (2006): Visualization of Complex Noise Environment by Virtual Reality Technologies. *Symposium of the campaign "Science in the Public Service"*. Environmental Protection Department, Hong Kong Institution of Science, Hong Kong, p. 8.

[261] Czerwinski A, Kolbe TH, Plümer L, Elke SM (2006): Spatial data infrastructure techniques for flexible noise mapping strategies. K Tochtermann, A Scharl, eds., *Proceedings of the 20th International Conference on Environmental Informatics-Managing Environmental Knowledge*. Graz, Austria, pp. 99–106.

[262] Kurakula V (2007): *A GIS-Based Approach for 3D Noise Modelling Using 3D City Models*. Master's thesis, International Institute for Geoinformation Science and Earth Observation, Enschede, the Netherlands.

[263] Billen R, Zlatanova S (2003): 3D spatial relationships model: a useful concept for 3D cadastre? *Computers, Environment and Urban Systems*, 27(4): 411–425.

[264] Stoter JE, van Oosterom PJM (2005): Technological aspects of a full 3D cadastral registration. *International Journal of Geographical Information Science*, 19(6): 669–696.

[265] Stoter J, Ploeger H, van Oosterom P (2013): 3D cadastre in the Netherlands: Developments and international applicability. *Computers, Environment and Urban Systems*, 40: 56–67.

[266] van Oosterom P (2013): Research and development in 3D cadastres. *Computers, Environment and Urban Systems*, 40: 1–6.

[267] Çağdaş V (2013): An Application Domain Extension to CityGML for immovable property taxation: A Turkish case study. *International Journal of Applied Earth Observation and Geoinformation*, 21: 545–555.

[268] Frédéricque B, Raymond K, van Prooijen K (2011): 3D GIS Applied to Cadastre. Tech. rep., Bentley Systems, Incorporated.

[269] Guo R, Li L, Ying S, Luo P, He B, Jiang R (2013): Developing a 3D cadastre for the administration of urban land use: A case study of Shenzhen, China. *Computers, Environment and Urban Systems*, 40: 46–55.

[270] Pouliot J, Roy T, Fouquet-Asselin G, Desgroseilliers J (2011): 3D Cadastre in the Province of Quebec: A First Experiment for the Construction of a Volumetric Representation. *Advances in 3D Geo-Information Sciences*, Springer Berlin Heidelberg, pp. 149–162.

[271] Soon KH (2012): A conceptual framework of representing semantics for 3D Cadastre in Singapore. *3rd international workshop on 3D Cadastres*. Shenzen, China, pp. 361–379.

[272] Vandysheva N, Sapelnikov S, van Oosterom PJM, De Vries ME, Spiering B, Wouters R, Hoogeveen A, Penkov V (2012): The 3D cadastre prototype and pilot in the Russian Federation. *FIG Working Week 2012*. International Federation of Surveyors (FIG), Rome, Italy.

[273] Stoter J, Ploeger H, Roes R, van der Riet E, Biljecki F, Ledoux H (2016): First 3D Cadastral Registration of Multi-level Ownerships Rights in the Netherlands. *5th International FIG Workshop on 3D Cadastres*. Athens, Greece, pp. 491–504.

[274] Shojaei D, Kalantari M, Bishop ID, Rajabifard A, Aien A (2013): Visualization requirements for 3D cadastral systems. *Computers, Environment and Urban Systems*, 41: 39–54.

[275] Pouliot J, Wang C, Hubert F, Fuchs V (2014): Empirical Assessment of the Suitability of Visual Variables to Achieve Notarial Tasks Established from 3D Condominium Models. *Innovations in 3D Geo-Information Sciences*, Springer International Publishing, pp. 195–210.

[276] Ranzinger M, Gleixner G (1997): GIS datasets for 3D urban planning. *Computers, Environment and Urban Systems*, 21(2): 159–173.

[277] Lewis JL, Casello JM, Groulx M (2012): Effective Environmental Visualization for Urban Planning and Design: Interdisciplinary Reflections on a Rapidly Evolving Technology. *Journal of Urban Technology*, 19(3): 85–106.

[278] Chen R (2011): The development of 3D city model and its applications in urban planning. *Proceedings of the 19th International Conference on Geoinformatics*. Shanghai, China, pp. 1–5.

[279] Pullar DV, Tidey ME (2001): Coupling 3D visualisation to qualitative assessment of built environment designs. *Landscape and Urban Planning*, 55(1): 29–40.

[280] Sabri S, Pettit CJ, Kalantari M, Rajabifard A, White M, Lade O, Ngo T (2015): What are Essential requirements in Planning for Future Cities using Open Data Infrastructures and 3D Data Models? *14th International Conference on Computers in Urban Planning and Urban Management*. Cambridge, MA USA, pp. 314(1–17).

[281] Métral C, Falquet G, Vonlanthen M (2007): An Ontology-based Model for Urban Planning Communication. *Ontologies for Urban Development*, Springer, Berlin, Heidelberg, pp. 61–72.

[282] Kibria MS, Zlatanova S, Itard L, Dorst M (2009): GeoVEs as Tools to Communicate in Urban Projects: Requirements for Functionality and Visualization. *3D Geo-Information Sciences*, Springer Berlin Heidelberg, pp. 379–395.

[283] Appleton K, Lovett A (2003): GIS-based visualisation of rural landscapes: defining 'sufficient' realism for environmental decision-making. *Landscape and Urban Planning*, 65(3): 117–131.

[284] Benner J, Geiger A, Häfele KH (2010): Concept for building licensing based on standardized 3d geo information. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XXXVIII-4/W15: 9–12.

[285] Leszek K (2015): Environmental and Urban Spatial Analysis Based on a 3D City Model. *Computational Science and Its Applications – ICCSA 2015*, Springer International Publishing, pp. 633–645.

[286] Isikdag U, Zlatanova S (2010): Interactive modelling of buildings in Google Earth: A 3D tool for Urban Planning. *Developments in 3D Geo-Information Sciences*, Springer, Berlin, Heidelberg, pp. 52–70.

[287] Lu S, Wang F (2014): Computer aided design system based on 3D GIS for park design. *Computer, Intelligent Computing and Education Technology*, CRC Press, pp. 413–416.

[288] Moser J, Albrecht F, Kosar B (2010): Beyond Visualisation - 3D GIS Analyses for Virtual City Models. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XXXVIII-4/W15: 143–146.

[289] Kaňuk J, Gallay M, Hofierka J (2015): Generating time series of virtual 3-D city models using a retrospective approach. *Landscape and Urban Planning*, 139: 40–53.

[290] Guney C, Akdag Girginkaya S, Cagdas G, Yavuz S (2012): Tailoring a geomodel for analyzing an urban skyline. *Landscape and Urban Planning*, 105(1-2): 160–173.

[291] Czyńska K, Rubinowicz P (2014): Application of 3D virtual city models in urban analyses of tall buildings – today practice and future challenges. *Architecturae et Artibus*, 19: 9–13.

[292] Chun W, Ge C, Yanyan L, Horne M (2008): Virtual-Reality Based Integrated Traffic Simulation for Urban Planning. *Proceedings of the 2008 International Conference on Computer Science and Software Engineering*. IEEE, Wuhan, China, pp. 1137–1140.

[293] Liu Y, Gevers T, Li X (2015): Estimation of Sunlight Direction Using 3D Object Models. *IEEE Transactions on Image Processing*, 24(3): 932–942.

[294] Auer S, Hinz S, Bamler R (2010): Ray-Tracing Simulation Techniques for Understanding High-Resolution SAR Images. *IEEE Transactions on Geoscience and Remote Sensing*, 48(3): 1445–1456.

[295] Franceschetti G, Iodice A, Riccio D, Ruello G (2003): SAR raw signal simulation for urban structures. *IEEE Transactions on Geoscience and Remote Sensing*, 41(9): 1986–1995.

[296] Margarit G, Mallorqui JJ, Pipia L (2010): Polarimetric Characterization and Temporal Stability Analysis of Urban Target Scattering. *IEEE Transactions on Geoscience and Remote Sensing*, 48(4): 2038–2048.

[297] Franceschetti G, Iodice A, Riccio D (2002): A canonical problem in electromagnetic backscattering from buildings. *IEEE Transactions on Geoscience and Remote Sensing*, 40(8): 1787–1801.

[298] Margarit G, Mallorqui JJ, Lopez-Martinez C (2007): Grecosar, a SAR simulator for complex targets: Application to urban environments. *2007 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, Barcelona, Spain, pp. 4160–4163.

[299] Zlatanova S, Beetz J (2012): 3D spatial information infrastructure: The case of Port Rotterdam. T Leduc, G Moreau, R Billen, eds., *Usage, Usability, and Utility of 3D City Models – European COST Action TU0801*. EDP Sciences, Nantes, France, p. 03010.

[300] Tang K, Xu JK (2012): Research on Application of Airport Simulation System Based on 3D GIS. *Applied Mechanics and Materials*, 198-199: 717–720.

[301] Liu R, Issa RRA (2012): 3D Visualization of Sub-Surface Pipelines in Connection with the Building Utilities: Integrating GIS and BIM for Facility Management. *International Conference on Computing in Civil Engineering*. American Society of Civil Engineers, Clearwater Beach, Florida, United States, pp. 341–348.

[302] Hijazi IH, Ehlers M, Zlatanova S (2012): NIBU: a new approach to representing and analysing interior utility networks within 3D geo-information systems. *International Journal of Digital Earth*, 5(1): 22–42.

[303] Becker T, Nagel C, Kolbe TH (2013): Semantic 3D Modeling of Multi-Utility Networks in Cities for Analysis and 3D Visualization. *Progress and New Trends in 3D Geoinformation Sciences*, Springer, Berlin, Heidelberg, pp. 41–62.

[304] Løvset T, Ulvang DM, Bekkvik TC, Villanger K, Viola I (2013): Rule-based method for automatic scaffold assembly from 3D building models. *Computers and Graphics*, 37(4): 256–268.

[305] Kwan MP, Lee J (2005): Emergency response after 9/11: the potential of real-time 3D GIS for quick emergency response in micro-spatial environments. *Computers, Environment and Urban Systems*, 29(2): 93–113.

[306] Tashakkori H, Rajabifard A, Kalantari M (2015): A new 3D indoor/outdoor spatial model for indoor emergency response facilitation. *Building and Environment*, 89: 170–182.

[307] Chen LC, Wu CH, Shen TS, Chou CC (2014): The application of geometric network models and building information models in geospatial environments for fire-fighting simulations. *Computers, Environment and Urban Systems*, 45: 1–12.

[308] Corre Y, Lostanlen Y (2009): Three-Dimensional Urban EM Wave Propagation Model for Radio Network Planning and Optimization Over Large Areas. *IEEE Transactions on Vehicular Technology*, 58(7): 3112–3123.

[309] Lebherz M, Wiesbeck W, Krank W (1992): A versatile wave propagation model for the VHF/UHF range considering three-dimensional terrain. *IEEE Transactions on Antennas and Propagation*, 40(10): 1121–1131.

[310] Yang G, Pahlavan K, Lee JF, Dagen AJ, Vancraeynest J (1994): Prediction of radio wave propagation in four blocks of New York City using 3D ray tracing. *Proceedings of the 5th IEEE International Symposium on Wireless Networks Catching the mobile future*. IOS Press, The Hague, the Netherlands, pp. 263–267.

[311] Rautiainen T, Wolfle G, Hoppe R (2002): Verifying path loss and delay spread predictions of a 3D ray tracing propagation model in urban environment. *Proceedings of the IEEE 56th Vehicular Technology Conference*. IEEE, Vancouver, Canada, pp. 2470–2474.

[312] Wagen JF, Rizk K (2003): Radiowave propagation, building databases, and GIS: anything in common? A radio engineer's viewpoint. *Environment and Planning B: Planning and Design*, 30(5): 767–787.

[313] Yun Z, Iskander MF, Lim SY, He D, Martinez R (2007): Radio Wave Propagation Prediction Based on 3-D Building Structures Extracted From 2-D Images. *Antennas and Wireless Propagation Letters*, 6(99): 557–559.

[314] Beekhuizen J, Heuvelink GBM, Huss A, Bürgi A, Kromhout H, Vermeulen R (2014): Impact of input data uncertainty on environmental exposure assessment models: A case study for electromagnetic field modelling from mobile phone base stations. *Environmental Research*, 135: 148–155.

[315] Lee G (2015): 3D coverage location modeling of Wi-Fi access point placement in indoor environment. *Computers, Environment and Urban Systems*, 54: 326–335.

[316] Asawa T, Hoyano A, Yoshida T, Takata M (2012): A New Approach to Microclimate Analysis Using Airborne Remote Sensing, 3D-GIS and CFD Simulation. *Proceedings of the 1st Asia conference of International Building Performance Simulation Association*. Shanghai, China, pp. 1–8.

[317] Ujang U, Anton F, Abdul-Rahman A (2013): Unified Data Model of Urban Air Pollution Dispersion and 3D Spatial City Model: Groundwork Assessment towards Sustainable Urban Development for Malaysia. *Journal of Environmental Protection*, 4(7): 701–712.

[318] Musy M, Malys L, Morille B, Inard C (2015): The use of SOLENE-microclimat model to assess adaptation strategies at the district scale. *Urban Climate*, 14: 213–223.

[319] Lee D, Pietrzyk P, Donkers S, Liem V, van Oostveen J, Montazeri S, Boeters R, Colin J, Kastendeuch P, Nerry F, Menenti M, Gorte B, Verbree E (2013): Modeling and observation of heat losses from buildings: The impact of geometric detail on 3D heat flux modeling. *Proceedings of the 33rd European Association of Remote Sensing Laboratories (EARSeL) Symposium*. Matera, Italy, pp. 353–372.

[320] Amorim JH, Valente J, Pimentel C, Miranda AI, Borrego C (2012): Detailed modelling of the wind comfort in a city avenue at the pedestrian level. T Leduc, G Moreau, R Billen, eds., *Usage, Usability, and Utility of 3D City Models – European COST Action TU0801*. EDP Sciences, Nantes, France, pp. (03008)1–6.

[321] Janssen WD, Blocken B, van Hooff T (2013): Pedestrian wind comfort around buildings: Comparison of wind comfort criteria based on whole-flow field data for a complex case study. *Building and Environment*, 59: 547–562.

[322] Maragkogiannis K, Kolokotsa D, Maravelakis E, Konstantaras A (2014): Combining terrestrial laser scanning and computational fluid dynamics for the study of the urban thermal environment. *Sustainable Cities and Society*, 13: 207–216.

[323] Willenborg B (2015): *Simulation of explosions in urban space and result analysis based on CityGML-City Models and a cloud-based 3D-Webclient*. Master's thesis, Technical University Munich.

[324] Upadhyay G, Kämpf JH, Scartezzini JL (2014): Ground Temperature Modelling: The Case Study of Rue des Maraîchers in Geneva. *Eurographics Workshop on Urban Data Modelling and Visualisation*.

[325] Hsieh CM, Aramaki T, Hanaki K (2011): Managing heat rejected from air conditioning systems to save energy and improve the microclimates of residential buildings. *Computers, Environment and Urban Systems*, 35(5): 358–367.

[326] Zheng MH, Guo YR, Ai XQ, Qin T, Wang Q, Xu JM (2008): Coupling GIS with CFD modeling to simulate urban pollutant dispersion. *2010 International Conference on Mechanic Automation and Control Engineering (MACE)*. IEEE, Wuhan, China, pp. 1785–1788.

[327] Kunze C, Hecht R (2015): Semantic enrichment of building data with volunteered geographic information to improve mappings of dwelling units and population. *Computers, Environment and Urban Systems*, 53: 4–18.

[328] Wu SS, Qiu X, Wang L (2005): Population Estimation Methods in GIS and Remote Sensing: A Review. *GIScience & Remote Sensing*, 42(1): 80–96.

[329] Lwin K, Murayama Y (2009): A GIS Approach to Estimation of Building Population for Micro-spatial Analysis. *Transactions in GIS*, 13(4): 401–414.

[330] Wu Ss, Wang L, Qiu X (2008): Incorporating GIS Building Data and Census Housing Statistics for Sub-Block-Level Population Estimation. *The Professional Geographer*, 60(1): 121–135.

[331] Ural S, Hussain E, Shan J (2011): Building population mapping with aerial imagery and GIS data. *International Journal of Applied Earth Observation and Geoinformation*, 13(6): 841–852.

[332] Silván-Cárdenas JL, Wang L, Rogerson P, Wu C, Feng T, Kamphaus BD (2010): Assessing fine-spatial-resolution remote sensing for small-area population estimation. *International Journal of Remote Sensing*, 31(21): 5605–5634.

[333] Lu Z, Im J, Quackenbush L, Halligan K (2010): Population estimation based on multi-sensor data fusion. *International Journal of Remote Sensing*, 31(21): 5587–5604.

[334] Kressler F, Steinnocher K (2008): Object-oriented analysis of image and LiDAR data and its potential for a dasymetric mapping application. *Object-Based Image Analysis*, Springer Berlin Heidelberg, pp. 611–624.

[335] Dong P, Ramesh S, Nepali A (2010): Evaluation of small-area population estimation using LiDAR, Landsat TM and parcel data. *International Journal of Remote Sensing*, 31(21): 5571–5586.

[336] Lwin KK, Murayama Y (2011): Estimation of Building Population from LIDAR Derived Digital Volume Model. *Spatial Analysis and Modeling in Geographical Transformation Process*, Springer Netherlands, Dordrecht, pp. 87–98.

[337] Qiu F, Sridharan H, Chun Y (2010): Spatial Autoregressive Model for Population Estimation at the Census Block Level Using LIDAR-derived Building Volume Information. *Cartography and Geographic Information Science*, 37(3): 239–257.

[338] Alahmadi M, Atkinson PM, Martin D (2016): A Comparison of Small-Area Population Estimation Techniques Using Built-Area and Height Data, Riyadh, Saudi Arabia. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 9(5): 1959–1969.

[339] Sridharan H, Qiu F (2013): A Spatially Disaggregated Areal Interpolation Model Using Light Detection and Ranging-Derived Building Volumes. *Geographical Analysis*, 45(3): 238–258.

[340] Tutschku K (1998): Demand-based radio network planning of cellular mobile communication systems. *IEEE INFOCOM'98 Conference on Computer Communications Seventeenth Annual Joint Conference of the IEEE Computer and Communications Societies*. IEEE, San Francisco, CA, United States, pp. 1054–1061.

[341] Schneiderbauer S, Ehrlich D (2005): Population Density Estimations for Disaster Management: Case Study Rural Zimbabwe. *Geo-information for Disaster Management*, Springer Berlin Heidelberg, pp. 901–921.

[342] Akbar M, Aliabadi S, Patel R, Watts M (2013): A fully automated and integrated multi-scale forecasting scheme for emergency preparedness. *Environmental Modelling & Software*, 39: 24–38.

[343] Hildebrandt D, Timm R (2014): An assisting, constrained 3D navigation technique for multiscale virtual 3D city models. *Geoinformatica*, 18(3): 537–567.

[344] Slingsby A, Raper J (2008): Navigable space in 3D city models for pedestrians. P van Oosterom, S Zlatanova, F Penninga, E Fendel, eds., *Advances in 3D Geoinformation Systems*, Springer, Delft, the Netherlands, pp. 49–64.

[345] Thill JC, Dao THD, Zhou Y (2011): Traveling in the three-dimensional city: applications in route planning, accessibility assessment, location analysis and beyond. *Journal of Transport Geography*, 19(3): 405–421.

[346] Liu L, Zlatanova S (2011): A "door-to-door" Path-Finding Approach for Indoor Navigation. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, II-4: 45–51.

[347] Jamali A, Abdul-Rahman A, Boguslawski P, Kumar P, Gold CM (2015): An automated 3D modeling of topological indoor navigation network. *GeoJournal*, pp. 1–14.

[348] Nagel C (2014): *Spatio-Semantic Modelling of Indoor Environments for Indoor Navigation*. Ph.D. thesis, TU Berlin.

[349] Isikdag U, Zlatanova S, Underwood J (2013): A BIM-Oriented Model for supporting indoor navigation requirements. *Computers, Environment and Urban Systems*, 41: 112–123.

[350] Kim JS, Yoo SJ, Li KJ (2014): Integrating IndoorGML and CityGML for Indoor Space. D Pfoser, KJ Li, eds., *Web and Wireless Geographical Information Systems*, Springer Berlin Heidelberg, pp. 184–196.

[351] Kim Y, Kang H, Lee J (2014): Developing CityGML Indoor ADE to Manage Indoor Facilities. *Innovations in 3D Geo-Information Sciences*, Springer International Publishing, pp. 243–265.

[352] Sternberg H, Keller F, Willemsen T (2013): Precise indoor mapping as a basis for coarse indoor navigation. *Journal of Applied Geodesy*, 7(4): 231–246.

[353] Atila U, Karas IR, Abdul-Rahman A (2013): A 3D-GIS Implementation for Realizing 3D Network Analysis and Routing Simulation for Evacuation Purpose. *Progress and New Trends in 3D Geoinformation Sciences*, Springer Berlin Heidelberg, pp. 249–260.

[354] Atila U, Karas IR, Abdul-Rahman A (2013): Integration of CityGML and Oracle Spatial for implementing 3D network analysis solutions and routing simulation within 3D-GIS environment. *Geo-spatial Information Science*, 16(4): 221–237.

[355] Atila U, Karas IR, Turan MK, Abdul-Rahman A (2014): Automatic Generation of 3D Networks in CityGML and Design of an Intelligent Individual Evacuation Model for Building Fires Within the Scope of 3D GIS. *Innovations in 3D Geo-Information Sciences*, Springer International Publishing, pp. 123–142.

[356] Zhang L, Wang Y, Shi H, Zhang L (2012): Modeling and analyzing 3D complex building interiors for effective evacuation simulations. *Fire Safety Journal*, 53: 1–12.

[357] Meijers M, Zlatanova S, Pfeifer N (2005): 3D geoinformation indoors: structuring for evacuation. *Proceedings of the 1st ISPRS/EuroSDR/DGPF International workshop on Next Generation 3D City Models*. Bonn, Germany, pp. 1–6.

[358] Schaap J, Zlatanova S, van Oosterom P (2011): Towards a 3D geo-data model to support pedestrian routing in multimodal public transport travel advices. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XXXVIII-4/C21: 59–76.

[359] Kim K, Wilson JP (2014): Planning and visualising 3D routes for indoor and outdoor spaces using CityEngine. *Journal of Spatial Science*, 60(1): 179–193.

[360] Dao THD, Zhou Y, Thill JC, Delmelle E (2012): Spatio-temporal location modeling in a 3D indoor environment: the case of AEDs as emergency medical devices. *International Journal of Geographical Information Science*, 26(3): 469–494.

[361] Christodoulou S, Vamvatsikos D, Georgiou C (2011): A BIM-based framework for forecasting and visualizing seismic damage, cost and time to repair. *Proceedings of the European Conference on Product and Process Modelling*. CRC Press, Cork, Ireland, pp. 33–38.

[362] Kemec S, Zlatanova S, Duzgun HS (2010): A Framework for Defining a 3D Model in Support of Risk Management. *Geographic Information and Cartography for Risk and Crisis Management*, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 69–82.

[363] Jain SK, Singh RD, Seth SM (2000): Design Flood Estimation Using GIS Supported GIUHApproach. *Water Resources Management*, 14(5): 369–376.

[364] Liu YB, Gebremeskel S, De Smedt F, Hoffmann L, Pfister L (2003): A diffusive transport approach for flow routing in GIS-based flood modeling. *Journal of Hydrology*, 283(1-4): 91–106.

[365] Schulte C, Coors V (2008): Development of a CityGML ADE for dynamic 3D flood information. *Joint ISCRAM-CHINA and GI4DM Conference on Informa- tion Systems for Crisis Management*. Harbin, China, pp. 1–12.

[366] Varduhn V, Mundani RP, Rank E (2015): Multiresolution Models: Recent Progress in Coupling 3D Geometry to Environmental Numerical Simulation. *3D Geoinformation Science*, Springer International Publishing, pp. 55–69.

[367] Amirebrahimi S, Rajabifard A, Mendis P, Ngo T (2015): A framework for a microscale flood damage assessment and visualization for a building using BIM–GIS integration. *International Journal of Digital Earth*, 9(4): 363–386.

[368] Liu J, Gong JH, Liang JM, Li Y, Kang LC, Song LL, Shi SX (2016): A quantitative method for storm surge vulnerability assessment – a case study of Weihai city. *International Journal of Digital Earth*, Advance online publication.

[369] Sharkawi KH, Abdul-Rahman A (2014): Towards an Efficient City Inventory Management System for Urban Authorities in Developing Countries–The Case of 3D Change Detection. *The Electronic Journal of Information Systems in Developing Countries*, 60(3): 1–13.

[370] Pedrinis F, Morel M, Gesquière G (2015): Change Detection of Cities. *3D Geoinformation Science*, Springer International Publishing, pp. 123–139.

[371] Qin R (2014): Change detection on LOD 2 building models with very high resolution spaceborne stereo imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 96: 179–192.

[372] GeoSignum (2015): Rotterdam analyses city lidar data with new technique. URL https://www.gim-international.com/.

[373] Ahmed FC, Sekar SP (2015): Using Three-Dimensional Volumetric Analysis in Everyday Urban Planning Processes. *Applied Spatial Analysis and Policy*, 8(4): 393–408.

[374] Ghassoun Y, Löwner MO, Weber S (2015): Exploring the Benefits of 3D City Models in the Field of Urban Particles Distribution Modelling—A Comparison of Model Results. *Advances in 3D Geoinformation*, Springer International Publishing, pp. 193–205.

[375] Hamaina R, Leduc T, Moreau G (2014): A New Method to Characterize Density Adapted to a Coarse City Model. V Popovich, C Claramunt, M Schrenk, K Korolenko, eds., *Information Fusion and Geographic Information Systems (IF and GIS 2013)*, Springer Berlin Heidelberg, pp. 249–263.

[376] Ghassoun Y, Ruths M, Löwner MO, Weber S (2015): Intra-urban variation of ultrafine particles as evaluated by process related land use and pollutant driven regression modelling. *Science of The Total Environment*, 536: 150–160.

[377] Ghassoun Y, Löwner MO (2016): Comparison of 2D & 3D Parameter-Based Models in Urban Fine Dust Distribution Modelling. *Advances in 3D Geoinformation*, Springer International Publishing, Cham, pp. 231–246.

[378] Roßmann J, Hoppen M, Bücken A (2015): GML-Based Data Management and Semantic World Modelling for a 4D Forest Simulation and Information System. *International Journal of 3-D Information Modeling*, 3(3): 50–67.

[379] Rahlf J, Breidenbach J, Solberg S, Næsset E, Astrup R (2014): Comparison of four types of 3D data for timber volume estimation. *Remote sensing of Environment*, 155: 325–333.

[380] Piccoli C (2013): CityEngine for Archaeology. *Presentation for the Mini Conference 3D GIS for Mapping the Via Appia*. Amsterdam, Netherlands.

[381] De Hond R (2013): Mapping the Via Appia - 3D GIS. *Presentation for the Mini Conference 3D GIS for Mapping the Via Appia*. Amsterdam, Netherlands.

[382] De Roo B, Bourgeois J, De Maeyer P (2016): Usability Assessment of a Virtual Globe-Based 4D Archaeological GIS. *Advances in 3D Geoinformation*, Springer International Publishing, Cham, pp. 323–335.

[383] Gaiani M, Gamberini E, Tonelli G (2001): VR as work tool for architectural & archaeological restoration: the ancient Appian way 3D web virtual GIS. *Proceedings of the Seventh International Conference on Virtual Systems and Multimedia*. IEEE, Berkeley, CA, USA, pp. 86–95.

[384] Apollonio FI, Gaiani M, Benedetti B (2012): 3D reality-based artefact models for the management of archaeological sites using 3D Gis: a framework starting from the case study of the Pompeii Archaeological area. *Journal of Archaeological Science*, 39(5): 1271–1287.

[385] Losier LM, Pouliot J, Fortin M (2007): 3D geometrical modeling of excavation units at the archaeological site of Tell 'Acharneh (Syria). *Journal of Archaeological Science*, 34(2): 272–288.

[386] Katsianis M, Tsipidis S, Kotsakis K, Kousoulakou A (2008): A 3D digital workflow for archaeological intra-site research using GIS. *Journal of Archaeological Science*, 35(3): 655–667.

[387] Wüst T, Nebiker S, Landolt R (2004): Applying the 3D GIS DILAS to Archaeology and Cultural Heritage Projects Requirements and First Results. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XXXV/P-B5: 407–412.

[388] von Schwerin J, Richards-Rissetto H, Remondino F, Spera MG, Auer M, Billen N, Loos L, Stelson L, Reindel M (2016): Airborne LiDAR acquisition, post-processing and accuracy-checking for a 3D WebGIS of Copan, Honduras. *Journal of Archaeological Science: Reports*, 5: 85–104.

[389] Auer M, Agugiaro G, Billen N, Loos L, Zipf A (2014): Web-based Visualization and Query of semantically segmented multiresolution 3D Models in the Field of Cultural Heritage. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, II-5: 33–39.

[390] von Schwerin J, Richards-Rissetto H, Remondino F, Agugiaro G, Girardi G (2013): The MayaArch3D project: A 3D WebGIS for analyzing ancient architecture and landscapes. *Literary and Linguistic Computing*, 28(4): 736–753.

[391] Succar B (2009): Building information modelling framework: A research and delivery foundation for industry stakeholders. *Automation in Construction*, 18(3): 357–375.

[392] BIM Forum (2016): Level of development specification. URL http://www.bimforum.org/lod.

[393] Vilgertshofer S, Jubierre JR, Borrmann A (2016): IfcTunnel – A proposal for a multi-scale extension of the IFC data model for shield tunnels under consideration of downward compatibility aspects. *11th European Conference on Product and Process Modelling*. Limassol, Cyprus.

[394] Volk R, Stengel J, Schultmann F (2014): Building Information Modeling (BIM) for existing buildings — Literature review and future needs. *Automation in Construction*, 38: 109–127.

[395] Bolpagni M, Ciribini ALC (2016): The Information Modeling and the Progression of Data-Driven Projects. *CIB World Building Congress*. Tampere, Finland, pp. 296–307.

[396] Biljecki F, Ledoux H, Stoter J, Zhao J (2014): Formalisation of the level of detail in 3D city modelling. *Computers, Environment and Urban Systems*, 48: 1–15.

[397] Biljecki F, Zhao J, Stoter J, Ledoux H (2013): Revisiting the concept of level of detail in 3D city modelling. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, II-2/W1: 63–74.

[398] Çöltekin A, Reichenbacher T (2011): High Quality Geographic Services and Bandwidth Limitations. *Future Internet*, 3(4): 379–396.

[399] Mao B (2011): *Visualisation and Generalisation of 3D City Models*. Ph.D. thesis, KTH Royal Institute of Technology, Stockholm, Sweden.

[400] Luebke D, Reddy M, Cohen JD, Varshney A, Watson B, Huebner R (2003): *Level of detail for 3D graphics*. Morgan Kaufmann Pub, San Francisco.

[401] Meng L, Forberg A (2007): 3D building generalisation. W Mackaness, A Ruas, T Sarjakoski, eds., *Challenges in the portrayal of geographic information: Issues of Generalisation and Multi Scale Representation*, Elsevier Science, Amsterdam, the Netherlands, pp. 211–232.

[402] Glander T, Döllner J (2008): Automated cell based generalization of virtual 3D city models with dynamic landmark highlighting. *Proceedings of the 11th ICA Workshop on Generalization and Multiple Representation*. Montpellier, France, pp. 1–14.

[403] Forberg A (2007): Generalization of 3D building data based on a scale-space approach. *ISPRS Journal of Photogrammetry and Remote Sensing*, 62(2): 104–111.

[404] Sester M (2007): 3D Visualization and Generalization. D Fritsch, ed., *Proceedings of the 51st Photogrammetric Week '07*. Stuttgart, Germany, pp. 285–295.

[405] Goetz M (2013): Towards generating highly detailed 3D CityGML models from OpenStreetMap. *International Journal of Geographical Information Science*, 27(5): 845–865.

[406] Granshaw SI (2016): Photogrammetric Terminology: Third Edition. *The Photogrammetric Record*, 31(154): 210–252.

[407] Lemmens M (2011): Quality of Geoinformation. *Geo-information. Technologies, Applications and the Environment*, Springer Netherlands, pp. 211–227.

[408] Gröger G, Plümer L (2012): CityGML – Interoperable semantic 3D city models. *ISPRS Journal of Photogrammetry and Remote Sensing*, 71: 12–33.

[409] Kolbe TH, Gröger G, Plümer L (2005): CityGML: Interoperable Access to 3D City Models. P van Oosterom, S Zlatanova, E Fendel, eds., *Proceedings of the First International Symposium on Geo-Information for Disaster Management (GI4DM)*. Delft, the Netherlands, pp. 883–899.

[410] Kolbe TH, Nagel C, Stadler A (2009): CityGML–OGC Standard for Photogrammetry. D Fritsch, ed., *Proceedings of the 52nd Photogrammetric Week '09*. Stuttgart, Germany, pp. 265–277.

[411] Löwner MO, Gröger G, Benner J, Biljecki F, Nagel C (2016): Proposal for a new LOD and multi-representation concept for CityGML. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, IV-2/W1: 3–12.

[412] Kolbe TH (2009): Representing and exchanging 3D city models with CityGML. S Zlatanova, J Lee, eds., *3D Geo-Information Sciences*, Springer Berlin Heidelberg, pp. 15–31.

[413] Albert J, Bachmann M, Hellmeier A (2003): Zielgruppen und Anwendungen für Digitale Stadtmodelle und Digitale Geländemodell. Tech. rep., SIG3D.

[414] Gröger G, Kolbe TH, Drees R, Kohlhaas A, Müller H, Knospe F, Gruber U, Krause U (2004): Das interoperable 3D-Stadtmodell der SIG 3D der GDI NRW. Tech. rep., Initiative Geodaten Infrastruktur NRW.

[415] Gröger G, Benner J, Dörschlag D, Drees R, Gruber U, Leinemann K, Löwner MO (2005): Das interoperable 3D-Stadtmodell der SIG 3D. *Zeitschrift für Geodäsie, Geoinformation und Landmanagement*, 130(6): 343–353.

[416] Kolbe TH, Burger B, Cantzler B (2015): CityGML goes to Broadway. *Photogrammetric Week '15*. Stuttgart, Germany, pp. 343–356.

[417] Stoter J, Ledoux H, Reuvers M, van den Brink L, Klooster R, Janssen P, Beetz J, Penninga F, Vosselman G (2013): Establishing and implementing a national 3D Standard in The Netherlands. *Photogrammetrie - Fernerkundung - Geoinformation*, 2013(4): 381–392.

[418] Chinese Ministry of Housing and Urban-Rural Development (2010): *Technical specification for three dimensional city modeling*. PRC Industry Standard 82765. China Building Industry Press.

[419] Thiemann F (2003): 3D-Gebäude-Generalisierung. WG Koch, ed., *Theorie 2003 - Vorträge der Dresdner Sommerschule für Kartographie*. Dresden, Germany, pp. 52–61.

[420] Schilcher M, Roschlaub R, Guo Z (1998): Vom 2D-GIS zum 3D-Stadtmodell durch Kombination von GIS-, CAD- und Animationstechniken. *Tagungsband ACS '98, Fachseminar Geoinformationssysteme*. Frankfurt, Germany, p. 12.

[421] Blom ASA (2011): Blom3D™ Whitepaper for Blom partners, clients and developers. Tech. rep., Oslo, Norway.

[422] Vertex Modelling (2013): Wide Area Models. URL http://www.vertexmodelling.co.uk/site/about/wide-area-models/.

[423] NAVTEQ (2011): *3D Landmarks - A product guide for developers*. NAVTEQ.

[424] CyberCity 3D (2013): 3D Modeling: 3D city and area models. URL http://www.cybercity3d.com/index.php?option=com_content&view=article&id=97&Itemid=13.

[425] Sanborn (2014): 3D Cities™. Sanborn. URL http://www.sanborn.com/sites/default/files/files/3DCitiesv6_0.pdf.

[426] Vande Velde L (2005): Tele Atlas 3D navigable maps. G Gröger, TH Kolbe, eds., *EuroSDR International Workshop on Next Generation 3D City Models*. pp. 47–50.

[427] Glasgow City Council (2009): Urban model Specification. Glasgow City Council. Development and Regeneration Services. Spec/09/01/v001. URL http://www.glasgow.gov.uk/CHttpHandler.ashx?id=2078&p=0.

[428] Hochtief ViCon (2011): Lusail City. 3D Model Specifications High Rise Buildings. Tech. rep., Lusail City, Qatar. URL http://www.lusail.com/English/The-Visitor-Center/lusailstarterpack/Documents/Section%205/Digital%203D%20Model%20Specifications.pdf.

[429] City of Wollongong (2010): City Centre 3D Model Specifications. Document Z10/56243. URL http://www.wollongong.nsw.gov.au/customerserviceonline/factsheet/Factsheets/City%20Centre%203D%20Model%20Specifications.pdf.

[430] Chen M (2013): *Comparison of 3D Tree Parameters*. Master's thesis, Wageningen University and Research Centre, Wageningen, the Netherlands.

[431] Rafiee A, Dias E, Koomen E (2013): Between Green and Grey: Towards a New Green Volume Indicator for Cities. S Geertman, J Stillwell, F Toppen, eds., *Proceedings of the 13th International Conference on Computers in Urban Planning and Urban Management (CUPUM)*. Utrecht, the Netherlands, pp. 1–18.

[432] Tamminga G, van den Brink L, van Lint H, Stoter J, Hoogendoorn S (2013): Towards GIS compliant data structures for traffic and transportation models. *Transportation Research Board 92nd Annual Meeting 2013. Transportation Data Interoperability: Recent Research (Session 283)*. Washington, D.C., USA, p. 18.

[433] Fan H, Meng L (2012): A three-step approach of simplifying 3D buildings modeled by CityGML. *International Journal of Geographical Information Science*, 26(6): 1091–1107.

[434] Benner J, Geiger A, Gröger G, Häfele KH, Löwner MO (2013): Enhanced LOD concepts for virtual 3D city models. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, II-2/W1: 51–61.

[435] Bandrova T, Bonchev S (2013): 3D Maps – Scale, Accuracy, Level of Detail. *Proceedings of the 26th International Cartographic Conference*. Dresden, Germany, p. 16.

[436] Löwner MO, Benner J, Gröger G, Häfele KH (2013): New Concepts for Structuring 3D City Models – An Extended Level of Detail Concept for CityGML Buildings. *Spatial Information Theory. Cognitive and Computational Foundations of Geographic Information Science*, Springer Berlin Heidelberg, pp. 466–480.

[437] Garland M, Heckbert PS (1997): Surface Simplification Using Quadric Error Metrics. *Proceedings of the 24th annual conference on Computer graphics and interactive techniques SIGGRAPH 97*. ACM Press/Addison-Wesley Publishing Co., Los Angeles, CA, United States, pp. 209–216.

[438] Remondino F, El-Hakim S (2006): Image-based 3D Modelling: A Review. *The Photogrammetric Record*, 21(115): 269–291.

[439] Schilling A, Bolling J, Nagel C (2016): Using glTF for streaming CityGML 3D city models. *Web3D '16*. ACM Press, Anaheim, CA, USA, pp. 109–116.

[440] Petit M (2016): Requirement templates for distributed 3D graphics. *Web3D '16*. ACM Press, Anaheim, CA, USA, pp. 61–68.

[441] Polys NF, Singh A, Sforza P (2016): A novel level-of-detail technique for virtual city environments. *Web3D '16*. ACM Press, Anaheim, CA, USA, pp. 183–184.

[442] Çöltekin A (2006): *Foveation for 3D visualization and stereo imaging*. Ph.D. thesis, Helsinki University of Technology, Institute of Photogrammetry and Remote Sensing, Helsinki, Finland.

[443] Clark JH (1976): Hierarchical Geometric Models for Visible Surface Algorithms. *Communications of the ACM*, 19(10): 547–554.

[444] Häfele KH (2011): CityGML Model of the FJK-Haus. Institut für Angewandte Informatik (IAI), Karlsruher Institut für Technologie.

[445] Cignoni P, Corsini M, Ranzuglia G (2008): MeshLab: an Open-Source 3D Mesh Processing System. *Ercim news*, 73: 45–46.

[446] Erikson C, Manocha D, Baxter III WV (2001): HLODs for faster display of large static and dynamic environments. JF Hughes, CH Séquin, eds., *Proceedings of the I3D '01 ACM Symposium on Interactive 3D Graphics*. ACM Press, Chapel Hill, NC, USA, pp. 111–120.

[447] Reddy M (1997): *Perceptually Modulated Level of Detail for Virtual Environments*. Ph.D. thesis, University of Edinburgh.

[448] Hausdorff F (1914): *Grundzüge der Mengenlehre*. Verlag von Veit & Comp., Leipzig, Germany.

[449] Garland M (1999): *Quadric-Based Polygonal Surface Simplification*. Ph.D. thesis, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, United States.

[450] Pasman W, Jansen FW (2002): Scheduling level of detail with guaranteed quality and cost. MG Wagner, KS Candan, MT Beitler, eds., *Proceedings of the 7th international conference on 3D Web technology*. ACM, Tempe, AZ, USA, pp. 43–51.

[451] Constantinescu Z (2000): Levels of Detail: An Overview. *Journal of the Lithuanian Association of Nonlinear Analysts*, 5(1): 39–52.

[452] Funkhouser TA, Séquin CH (1993): Adaptive display algorithm for interactive frame rates during visualization of complex virtual environments. *Proceedings of the 20th annual conference on Computer graphics and interactive techniques (SIGGRAPH 93)*. ACM, Anaheim, CA, USA, pp. 247–254.

[453] Coors V, Flick S (1998): Integrating Levels of Detail in a Web-based 3D-GIS. R Laurini, K Makki, N Pissinou, eds., *Proceedings of the 6th ACM international symposium on Advances in Geographic Information Systems*. ACM, Washington, DC, USA, pp. 40–45.

[454] Danovaro E, De Floriani L, Magillo P, Puppo E, Sobrero D (2006): Level-of-detail for data analysis and exploration: a historical overview and some new perspectives. *Computers and Graphics*, 30(3): 334–344.

[455] Kada M, Wichmann A, Hermes T (2015): Smooth transformations between generalized 3D building models for visualization purposes. *Cartography and Geographic Information Science*, 42(4): 306–314.

[456] Xie J, Zhang L, Li J, Wang H, Yang L (2012): Automatic simplification and visualization of 3D urban building models. *International Journal of Applied Earth Observation and Geoinformation*, 18: 222–231.

[457] Kada M, Wichmann A, Filippovska Y, Hermes T (2016): Animation strategies for smooth transformations between discrete LODs of 3D building models. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLI-B2: 413–420.

[458] De Berg M, Dobrindt KTG (1998): On levels of detail in terrains. *Graphical Models and Image Processing*, 60(1): 1–12.

[459] Suárez JP, Trujillo A, Santana JM, de la Calle M, Gómez-Deck D (2015): An efficient terrain Level of Detail implementation for mobile devices and performance study. *Computers, Environment and Urban Systems*, 52: 21–33.

[460] Rossignac J (2005): Shape complexity. *The Visual Computer*, 21(12): 985–996.

[461] Mandelbrot B (1967): How Long Is the Coast of Britain? Statistical Self-Similarity and Fractional Dimension. *Science*, 156(3775): 636–638.

[462] Frédéricque B, Daniel S, Bédard Y, Paparoditis N (2005): Knowledge-based process management to populate databases with 3D multi representation of buildings. C Heipke, K Jacobsen, M Gerke, eds., *ISPRS Hannover Workshop 2005: High-Resolution Earth Imaging for Geospatial Information*. Hannover, Germany, p. 6.

[463] Rua H, Alvito P (2011): Living the past: 3D models, virtual reality and game engines as tools for supporting archaeology and the reconstruction of cultural heritage – the case-study of the Roman villa of Casal de Freiria. *Journal of Archaeological Science*, 38(12): 3296–3308.

[464] Billen R, Cutting-Decelle AF, Métral C, Falquet G, Zlatanova S, Marina O (2015): Challenges of Semantic 3D City Models:. *International Journal of 3-D Information Modeling*, 4(2): 68–76.

[465] Stadler A, Kolbe TH (2007): Spatio-semantic coherence in the integration of 3D city models. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XXXVI-2/C43: 8.

[466] Malinverni ES, Tassetti AN (2013): GIS-based smart cartography using 3D modeling. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XL-2-W2: 47–52.

[467] Heuvelink GBM (1998): Uncertainty analysis in environmental modelling under a change of spatial scale. *Nutrient Cycling in Agroecosystems*, 50(1): 255–264.

[468] Chrisman NR (1991): The error component in spatial data. PA Longley, MF Goodchild, DJ Maguire, DW Rhind, eds., *Geographical Information Systems*, John Wiley & Sons, New York, pp. 165–174.

[469] Oude Elberink S, Vosselman G (2011): Quality analysis on 3D building models reconstructed from airborne laser scanning data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66(2): 157–165.

[470] Chaudhry OZ, Mackaness WA, Regnauld N (2009): A functional perspective on map generalisation. *Computers, Environment and Urban Systems*, 33(5): 349–362.

[471] Šuba R, Meijers M, Oosterom P (2016): Continuous Road Network Generalization throughout All Scales. *ISPRS International Journal of Geo-Information*, 5(8): 145–21.

[472] Luo Y, He J, He Y (2017): A rule-based city modeling method for supporting district protective planning. *Sustainable Cities and Society*, 28: 277–286.

[473] Goodchild MF (2001): Metrics of scale in remote sensing and GIS. *International Journal of Applied Earth Observation and Geoinformation*, 3(2): 114–120.

[474] Weibel R (1997): Generalization of spatial data: Principles and selected algorithms. *Algorithmic Foundations of Geographic Information Systems*, Springer, Berlin, Heidelberg, pp. 99–152.

[475] Montello DR, Golledge RG (1998): Scale and Detail in the Cognition of Geographic Information. Tech. rep., Santa Barbara, CA, USA.

[476] Mark DM (2003): Geographic information science: Defining the field. M Duckham, MF Goodchild, MF Worboys, eds., *Foundations of Geographic Information Science*, New York, NY, pp. 1–15.

[477] Jiang B, Brandt S (2016): A Fractal Perspective on Scale in Geography. *ISPRS International Journal of Geo-Information*, 5(6): 95.

[478] Aronoff S (1989): *Geographic Information Systems: A Management Perspective*. WDL Publications, Ottawa, Canada.

[479] van Oort P (2006): *Spatial data quality: from description to application*. Ph.D. thesis, Wageningen Universiteit, Wageningen, the Netherlands.

[480] Meroney RN, Leitl BM, Rafailidis S, Schatzmann M (1999): Wind-tunnel and numerical modeling of flow and dispersion about several building shapes. *Journal of Wind Engineering and Industrial Aerodynamics*, 81(1-3): 333–345.

[481] Picaut J, Simon L (2001): A scale model experiment for the study of sound propagation in urban areas. *Applied Acoustics*, 62(3): 327–340.

[482] Ismail MR, Oldham DJ (2005): A scale model investigation of sound reflection from building façades. *Applied Acoustics*, 66(2): 123–147.

[483] Coors V (2001): 3D-GIS in Networking Environments. *Proceedings of the International Workshop on 3D Cadastres, Registration of properties in strata*. Delft, the Netherlands, pp. 159–168.

[484] Kada M (2007): Scale-Dependent Simplification of 3D Building Models Based on Cell Decomposition and Primitive Instancing. *Spatial Information Theory*, Springer Berlin Heidelberg, pp. 222–237.

[485] Hagedorn B, Trapp M, Glander T, Döllner J (2009): Towards an Indoor Level-of-Detail Model for Route Visualization. *10th International Conference on Mobile Data Management: Systems, Services and Middleware*. IEEE, Taipei, Taiwan, pp. 692–697.

[486] Kemec S, Zlatanova S, Duzgun S (2012): A new LoD definition hierarchy for 3D city models used for natural disaster risk communication tool. T Bandrova, C König, eds., *Eurographics Workshop on Urban Data Modelling and Visualisation 2016*. Albena, Bulgaria, pp. 95–104.

[487] Morton PJ, Horne M, Dalton RC, Thompson EM (2012): Virtual City Models: Avoidance of Obsolescence. *Digital Physicality - Proceedings of the 30th eCAADe Conference*. Prague, Czech Republic, pp. 213–224.

[488] Li KJ, Lee J (2010): Indoor Spatial Awareness Initiative and Standard for Indoor Spatial Data. *Proceedings of the IROS 2010 Workshop on Standardization for Service Robots: Current Status and Future Directions*. Taipei, Taiwan, p. 5.

[489] ISO (2004): ISO 6707-1:2004(E): Building and civil engineering – Vocabulary. International Organisation for Standardisation.

[490] Zwolinski A, Jarzemski M (2015): Computing and monitoring potential of public spaces by shading analysis using 3D lidar data and advanced image analysis. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XL-7-W3: 743–750.

[491] Tempfli K, Pilouk M (1996): Practicable photogrammetry for 3D-GIS. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XXXI-B4: 859–867.

[492] Döllner J (2005): Continuous level-of-detail modeling of buildings in 3D city models. C Shahabi, O Boucelma, eds., *GIS '05 Proceedings of the 13th annual ACM international workshop on Geographic information systems*. Bremen, Germany, pp. 173–181.

[493] Deakin M, Campbell F, Reid A, Orsinger J (2014): The Urban Context. *The Mass Retrofitting of an Energy Efficient—Low Carbon Zone*, Springer London, London, pp. 23–38.

[494] Billger M, Thuvander L, Wästberg BS (2016): In search of visualization challenges: The development and implementation of visualization tools for supporting dialogue in urban planning processes. *Environment and Planning B: Planning and Design*, Advance online publication.

[495] Rautenbach V, Çöltekin A, Coetzee S (2015): Exploring the impact of visual complexity levels in 3D city models on the accuracy of individuals' orientation and cognitive maps. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, II-3/W5: 499–506.

[496] Goodchild MF (2001): Models of Scale and Scales of Modelling. NJ Tate, PM Atkinson, eds., *Modelling Scale in Geographical Information Science*, Wiley, Chichester, UK, pp. 3–10.

[497] Touya G, Reimer A (2015): Inferring the Scale of OpenStreetMap Features. JJ Arsanjani, A Zipf, P Mooney, M Helbich, eds., *OpenStreetMap in GIScience*, Springer International Publishing, pp. 81–99.

[498] Wong K, Ellul C (2016): Using geometry-based metrics as part of fitness-for-purpose evaluations of 3D city models. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, IV-2-W1: 129–136.

[499] Löwner MO, Gröger G (2016): Evaluation Criteria for Recent LoD Proposals for City-GML Buildings. *Photogrammetrie - Fernerkundung - Geoinformation*, 2016(1): 31–43.

[500] Biljecki F, Ledoux H, Stoter J (2016): An improved LOD specification for 3D building models. *Computers, Environment and Urban Systems*, 59: 25–37.

[501] Verdie Y, Lafarge F, Alliez P (2015): LOD Generation for Urban Scenes. *ACM Transactions on Graphics*, 34(3): 1–14.

[502] Musialski P, Wonka P, Aliaga DG, Wimmer M, van Gool L, Purgathofer W (2013): A Survey of Urban Reconstruction. *Computer Graphics Forum*, 32(6): 146–177.

[503] Tolmer CE, Castaing C, Diab Y, Morand D (2013): CityGML and IFC: Going further than LOD. *Digital Heritage International Congress DigitalHeritage*. IEEE, Marseille, France, pp. 645–648.

[504] Stoter J, Vosselman G, Goos J, Zlatanova S, Verbree E, Klooster R, Reuvers M (2011): Towards a National 3D Spatial Data Infrastructure: Case of The Netherlands. *Photogrammetrie - Fernerkundung - Geoinformation*, 2011(6): 405–420.

[505] Noskov A, Doytsher Y (2014): Preparing Simplified 3D Scenes of Multiple LODs of Buildings in Urban Areas Based on a Raster Approach and Information Theory. *Thematic Cartography for the Society*, Springer, pp. 221–236.

[506] Deng Y, Cheng JCP, Anumba C (2016): Mapping between BIM and 3D GIS in different levels of detail using schema mediation and instance comparison. *Automation in Construction*, 67: 1–21.

[507] Stoter J, Roensdorf C, Home R, Capstick D, Streilein A, Kellenberger T, Bayers E, Kane P, Dorsch J, Woźniak P, Lysell G, Lithen T, Bucher B, Paparoditis N, Ilves R (2015): 3D Modelling with National Coverage: Bridging the Gap Between Research and Practice. *Advances in 3D Geoinformation*, Springer International Publishing, Cham, pp. 207–225.

[508] Stoter J, Vallet B, Lithen T, Pla M, Woźniak P, Kellenberger T, Streilein A, Ilves R, Ledoux H (2016): State-of-the-art of 3D national mapping in 2016. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLI-B4: 653–660.

[509] Fan H, Zipf A, Fu Q, Neis P (2014): Quality assessment for building footprints data on OpenStreetMap. *International Journal of Geographical Information Science*, 28(4): 700–719.

[510] Gröger G, Coors V (2011): Modeling Guide for 3D Objects. Tech. rep., SIG3D Quality Working Group. URL http://files.sig3d.de/file/ag-qualitaet/20110531_Regeln_GML_final_EN.pdf.

[511] Wagner D, Wewetzer M, Bogdahn J, Alam N, Pries M, Coors V (2012): Geometric-Semantical Consistency Validation of CityGML Models. *Progress and New Trends in 3D Geoinformation Sciences*, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 171–192.

[512] Coors V, Wagner D (2015): CityGML Quality Interoperability Experiment des OGC. *DGPF Tagungsband. Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V.*, 24: 288–295.

[513] Stoter J, Vosselman G, Dahmen C, Oude Elberink S, Ledoux H (2014): CityGML Implementation Specifications for a Countrywide 3D Data Set: The Case of The Netherlands. *Photogrammetric Engineering and Remote Sensing*, 80(10): 13–21.

[514] He S, Besuievsky G, Tourre V, Patow G, Moreau G (2012): All range and heterogeneous multiscale 3D city models. T Leduc, G Moreau, R Billen, eds., *Usage, Usability, and Utility of 3D City Models – European COST Action TU0801*. EDP Sciences, Nantes, France.

[515] He S, Moreau G, Martin JY (2013): Footprint-Based Generalization of 3D Building Groups at Medium Level of Detail for Multi-Scale Urban Visualization. *International Journal On Advances in Software*, 5(3 and 4): 378–388.

[516] Borrmann A, Flurl M, Jubierre JR, Mundani RP, Rank E (2014): Synchronous collaborative tunnel design based on consistency-preserving multi-scale models. *Advanced Engineering Informatics*, 28(4): 499–517.

[517] Breunig M, Hinz S, Menninghaus M, Vögtle T, Rank E, Schilcher M, Borrmann A, Mundani RP, Ji Y, Donaubauer A, Steuer H (2011): Towards Computer-Aided Collaborative Subway Track Planning in Multi-Scale 3D City and Building Models. *Proceedings of the Joint ISPRS Workshop on 3D City Modelling & Applications and the 6th 3D GeoInfo Conference*. Wuhan, China, pp. 1–9.

[518] van Berlo LAHM, Bomhof F (2014): Creating the Dutch National BIM Levels of Development. *2014 International Conference on Computing in Civil and Building Engineering*. American Society of Civil Engineers, Orlando, FL, United States, pp. 129–136.

[519] Tolmer CE, Castaing C, Morand D, Diab Y (2014): Structuration des informations pour les projets d'infrastructures. Proposition de niveaux complémentaires aux Level Of Detail et Level Of Development. *Proceedings of the Conférence SCAN'14*. Luxembourg, Luxembourg, pp. 1–10.

[520] Wate P, Srivastav SK, Saran S, Murthy YVNK (2013): Formulation of hierarchical framework for 3D-GIS data acquisition techniques in context of Level-of-Detail (LoD). *Proceedings of the 2013 IEEE Second International Conference on Image Information Processing (ICIIP-2013)*. Shimla, India, pp. 154–159.

[521] Vosselman G, Dijkman S (2001): 3D building model reconstruction from point clouds and ground plans. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XXXIV-3/W4: 37–44.

[522] Billen R, Zaki C, Servières M, Moreau G, Hallot P (2012): Developing an ontology of space: Application to 3D city modeling. T Leduc, G Moreau, R Billen, eds., *Usage, Usability, and Utility of 3D City Models – European COST Action TU0801*. EDP Sciences, Nantes, France.

[523] Kang HY, Lee J (2014): A Study on the LOD(Level of Detail) Model for Applications based on Indoor Space Data. *Journal of the Korean Society of Surveying, Geodesy, Photogrammetry and Cartography*, 32(2): 143–151.

[524] Habib AF, Zhai R, KIM C (2010): Generation of Complex Polyhedral Building Models by Integrating Stereo-Aerial Imagery and Lidar Data. *Photogrammetric Engineering and Remote Sensing*, 76(5): 609–623.

[525] Xiong B, Jancosek M, Oude Elberink S, Vosselman G (2015): Flexible building primitives for 3D building modeling. *ISPRS Journal of Photogrammetry and Remote Sensing*, 101: 275–290.

[526] AdV (2011): Produktstandard für 3D-Gebäudemodelle. Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland (Working Committee of the Surveying Authorities of the States of the Federal Republic of Germany). Available online at http://www.adv-online.de, last accessed on 12 February 2016.

[527] AdV (2013): Modellierungsbeispiele für 3D-Gebäudemodelle. Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland (Working Committee of the Surveying Authorities of the States of the Federal Republic of Germany). Available online at http://www.adv-online.de, last accessed on 12 February 2016.

[528] Blaauboer J, Goos J, Ledoux H, Penninga F, Reuvers M, Stoter J, Vosselman G (2013): Technical specifications for the reconstruction of 3D IMGeo CityGML data. Tech. rep., Apeldoorn, the Netherlands.

[529] Zhao J, Zhu Q, Du Z, Feng T, Zhang Y (2012): Mathematical morphology-based generalization of complex 3D building models incorporating semantic relationships. *ISPRS Journal of Photogrammetry and Remote Sensing*, 68: 95–111.

[530] Li Q, Sun X, Yang B, Jiang S (2013): Geometric structure simplification of 3D building models. *ISPRS Journal of Photogrammetry and Remote Sensing*, 84: 100–113.

[531] Andujar C, Brunet P, Chica A, Navazo I (2010): Visualization of Large-Scale Urban Models through Multi-Level Relief Impostors. *Computer Graphics Forum*, 29(8): 2456–2468.

[532] Huang H, Brenner C, Sester M (2011): 3D building roof reconstruction from point clouds via generative models. D Agrawal, I Cruz, CS Jensen, E Ofek, E Tanin, eds., *Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*. ACM Press, Chicago, IL, United States, pp. 16–24.

[533] Verma V, Kumar R, Hsu S (2006): 3D Building Detection and Modeling from Aerial LIDAR Data. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)*. IEEE, pp. 2213–2220.

[534] Burochin JP, Vallet B, Brédif M, Mallet C, Brosset T, Paparoditis N (2014): Detecting blind building façades from highly overlapping wide angle aerial imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 96: 193–209.

[535] Becker S (2011): Towards Complete LOD3 Models – Automatic Interpretation of Building Structures. D Fritsch, ed., *Proceedings of the 53rd Photogrammetric Week '11*. pp. 39–56.

[536] Zhu Q, Hu MY (2010): Semantics-based 3D dynamic hierarchical house property model. *International Journal of Geographical Information Science*, 24(2): 165–188.

[537] Jarząbek-Rychard M, Borkowski A (2016): 3D building reconstruction from ALS data using unambiguous decomposition into elementary structures. *ISPRS Journal of Photogrammetry and Remote Sensing*, 118: 1–12.

[538] Arefi H, Reinartz P (2012): Multi-level building reconstruction for automatic enhancement of high resolution DSMs. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XXXIX-B2: 11–16.

[539] Tarantino E, Figorito B (2011): Extracting Buildings from True Color Stereo Aerial Images Using a Decision Making Strategy. *Remote Sensing*, 3(8): 1553–1567.

[540] McClune AP, Mills JP, Miller PE, Holland DA (2016): Automatic 3D building reconstruction from a dense image matching dataset. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLI-B3: 641–648.

[541] Municipality of Rotterdam (2016): Digital 3D Building Model Rotterdam: Descriptive Document (European public tendering procedure) on the realisation of a digital 3D building model of the city of Rotterdam (1-308-16).

[542] Ioannidis C, Verykokou S, Soile S, Potsiou C (2015): 5D Multi-Purpose Land Information System. *Eurographics Workshop on Urban Data Modelling and Visualisation*. The Eurographics Association, Delft, Netherlands, pp. 19–24.

[543] Frommholz D, Linkiewicz M, Meissner H, Dahlke D, Poznanska A (2015): Extracting semantically annotated 3D building models with textures from oblique aerial imagery. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XL-3/W2: 53–58.

[544] Kartverket (2014): SOSI Del 3 Produktspesifikasjon for Felles KartdataBase (FKB).

[545] Touya G, Brando-Escobar C (2013): Detecting Level-of-Detail Inconsistencies in Volunteered Geographic Information Data Sets. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 48(2): 134–143.

[546] Biljecki Z (2007): *Concept and Implementation of Croatian Topographic Information System—CROTIS*. Ph.D. thesis, Institut für Photogrammetrie und Fernerkundung, Technische Universität Wien.

[547] Guskov I, Brewington B (2015): US20150187130A1, United States Patent Application.

[548] Buyuksalih I, Isikdag U, Zlatanova S (2013): Exploring the processes of generating LOD (0-2) CityGML models in greater municipality of Istanbul. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XL-2/W2: 19–24.

[549] Over M, Schilling A, Neubauer S, Zipf A (2010): Generating web-based 3D City Models from OpenStreetMap: The current situation in Germany. *Computers, Environment and Urban Systems*, 34(6): 496–507.

[550] Ordnance Survey (2014): *OS MasterMap Topography Layer – Building Height Attribute. Getting started guide*, 1.0 edn.

[551] SwissTopo (2010): swissBUILDINGS3D 1.0. Vereinfachte 3D-Gebäude der Schweiz. Tech. rep.

[552] Sargent I, Holland D, Harding J (2015): The Building Blocks of User-Focused 3D City Models. *ISPRS International Journal of Geo-Information*, 4(4): 2890–2904.

[553] Baig SU, Abdul-Rahman A (2013): Generalization of buildings within the framework of CityGML. *Geo-spatial Information Science*, 16(4): 247–255.

[554] Diakité AA, Damiand G, Van Maercke D (2014): Topological Reconstruction of Complex 3D Buildings and Automatic Extraction of Levels of Detail. G Besuievsky, V Tourre, eds., *Eurographics Workshop on Urban Data Modelling and Visualisation*. Strasbourg, France, pp. 25–30.

[555] El-Mekawy M, Östman A, Shahzad K (2011): Towards Interoperating CityGML and IFC Building Models: A Unified Model Based Approach. TH Kolbe, G König, C Nagel, eds., *Advances in 3D Geoinformation*, Springer, Berlin, Heidelberg, pp. 73–93.

[556] van den Brink L, Stoter J, Zlatanova S (2013): Establishing a national standard for 3D topographic data compliant to CityGML. *International Journal of Geographical Information Science*, 27(1): 92–113.

[557] Jurelionis A, Bouris D (2016): Impact of Urban Morphology on Infiltration-Induced Building Energy Consumption. *Energies*, 9(3): 177.

[558] Petrescu F, Aldea M, Luca O, Iacoboaea C, Gaman F, Parlow E (2016): 3D geo-information in urban climate studies. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLII-2-W2: 51–55.

[559] Gesquière G, Manin A (2012): 3D Visualization of Urban Data Based on CityGML with WebGL. *International Journal of 3-D Information Modeling*, 1(3): 1–15.

[560] Peronato G, Bonjour S, Stoeckli J, Rey E, Andersen M (2016): Sensitivity of calculated solar irradiation to the level of detail: insights from the simulation of four sample buildings in urban areas. *Proceedings of the 32nd International Conference on Passive and Low Energy Architecture*. Los Angeles, CA, United States, pp. 685–690.

[561] Jaugsch F, Löwner MO (2016): Estimation of solar energy on vertical 3D building walls on city quarter scale. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLII-2-W2: 135–143.

[562] Ha J, Lee S, Park C (2016): Temporal Effects of Environmental Characteristics on Urban Air Temperature: The Influence of the Sky View Factor. *Sustainability*, 8(9): 895.

[563] Götzelmann T, Guercke R, Brenner C, Sester M (2009): Terrain-dependent aggregation of 3D city models. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XXXVIII-2/W11: 5.

[564] Mayer H (2005): Scale-spaces for generalization of 3D buildings. *International Journal of Geographical Information Science*, 19(8-9): 975–997.

[565] Agugiaro G (2014): From sub-optimal datasets to a CityGML-compliant 3D city model: experiences from Trento, Italy. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XL-4: 7–13.

[566] Biljecki F, Ledoux H, Du X, Stoter J, Soon KH, Khoo VHS (2016): The most common geometric and semantic errors in CityGML datasets. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, IV-2/W1: 13–22.

[567] Agugiaro G (2016): First steps towards an integrated CityGML-based 3D model of Vienna. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, III-4: 139–146.

[568] Yang L, Zhang L, Ma J, Xie J, Liu L (2011): Interactive visualization of multi-resolution urban building models considering spatial cognition. *International Journal of Geographical Information Science*, 25(1): 5–24.

[569] Prandi F, De Amicis R, PIffer S, Soave M, Cadzow S, Boix EG, D'Hondt E (2013): Using CityGML to deploy smart-city services for urban ecosystems. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XL-4/W1: 87–92.

[570] Ellul C, Altenbuchner J (2013): LOD 1 VS. LOD 2 – Preliminary investigations into differences in mobile rendering performance. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, II-2/W1: 129–138.

[571] Anders KH (2005): Level of detail generation of 3D building groups by aggregation and typification. *Proceedings of the 22nd International Cartographic Conference: Mapping approaches into a changing world*. La Coruña, Spain.

[572] Emem O, Batuk F (2004): Generating Precise and Accurate 3D City Models Using Photogrammetric Data . *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XXXV/B4: 431–436.

[573] Commandeur T (2012): *Footprint decomposition combined with point cloud segmentation for producing valid 3D models*. Master's thesis, Delft University of Technology.

[574] Oude Elberink S, Stoter J, Ledoux H, Commandeur T (2013): Generation and Dissemination of a National Virtual 3D City and Landscape Model for the Netherlands. *Photogrammetric Engineering and Remote Sensing*, 79(2): 147–158.

[575] Xiong B, Oude Elberink S, Vosselman G (2016): Footprint map partitioning using airborne laser scanning data. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, III-3: 241–247.

[576] Alexander C, Smith-Voysey S, Jarvis C, Tansey K (2009): Integrating building footprints and LiDAR elevation data to classify roof structures and visualise buildings. *Computers, Environment and Urban Systems*, 33(4): 285–292.

[577] Haala N, Brenner C (1999): Virtual city models from laser altimeter and 2D map data. *Photogrammetric Engineering and Remote Sensing*, 65(7): 787–795.

[578] Haala N, Kada M (2010): An update on automatic 3D building reconstruction. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65(6): 570–580.

[579] Mao B, Harrie L, Ban Y (2012): Detection and typification of linear structures for dynamic visualization of 3D city models. *Computers, Environment and Urban Systems*, 36(3): 233–244.

[580] Brenner C (2005): Building reconstruction from images and laser scanning. *International Journal of Applied Earth Observation and Geoinformation*, 6(3-4): 187–198.

[581] Kada M, McKinley L (2009): 3D building reconstruction from LiDAR based on a cell decomposition approach. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XXXVIII-3/W4: 47–52.

[582] SwissTopo (2014): Produktinformation swissBUILDINGS3D 2.0.

[583] Steinhage V, Behley J, Meisel S, Cremers AB (2010): Automated updating and maintenance of 3D city models. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XXXVIII-4-8-2/W9: 1–6.

[584] Oude Elberink S (2010): *Acquisition of 3D topography: automated 3D road and building reconstruction using airborne laser scanner data and topographic maps*. Ph.D. thesis, ITC, University of Twente, Enschede, the Netherlands.

[585] Bremer M, Mayr A, Wichmann V, Schmidtner K, Rutzinger M (2016): A new multi-scale 3D-GIS-approach for the assessment and dissemination of solar income of digital city models. *Computers, Environment and Urban Systems*, 57: 144–154.

[586] van den Brink L, Stoter J, Zlatanova S (2013): UML-Based Approach to Developing a CityGML Application Domain Extension. *Transactions in GIS*, 17(6): 920–942.

[587] Schilling A, Hagedorn B, Coors V (2012): Final report of the OGC 3D Portrayal Interoperability Experiment. Tech. Rep. 12-075.

[588] Schwalbe E, Maas HG, Seidel F (2005): 3D building model generation from airborne laser scanner data using 2D GIS data and orthogonal point cloud projections. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XXXVI-3/W19: 209–214.

[589] Förstner W (1999): 3D-City Models: Automatic and Semiautomatic Acquisition Methods. D Fritsch, RH Spiller, eds., *Proceedings of the 47th Photogrammetric Week '99*. Stuttgart, Germany, pp. 291–303.

[590] Henn A, Gröger G, Stroh V, Plümer L (2013): Model driven reconstruction of roofs from sparse LIDAR point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing*, 76: 17–29.

[591] Ben Fekih Fradj N, Löwner MO (2012): Abschätzung des nutzbaren Dachflächenanteils für Solarenergie mit CityGML-Gebäudemodellen und Luftbildern. MO Löwner, F Hillen, R Wohlfahrt, eds., *Geoinformatik 2012 "Mobilität und Umwelt". Konferenzband zur Tagung Geoinformatik*. Braunschweig, Germany, pp. 171–177.

[592] Garnett R, Freeburn JT (2014): Visual Acceptance of Library-Generated CityGML LOD3 Building Models. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 49(4): 218–224.

[593] El Meouche R, Rezoug M, Hijazi I, Maes D (2013): Automatic Reconstruction of 3D Building Models from Terrestrial Laser Scanner Data. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, II-4-W1: 7–12.

[594] Akmalia R, Setan H, Majid Z, Suwardhi D, Chong A (2014): TLS for generating multi-LOD of 3D building model. *IOP Conference Series: Earth and Environmental Science. Proceedings of the 8th International Symposium of the Digital Earth (ISDE8)*, 18: 1–8.

[595] Truong-Hong L, Laefer DF (2015): Quantitative evaluation strategies for urban 3D model generation from remote sensing data. *Computers and Graphics*, 49: 82–91.

[596] Donkers S, Ledoux H, Zhao J, Stoter J (2016): Automatic conversion of IFC datasets to geometrically and semantically correct CityGML LOD3 buildings. *Transactions in GIS*, 20(4): 547–569.

[597] Isikdag U, Zlatanova S (2009): Towards Defining a Framework for Automatic Generation of Buildings in CityGML Using Building Information Models. *3D Geo-Information Sciences*, Springer Berlin Heidelberg, pp. 79–96.

[598] de Laat R, van Berlo L (2011): Integration of BIM and GIS: The Development of the CityGML GeoBIM Extension. *Advances in 3D Geo-Information Sciences*, Springer Berlin Heidelberg, pp. 211–225.

[599] Yin X, Wonka P, Razdan A (2009): Generating 3D Building Models from Architectural Drawings: A Survey. *IEEE Computer Graphics and Applications*, 29(1): 20–30.

[600] Xiao J, Fang T, Zhao P, Lhuillier M, Quan L, Xiao J, Fang T, Zhao P, Lhuillier M, Quan L (2009): Image-based street-side city modeling. *ACM Transactions on Graphics*, 28(5): 114.

[601] Martinović A, Knopp J, Riemenschneider H, van Gool L (2015): 3D All The Way: Semantic Segmentation of Urban Scenes from Start to End in 3D. *CVPR 2015: The 28th IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, Boston, United States, pp. 4456–4465.

[602] Becker S (2009): Generation and application of rules for quality dependent façade reconstruction. *ISPRS Journal of Photogrammetry and Remote Sensing*, 64(6): 640–653.

[603] van Gool L, Martinović A (2013): Towards Semantic City Models. D Fritsch, ed., *Proceedings of the 54th Photogrammetric Week '13*. Stuttgart, Germany, pp. 217–232.

[604] Nguatem W, Drauschke M, Mayer H (2016): Automatic generation of building models with levels of detail 1–3. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLI-B3: 649–654.

[605] Göçer Ö, Hua Y, Göçer K (2016): A BIM-GIS integrated pre-retrofit model for building data mapping. *Building Simulation*, 9(5): 513–527.

[606] Biljecki F, Ledoux H, Stoter J, Vosselman G (2016): The variants of an LOD of a 3D building model and their influence on spatial analyses. *ISPRS Journal of Photogrammetry and Remote Sensing*, 116: 42–54.

[607] Biljecki F, Ledoux H, Stoter J (2014): Height references of CityGML LOD1 buildings and their influence on applications. M Breunig, AD Mulhim, E Butwilowski, PV Kuper, J Benner, KH Häfele, eds., *Proceedings of the 9th 3D GeoInfo Conference 2014*. Dubai, UAE.

[608] Tomljenovic I, Höfle B, Tiede D, Blaschke T (2015): Building Extraction from Airborne Laser Scanning Data: An Analysis of the State of the Art. *Remote Sensing*, 7(4): 3826–3862.

[609] Rottensteiner F (2003): Automatic generation of high-quality building models from lidar data. *IEEE Computer Graphics and Applications*, 23(6): 42–50.

[610] Sirmacek B, Taubenbock H, Reinartz P, Ehlers M (2012): Performance Evaluation for 3-D City Model Generation of Six Different DSMs From Air- and Spaceborne Sensors. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 5(1): 59–70.

[611] Hillsman EL, Rhoda R (1978): Errors in measuring distances from populations to service centers. *The Annals of Regional Science*, 12(3): 74–88.

[612] Miller HJ (1996): GIS and geometric representation in facility location problems. *International Journal of Geographical Information Systems*, 10(7): 791–816.

[613] Murray AT, O'Kelly ME (2002): Assessing representation error in point-based coverage modeling. *Journal of Geographical Systems*, 4(2): 171–191.

[614] Cromley RG, Lin J, Merwin DA (2012): Evaluating representation and scale error in the maximal covering location problem using GIS and intelligent areal interpolation. *International Journal of Geographical Information Science*, 26(3): 495–517.

[615] Pedrinis F, Gesquière G (2016): Reconstructing 3D Building Models with the 2D Cadastre for Semantic Enhancement. *Advances in 3D Geoinformation*, Springer International Publishing, Cham, pp. 119–135.

[616] Oude Elberink S (2008): Problems in automated building reconstruction based on dense airborne laser scanning data. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XXXVII: 93–98.

[617] INSPIRE Thematic Working Group Buildings (2013): D2.8.III.2 INSPIRE Data Specification on Buildings – Technical Guidelines. European Commission Joint Research Centre.

[618] Gröger G, Plümer L (2013): The Interoperable Building Model of the European Union. A Abdul-Rahman, P Boguslawski, F Anton, MN Said, KM Omar, eds., *Geoinformation for Informed Decisions*, Springer International Publishing, pp. 1–17.

[619] He Y, Zhang C, Awrangjeb M, Fraser CS (2012): Automated reconstruction of walls from airborne LiDAR data for complete 3D building modelling. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XXXIX-B3: 115–120.

[620] Hermosilla T, Ruiz LA, Recio JA, Cambra-López M (2012): Assessing contextual descriptive features for plot-based classification of urban areas. *Landscape and Urban Planning*, 106(1): 124–137.

[621] Zhang K, Yan J, Chen SC (2006): Automatic Construction of Building Footprints From Airborne LIDAR Data. *IEEE Transactions on Geoscience and Remote Sensing*, 44(9): 2523–2533.

[622] Brasebin M, Perret J, Mustiere S, Weber C (2016): A Generic Model to Exploit Urban Regulation Knowledge. *ISPRS International Journal of Geo-Information*, 5(2): 14.

[623] Eeftens M, Beekhuizen J, Beelen R, Wang M, Vermeulen R, Brunekreef B, Huss A, Hoek G (2013): Quantifying urban street configuration for improvements in air pollution models. *Atmospheric Environment*, 72: 1–9.

[624] Bruse M, Nouvel R, Wate P, Kraut V, Coors V (2015): An Energy-Related CityGML ADE and Its Application for Heating Demand Calculation. *International Journal of 3-D Information Modeling*, 4(3): 59–77.

[625] Ordnance Survey (2014): *OS MasterMap Topography Layer. User guide and technical specification*, 1.12 edn.

[626] Arefi H, Engels J, Hahn M, Mayer H (2008): Levels of detail in 3D building reconstruction from LiDAR data. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XXXVII-B3b: 485–490.

[627] Goetz M, Zipf A (2012): Towards Defining a Framework for the Automatic Derivation of 3D CityGML Models from Volunteered Geographic Information. *International Journal of 3-D Information Modeling*, 1(2): 1–16.

[628] ISO (2013): ISO 19157:2013 – Geographic information – Data quality.

[629] Biljecki F, Ledoux H, Stoter J (2016): Generation of multi-LOD 3D city models in CityGML with the procedural modelling engine Random3Dcity. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, IV-4/W1: 51–59.

[630] Xie J, Feng CC (2016): An Integrated Simplification Approach for 3D Buildings with Sloped and Flat Roofs. *ISPRS International Journal of Geo-Information*, 5(8): 128.

[631] Paroush J (1978): On Quality Discrimination. *Southern Economic Journal*, 45(2): 592.

[632] Johansson T, Segerstedt E, Olofsson T, Jakobsson M (2016): Revealing Social Values by 3D City Visualization in City Transformations. *Sustainability*, 8(2): 195.

[633] Ledoux H (2013): On the Validation of Solids Represented with the International Standards for Geographic Information. *Computer-Aided Civil and Infrastructure Engineering*, 28(9): 693–706.

[634] Martinović A (2015): *Inverse Procedural Modeling of Buildings*. Ph.D. thesis, KU Leuven.

[635] Wonka P, Wimmer M, Ribarsky W, Sillion F (2003): Instant architecture. *ACM Transactions on Graphics*, 22(3): 669–677.

[636] Kelly T, Wonka P (2011): Interactive architectural modeling with procedural extrusions. *ACM Transactions on Graphics*, 30(2): 1–15.

[637] Rodrigues R, Coelho A, Reis LP (2010): Data model for procedural modelling from textual descriptions. *IEEE Congress on Evolutionary Computation (CEC)*. IEEE, Barcelona, Spain, pp. 1–8.

[638] Beneš J, Wilkie A, Křivánek J (2014): Procedural Modelling of Urban Road Networks. *Computer Graphics Forum*, 33(6): 132–142.

[639] Lintermann B, Deussen O (1999): Interactive modeling of plants. *IEEE Computer Graphics and Applications*, 19(1): 56–65.

[640] Koehl M, Roussel F (2015): Procedural modelling for reconstruction of historic monuments. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, II-5-W3: 137–144.

[641] Vanegas CA, Kelly T, Weber B, Halatsch J, Aliaga DG, Müller P (2012): Procedural Generation of Parcels in Urban Modeling. *Computer Graphics Forum*, 31(2): 681–690.

[642] Tsiliakou E, Labropoulos T, Dimopoulou E (2014): Procedural Modeling in 3D GIS Environment. *International Journal of 3-D Information Modeling*, 3(3): 17–34.

[643] Besuievsky G, Patow G (2014): Recent Advances on LoD for Procedural Urban Models. *Proceedings of the Workshop on Processing Large Geospatial Data*. Cardiff, United Kingdom.

[644] Rautenbach V, Bevis Y, Coetzee S, Combrinck C (2015): Evaluating procedural modelling for 3D models of informal settlements in urban design activities. *South African Journal of Science*, 111(11/12): 10.

[645] Rautenbach V, Coetzee S, Çöltekin A (2016): Investigating the use of 3D geovisualizations for urban design in informal settlement upgrading in South Africa. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLI-B2: 425–431.

[646] Stouffs R, Janssen P (2017): A Rule-Based Generative Analysis Approach for Urban Planning. *Morphological Analysis of Cultural DNA*, Springer Singapore, pp. 125–136.

[647] Hermosilla T, Ruiz LA, Recio JA, Estornell J (2011): Evaluation of Automatic Building Detection Approaches Combining High Resolution Images and LiDAR Data. *Remote Sensing*, 3(6): 1188–1210.

[648] Zhang W, Wang H, Chen Y, Yan K, Chen M (2014): 3D Building Roof Modeling by Optimizing Primitive's Parameters Using Constraints from LiDAR Data and Aerial Imagery. *Remote Sensing*, 6(9): 8107–8133.

[649] Sinning-Meister M, Gruen A, Dan H (1996): 3D city models for CAAD-supported analysis and design of urban areas. *ISPRS Journal of Photogrammetry and Remote Sensing*, 51(4): 196–208.

[650] Hammoudi K, Dornaika F (2011): A Featureless Approach to 3D Polyhedral Building Modeling from Aerial Images. *Sensors*, 11(1): 228–259.

[651] ISO (2008): ISO/IEC 9834-8:2008(E) – Information technology — Open Systems Interconnection — Procedures for the operation of OSI Registration Authorities: Generation and registration of Universally Unique Identifiers (UUIDs) and their use as ASN.1 Object Identifier components.

[652] OGC (2012): OGC Geography Markup Language (GML) — Extended schemas and encoding rules 3.3.0. Open Geospatial Consortium.

[653] ISO (2003): 19107:2003(E) – Geographic information – Spatial schema.

[654] Zhao J, Ledoux H, Stoter J (2013): Automatic repair of CityGML LoD2 buildings using shrink-wrapping. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, II-2/W1: 309–317.

[655] OGC (2016): OGC CityGML Quality Interoperability Experiment. Tech. Rep. OGC 16-064.

[656] Doodman S, Afghantoloee A, Mostafavi MA, Karimipour F (2014): 3D extension of the vor algorithm to determine and optimize the coverage of geosensor networks. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XL-2/W3: 103–108.

[657] Steuer H, Machl T, Sindram M, Liebel L, Kolbe TH (2015): Voluminator—Approximating the Volume of 3D Buildings to Overcome Topological Errors. *AGILE 2015*, Springer International Publishing, pp. 343–362.

[658] Peters R (2017): *Point cloud modelling with the 3D Medial Axis Transform*. Ph.D. thesis, Delft University of Technology, In preparation.

[659] Neuville R, Poux F, Hallot P, Billen R (2016): Towards a normalised 3D geovisualisation: the viewpoint management. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, IV-2-W1: 179–186.

[660] Tsigkanos C, Kehrer T, Ghezzi C (2017): Modeling and Verification of Evolving Cyber-Physical Spaces. *Submitted to the 11th joint meeting of the European Software Engineering Conference and the ACM SIGSOFT Symposium on the Foundations of Software Engineering (under review)*. Paderborn, Germany, pp. 1–11.

[661] Becker S, Peter M, Fritsch D, Philipp D, Baier P, Dibak C (2013): Combined Grammar for the Modeling of Building Interiors. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, II-4/W1: 1–6.

[662] Peter M, Becker S, Fritsch D (2013): Grammar supported indoor mapping. *Proceedings of the 26th International Cartographic Conference*. Dresden, Germany, pp. 1–18.

[663] Gröger G, Plümer L (2010): Derivation of 3D Indoor Models by Grammars for Route Planning. *Photogrammetrie - Fernerkundung - Geoinformation*, 2010(3): 191–206.

[664] Ilčík M, Wimmer M (2013): Challenges and ideas in procedural modeling of interiors. *Eurographics Workshop on Urban Data Modelling and Visualisation*. Eurographics Association, Girona, Spain, pp. 29–30.

[665] Biljecki F, Ledoux H, Stoter J (2017): Generating 3D city models without elevation data. *Computers, Environment and Urban Systems*, 64: 1–18.

[666] Hecht R, Meinel G, Buchroithner M (2015): Automatic identification of building types based on topographic databases – a comparison of different data sources. *International Journal of Cartography*, 1(1): 18–31.

[667] Hartmann A, Meinel G, Hecht R, Behnisch M (2016): A Workflow for Automatic Quantification of Structure and Dynamic of the German Building Stock Using Official Spatial Data. *ISPRS International Journal of Geo-Information*, 5(8): 142.

[668] Wendel J, Murshed SM, Sriramulu A, Nichersu A (2016): Development of a Web-Browser Based Interface for 3D Data—A Case Study of a Plug-in Free Approach for Visualizing Energy Modelling Results. *Progress in Cartography*, Springer International Publishing, Cham, pp. 185–205.

[669] Farr TG, Rosen PA, Caro E, Crippen R, Duren R, Hensley S, Kobrick M, Paller M, Rodriguez E, Roth L, Seal D, Shaffer S, Shimada J, Umland J, Werner M, Oskin M, Burbank D, Alsdorf D (2007): The Shuttle Radar Topography Mission. *Reviews of Geophysics*, 45(2): RG2004.

[670] Sathish Kumar D, Arya DS, Vojinovic Z (2013): Modeling of urban growth dynamics and its impact on surface runoff characteristics. *Computers, Environment and Urban Systems*, 41: 124–135.

[671] Varga M, Bašić T (2015): Accuracy validation and comparison of global digital elevation models over Croatia. *International Journal of Remote Sensing*, 36(1): 170–189.

[672] Uden M, Zipf A (2013): Open Building Models: Towards a Platform for Crowdsourcing Virtual 3D Cities. *Progress and New Trends in 3D Geoinformation Sciences*, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 299–314.

[673] Noskov A, Doytsher Y (2013): Hierarchical Quarters Model Approach toward 3D Raster Based Generalization of Urban Environments. *International Journal On Advances in Software*, 6(3&4): 343–353.

[674] Fan H, Zipf A (2016): Modelling the world in 3D from VGI/Crowdsourced data. *European Handbook of Crowdsourced Geographic Information*, London, pp. 435–446.

[675] Coors V, Hünlich K, On G (2009): Constraint-based Generation and Visualization of 3D City Models. *3D Geo-Information Sciences*, Springer Berlin Heidelberg, pp. 365–378.

[676] Şengül A (2010): *Extracting Semantic Building Models From Aerial Stereo Images and Conversion to CityGML*. Master's thesis, Istanbul Technical University and Technical University Berlin, Istanbul, Turkey.

[677] Wróżyński R, Sojka M, Pyszny K (2016): The application of GIS and 3D graphic software to visual impact assessment of wind turbines. *Renewable Energy*, 96: 625–635.

[678] Ratti C, Richens P (2004): Raster Analysis of Urban Form. *Environment and Planning B: Planning and Design*, 31(2): 297–309.

[679] Schläpfer M, Lee J, Bettencourt LM (2016): Urban Skylines: Building Heights and Shapes as Measures of City Size. Working paper, Santa Fe Institute.

[680] Santana JM, Wendel J, Trujillo A, Suárez JP, Simons A, Koch A (2017): Multimodal Location Based Services—Semantic 3D City Data as Virtual and Augmented Reality. *Progress in Location-Based Services 2016*, Springer International Publishing, Cham, pp. 329–353.

[681] Goetz M, Zipf A (2012): OpenStreetMap in 3D – Detailed insights on the current situation in Germany. *Proceedings of the AGILE'2012 International Conference on Geographic Information Science*. Avignon, April, pp. 288–292.

[682] Olivares García JM, Virgós Soriano LI, Velasco Martín-Varés A (2011): 3D Modeling and Representation of the Spanish Cadastral Cartography. *Proceedings of the 2nd International Workshop on 3D Cadastres*. Delft, the Netherlands, pp. 209–222.

[683] Chau KW, Wong SK, Yau Y, Yeung AKC (2007): Determining Optimal Building Height. *Urban Studies*, 44(3): 591–607.

[684] Dalmau F, Garcia-Almirall P, Domínguez E, Escudero D (2014): From Raw Data to Meaningful Information: A Representational Approach to Cadastral Databases in Relation to Urban Planning. *Future Internet*, 6(4): 612–639.

[685] Basiouka S, Potsiou C, Bakogiannis E (2015): OpenStreetMap for cadastral purposes: an application using VGI for official processes in urban areas. *Survey Review*, 47(344): 333–341.

[686] Figueiredo R, Martina M (2016): Using open building data in the development of exposure data sets for catastrophe risk modelling. *Natural Hazards and Earth System Sciences*, 16(2): 417–429.

[687] Bertaud A, Brueckner JK (2005): Analyzing building-height restrictions: predicted impacts and welfare costs. *Regional Science and Urban Economics*, 35(2): 109–125.

[688] Joshi KK, Kono T (2009): Optimization of floor area ratio regulation in a growing city. *Regional Science and Urban Economics*, 39(4): 502–511.

[689] Brasebin M, Perret J, Haëck C (2012): Towards a 3D geographic information system for the exploration of urban rules. *Urban and Regional Data Management*, CRC Press, pp. 37–49.

[690] Allani-Bouhoula N, Perrin JP (2008): Architectural rules for three-dimensional reconstruction. *CGIM '08 Proceedings of the Tenth IASTED International Conference on Computer Graphics and Imaging*. Innsbruck, Austria, pp. 237–242.

[691] Irvin RB, McKeown DM (1989): Methods for Exploiting the Relationship Between Buildings and Their Shadows in Aerial Imagery. EB Barrett, JJ Pearson, eds., *Proceedings of SPIE 1076 - Image Understanding and the Man-Machine Interface II*. International Society for Optics and Photonics, pp. 156–164.

[692] Lin C, Huertas A, Nevatia R (1994): Detection of buildings using perceptual grouping and shadows. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, Seattle, WA, pp. 62–69.

[693] Nevatia R, Lin C, Huertas A (1997): A System for Building Detection from Aerial Images. *Automatic Extraction of Man-Made Objects from Aerial and Space Images (II)*, Birkhäuser Basel, Basel, pp. 77–86.

[694] Sırmaçek B, Ünsalan C (2008): Building detection from aerial images using invariant color features and shadow information. *23rd International Symposium on Computer and Information Sciences (ISCIS)*. IEEE, Istanbul, Turkey, pp. 1–5.

[695] Peeters A, Etzion Y (2012): Automated recognition of urban objects for morphological urban analysis. *Computers, Environment and Urban Systems*, 36(6): 573–582.

[696] Lee T, Kim T (2013): Automatic building height extraction by volumetric shadow analysis of monoscopic imagery. *International Journal of Remote Sensing*, 34(16): 5834–5850.

[697] Cheng F, Thiel KH (1995): Delimiting the building heights in a city from the shadow in a panchromatic SPOT-image—Part 1. Test of forty-two buildings. *Remote Sensing*, 16(3): 409–415.

[698] Shao Y, Taff GN, Walsh SJ (2011): Shadow detection and building-height estimation using IKONOS data. *International Journal of Remote Sensing*, 32(22): 6929–6944.

[699] Liasis G, Stavrou S (2016): Satellite images analysis for shadow detection and building height estimation. *ISPRS Journal of Photogrammetry and Remote Sensing*, 119: 437–450.

[700] Tong X, Lin X, Feng T, Xie H, Liu S, Hong Z, Chen P (2013): Use of shadows for detection of earthquake-induced collapsed buildings in high-resolution satellite imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 79: 53–67.

[701] Tu J, Sui H, Feng W, Song Z (2016): Automatic building damage detection method using high-resolution remote sensing images and 3D GIS model. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, III-8: 43–50.

[702] Breiman L (2001): Random Forests. *Machine Learning*, 45(1): 5–32.

[703] Mutanga O, Adam E, Cho MA (2012): High density biomass estimation for wetland vegetation using WorldView-2 imagery and random forest regression algorithm. *International Journal of Applied Earth Observation and Geoinformation*, 18: 399–406.

[704] Zhou X, Yu W, Sullivan WC (2016): Making pervasive sensing possible: Effective travel mode sensing based on smartphones. *Computers, Environment and Urban Systems*, 58: 52–59.

[705] Conedera M, Tonini M, Oleggini L, Vega Orozco C, Leuenberger M, Pezzatti GB (2015): Geospatial approach for defining the Wildland-Urban Interface in the Alpine environment. *Computers, Environment and Urban Systems*, 52: 10–20.

[706] Grömping U (2009): Variable Importance Assessment in Regression: Linear Regression versus Random Forest. *The American Statistician*, 63(4): 308–319.

[707] Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M, Duchesnay É (2011): Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12(Oct): 2825–2830.

[708] VROM (2008): Geometrie en de BAG (Verdiepingsdocument). Dutch Ministry of Housing, Spatial Planning and the Environment. Version 1.9. URL http://www.kadaster.nl/web/artikel/download/BAG-verdiepingsdocument-geometrie-1.htm.

[709] Roschlaub R, Batscheider J (2016): An INSPIRE-konform 3D building model of Bavaria using cadastre information, lidar and image matching. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLI-B4: 747–754.

[710] Salimzadeh N, Sharif SA, Hammad A (2016): Visualizing and Analyzing Urban Energy Consumption: A Critical Review and Case Study. *Construction Research Congress 2016*. American Society of Civil Engineers, Reston, VA, pp. 1323–1331.

[711] Lwin KK, Sugiura K, Zettsu K (2016): Space–time multiple regression model for grid-based population estimation in urban areas. *International Journal of Geographical Information Science*, 30(8): 1579–1593.

[712] Neidhart H, Sester M (2004): Identifying building types and building clusters using 3-D laser scanning and GIS-data. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XXXV/B4: 715–720.

[713] Meinel G, Hecht R, Herold H (2009): Analyzing building stock using topographic maps and GIS. *Building Research & Information*, 37(5-6): 468–482.

[714] Vartanian O, Navarrete G, Chatterjee A, Fich LB, Gonzalez-Mora JL, Leder H, Modroño C, Nadal M, Rostrup N, Skov M (2015): Architectural design and the brain: Effects of ceiling height and perceived enclosure on beauty judgments and approach-avoidance decisions. *Journal of Environmental Psychology*, 41: 10–18.

[715] Alahmadi M, Atkinson P, Martin D (2013): Estimating the spatial distribution of the population of Riyadh, Saudi Arabia using remotely sensed built land cover and height data. *Computers, Environment and Urban Systems*, 41: 167–176.

[716] Bast H, Storandt S, Weidner S (2015): Fine-grained population estimation. *Proceedings of the 23rd SIGSPATIAL International Conference on Advances in Geographic Information Systems*. ACM, New York, NY, USA, pp. 17–10.

[717] NEN (2007): Nederlandse norm NEN 2580:2007 - Oppervlakten en inhouden van gebouwen - Termen, definities en bepalingsmethoden (Netherlands standard on Areas and volumes of buildings - Terms, definitions and methods of determination). Nederlands Normalisatie-instituut (Netherlands Standardization Institute).

[718] Dehbi Y, Gröger G, Plümer L (2016): Identification and Modelling of Translational and Axial Symmetries and their Hierarchical Structures in Building Footprints by Formal Grammars. *Transactions in GIS*, 20(5): 645–663.

[719] Belgiu M, Tomljenovic I, Lampoltshammer T, Blaschke T, Höfle B (2014): Ontology-Based Classification of Building Types Detected from Airborne Laser Scanning Data. *Remote Sensing*, 6(2): 1347–1366.

[720] Angel S, Parent J, Civco DL (2010): Ten compactness properties of circles: measuring shape in geography. *The Canadian Geographer*, 54(4): 441–461.

[721] Hecht R, Kunze C, Hahmann S (2013): Measuring Completeness of Building Footprints in OpenStreetMap over Space and Time. *ISPRS International Journal of Geo-Information*, 2(4): 1066–1091.

[722] Hintze JL, Nelson RD (1998): Violin Plots: A Box Plot-Density Trace Synergism. *The American Statistician*, 52(2): 181–184.

[723] Chai T, Draxler RR (2014): Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, 7(3): 1247–1250.

[724] Zandbergen PA (2008): Positional Accuracy of Spatial Data: Non-Normal Distributions and a Critique of the National Standard for Spatial Data Accuracy. *Transactions in GIS*, 12(1): 103–130.

[725] Macay Moreia JM, Nex F, Agugiaro G, Remondino F, Lim NJ (2013): From DSM to 3D building models: a quantitative evaluation. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XL-1/W1: 213–219.

[726] Gupta K, Bhardwaj A, Kumar P, Pushpalata P (2015): Procedural rule based 3D city modeling and visualization using high resolution satellite data. *International Journal of Advancement in Remote Sensing, GIS and Geography*, 3(2): 16–25.

[727] Al Amouri A, Kolbe TH (2009): Quality assessment of historical Baalbek's 3D city model. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XXXVIII-2/W11: 5.

[728] van der Hoeven F, Wandl A (2015): Amsterwarm: Mapping the landuse, health and energy-efficiency implications of the Amsterdam urban heat island. *Building Services Engineering Research and Technology*, 36(1): 67–88.

[729] Helbich M, van Emmichoven MJZ, Dijst MJ, Kwan MP, Pierik FH, de Vries SI (2016): Natural and built environmental exposures on children's active school travel: A Dutch global positioning system-based cross-sectional study. *Health & Place*, 39: 101–109.

[730] Koziatek O, Dragicevic S, Li S (2016): Geospatial modelling approach for 3D urban densification developments. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLI-B2: 349–352.

[731] Pittore M, Wieland M (2012): Toward a rapid probabilistic seismic vulnerability assessment using satellite and ground-based remote sensing. *Natural Hazards*, 68(1): 115–145.

[732] Wieland M, Pittore M, Parolai S, Zschau J, Moldobekov B, Begaliev U (2012): Estimating building inventory for rapid seismic vulnerability assessment: Towards an integrated approach based on multi-source imaging. *Soil Dynamics and Earthquake Engineering*, 36: 70–83.

[733] Mastrucci A, Marvuglia A, Popovici E, Leopold U, Benetto E (2016): Geospatial characterization of building material stocks for the life cycle assessment of end-of-life scenarios at the urban scale. *Resources, Conservation & Recycling*, Advance online publication.

[734] Taubenbock H, Standfuß I, Klotz M, Wurm M (2016): The Physical Density of the City—Deconstruction of the Delusive Density Measure with Evidence from Two European Megacities. *ISPRS International Journal of Geo-Information*, 5(11): 206.

[735] Amado M, Poggi F, Amado AR (2016): Energy efficient city: A model for urban planning. *Sustainable Cities and Society*, 26: 476–485.

[736] Zucker G, Judex F, Blöchle M, Köstl M, Widl E, Hauer S, Bres A, Zeilinger J (2016): A new method for optimizing operation of large neighborhoods of buildings using thermal simulation. *Energy and Buildings*, 125: 153–160.

[737] Xie Y, Weng A, Weng Q (2015): Population Estimation of Urban Residential Communities Using Remotely Sensed Morphologic Data. *IEEE Geoscience and Remote Sensing Letters*, 12(5): 1111–1115.

[738] Schebek L, Schnitzer B, Blesinger D, Köhn A, Miekley B, Linke HJ, Lohmann A, Motzko C, Seemann A (2016): Material stocks of the non-residential building sector: the case of the Rhine-Main area. *Resources, Conservation & Recycling*, Advance online publication.

[739] Bakillah M, Liang S, Mobasheri A, Jokar Arsanjani J, Zipf A (2014): Fine-resolution population mapping using OpenStreetMap points-of-interest. *International Journal of Geographical Information Science*, 28(9): 1940–1963.

[740] Biljecki F, Ledoux H, Stoter J (2015): Improving the consistency of multi-LOD CityGML datasets by removing redundancy. M Breunig, AD Mulhim, E Butwilowski, PV Kuper, J Benner, KH Häfele, eds., *3D Geoinformation Science*, Springer International Publishing, pp. 1–17.

[741] Arroyo Ohori K, Ledoux H, Biljecki F, Stoter J (2015): Modeling a 3D City Model and Its Levels of Detail as a True 4D Model. *ISPRS International Journal of Geo-Information*, 4(3): 1055–1075.

[742] ISO/TC 184 (2012): ISO/PAS 17506:2012 - Industrial automation systems and integration – COLLADA digital asset schema specification for 3D visualization of industrial data.

[743] Sony Computer Entertainment (2008): COLLADA – Digital Asset Schema Release 1.5.0. The Khronos Group.

[744] ISO (2013): ISO/IEC 19775-1:2013 Information technology – Computer graphics, image processing and environmental data representation – Extensible 3D (X3D) – Part 1: Architecture and base components.

[745] Gröger G, Plümer L (2009): How to achieve consistency for 3D city models. *Geoinformatica*, 15(1): 137–165.

[746] Stoter J, Visser T, van Oosterom P, Quak W, Bakker N (2011): A semantic-rich multi-scale information model for topography. *International Journal of Geographical Information Science*, 25(5): 739–763.

[747] Gröger G, Plümer L (2012): Transaction rules for updating surfaces in 3D GIS. *ISPRS Journal of Photogrammetry and Remote Sensing*, 69: 134–145.

[748] Gröger G, Plümer L (2005): How to Get 3-D for the Price of 2-D? Topology and Consistency of 3-D Urban GIS. *Geoinformatica*, 9(2): 139–158.

[749] Yang B, Huang R, Li J, Tian M, Dai W, Zhong R (2016): Automated Reconstruction of Building LoDs from Airborne LiDAR Point Clouds Using an Improved Morphological Scale Space. *Remote Sensing*, 9(1): 14.

[750] Scianna A (2013): Building 3D GIS data models using open source software. *Applied Geomatics*, 5(2): 119–132.

[751] Arkin EM, Chew LP, Huttenlocher DP, Kedem K, Mitchell JSB (1991): An efficiently computable metric for comparing polygonal shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(3): 209–216.

[752] Zheng Y, Cohen-Or D, Averkiou M, Mitra NJ (2014): Recurring part arrangements in shape collections. *Computer Graphics Forum*, 33(2): 115–124.

[753] Dilo A, van Oosterom P, Hofman A (2009): Constrained tGAP for generalization between scales: The case of Dutch topographic data. *Computers, Environment and Urban Systems*, 33(5): 388–402.

[754] Zhang X, Ai T, Stoter J, Zhao X (2014): Data matching of building polygons at multiple map scales improved by contextual information and relaxation. *ISPRS Journal of Photogrammetry and Remote Sensing*, 92: 147–163.

[755] Huffman D (1952): A Method for the Construction of Minimum-Redundancy Codes. *Proceedings of the IRE*, 40(9): 1098–1101.

[756] Salomon D (2007): *Data Compression—The Complete Reference*. Springer London, 4 edn.

[757] XML Core Working Group (2010): XML Linking Language (XLink) Version 1.1. URL http://www.w3.org/TR/xlink11/.

[758] Stadler A, Nagel C, König G, Kolbe TH (2009): Making Interoperability Persistent: A 3D Geo Database Based on CityGML. J Lee, S Zlatanova, eds., *3D Geo-Information Sciences*, Springer Berlin Heidelberg, pp. 175–192.

[759] Iwaszczuk D, Stilla U (2010): A concept for assignment of textures to partially occluded faces of 3D city models stored in CityGML. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XXXVIII-4/W15: 57–62.

[760] Li L, Luo F, Zhu H, Ying S, Zhao Z (2016): A two-level topological model for 3D features in CityGML. *Computers, Environment and Urban Systems*, 59: 11–24.

[761] Lake R, Burggraf D, Trninić M, Rae L (2004): *Geography mark-up language (GML). Foundation for the Geo-web*. John Wiley & Sons, Inc.

[762] Dalamagas T, Cheng T, Winkel KJ, Sellis T (2004): Clustering XML Documents by Structure. GA Vouros, T Panayiotopoulos, eds., *Methods and Applications of Artificial Intelligence*, Springer Berlin Heidelberg, pp. 112–121.

[763] Arroyo Ohori GAK (2016): *Higher-dimensional modelling of geographic information*. Ph.D. thesis, Delft University of Technology, Delft, the Netherlands.

[764] van Oosterom P, Meijers M (2014): Vario-scale data structures supporting smooth zoom and progressive transfer of 2D and 3D data. *International Journal of Geographical Information Science*, 28(3): 455–478.

[765] Meijers M (2011): *Variable-scale Geoinformation*. Ph.D. thesis, Delft University of Technology.

[766] Devogele T, Trevisan J, Raynal L (1996): Building a multi-scale database with scale-transition relationships. *Proceedings of the 7th International Symposium on Spatial Data Handling*. Delft, Netherlands, pp. 337–351.

[767] Arroyo Ohori K, Biljecki F, Stoter J, Ledoux H (2013): Manipulating higher dimensional spatial information. D Vandenbroucke, B Bucher, J Crompvoets, eds., *Proceedings of the 16th AGILE Conference on Geographic Information Science*. Leuven, Belgium, pp. 1–7.

[768] Biljecki F, Arroyo Ohori K, Ledoux H, Peters R, Stoter J (2016): Population Estimation Using a 3D City Model: A Multi-Scale Country-Wide Study in the Netherlands. *PLOS ONE*, 11(6): e0156808.

[769] Biljecki F, Ledoux H, Stoter J (2017): Does a Finer Level of Detail of a 3D City Model Bring an Improvement for Estimating Shadows? A Abdul-Rahman, ed., *Advances in 3D Geoinformation*, Springer International Publishing, Cham, pp. 31–47.

[770] Mackaness WA (2007): Understanding Geographic Space. *Generalisation of Geographic Information*, Elsevier, pp. 1–10.

[771] Deng Y, Cheng JCP (2015): Automatic Transformation of Different Levels of Detail in 3D GIS City Models in CityGML. *International Journal of 3-D Information Modeling*, 4(3): 1–21.

[772] Cervilla AR, Tabik S, Vías J, Mérida M, Romero LF (2016): Total 3D-viewshed Map: Quantifying the Visible Volume in Digital Elevation Models. *Transactions in GIS*, Advance online publication.

[773] Yu S, Yu B, Song W, Wu B, Zhou J, Huang Y, Wu J, Zhao F, Mao W (2016): View-based greenery: A three-dimensional assessment of city buildings' green visibility using Floor Green View Index. *Landscape and Urban Planning*, 152: 13–26.

[774] Grassi S, Klein TM (2016): 3D augmented reality for improving social acceptance and public participation in wind farms planning. *Journal of Physics: Conference Series*, 749: 012020.

[775] Hengl T (2006): Finding the right pixel size. *Computers & Geosciences*, 32(9): 1283–1298.

[776] Usery EL, Finn MP, Scheidt DJ, Ruhl S, Beard T, Bearden M (2004): Geospatial data resampling and resolution effects on watershed modeling: A case study using the agricultural nonpoint source pollution model. *Journal of Geographical Systems*, 6(3): 289–306.

[777] Guo-an T, Yang-he H, Strobl J, Wang-qing L (2001): The impact of resolution on the accuracy of hydrologic data derived from DEMs. *Journal of Geographical Sciences*, 11(4): 393–401.

[778] Booij MJ (2005): Impact of climate change on river flooding assessed with different spatial model resolutions. *Journal of Hydrology*, 303(1-4): 176–198.

[779] Chaubey I, Cotter AS, Costello TA, Soerens TS (2005): Effect of DEM data resolution on SWAT output uncertainty. *Hydrological Processes*, 19(3): 621–628.

[780] Ling Y, Ehlers M, Usery EL, Madden M (2008): Effects of spatial resolution ratio in image fusion. *International Journal of Remote Sensing*, 29(7): 2157–2167.

[781] Pogson M, Smith P (2015): Effect of spatial data resolution on uncertainty. *Environmental Modelling & Software*, 63: 87–96.

[782] Veregin H (2000): Quantifying positional error induced by line simplification. *International Journal of Geographical Information Science*, 14(2): 113–130.

[783] Cheung CK, Shi W (2004): Estimation of the Positional Uncertainty in Line Simplification in GIS. *The Cartographic Journal*, 41(1): 37–45.

[784] Ruiz Arias JA, Tovar Pescador J, Pozo Vázquez D, Alsamamra H (2009): A comparative analysis of DEM-based models to estimate the solar radiation in mountainous terrain. *International Journal of Geographical Information Science*, 23(8): 1049–1076.

[785] Hannibal C, Brown A, Knight M (2005): An assessment of the effectiveness of sketch representations in early stage digital design. *International Journal of Architectural Computing*, 3(1): 107–126.

[786] Strzalka A, Monien D, Koukofikis A, Eicker U (2015): Sensitivity analysis for minimization of input data for urban scale heat demand forecasting. *14th International Conference on Sustainable Energy Technologies SET*. Nottingham, United Kingdom, pp. 1–10.

[787] Deng Y, Cheng JCP, Anumba C (2016): A framework for 3D traffic noise mapping using data from BIM and GIS integration. *Structure and Infrastructure Engineering*, 12(10): 1267–1280.

[788] Barbano G, Egusquiza A (2015): Interconnection between Scales for Friendly and Affordable Sustainable Urban Districts Retrofitting. *Energy Procedia*, 78: 1853–1858.

[789] Brédif M (2013): Image-Based Rendering of LOD1 3D City Models for traffic-augmented Immersive Street-view Navigation. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, II-3/W3: 7–11.

[790] Fai S, Rafeiro J (2014): Establishing an Appropriate Level of Detail (LoD) for a Building Information Model (BIM) – West Block, Parliament Hill, Ottawa, Canada. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, II-5: 123–130.

[791] Anderson W, Guikema S, Zaitchik B, Pan W (2014): Methods for Estimating Population Density in Data-Limited Areas: Evaluating Regression and Tree-Based Models in Peru. *PLOS ONE*, 9(7): e100037.

[792] Hillson R, Alejandre JD, Jacobsen KH, Ansumana R, Bockarie AS, Bangura U, Lamin JM, Malanoski AP, Stenger DA (2014): Methods for Determining the Uncertainty of Population Estimates Derived from Satellite Imagery and Limited Survey Data: A Case Study of Bo City, Sierra Leone. *PLOS ONE*, 9(11): e112241.

[793] Welch R (1980): Monitoring urban population and energy utilization patterns from satellite Data. *Remote sensing of Environment*, 9(1): 1–9.

[794] Lu D, Weng Q, Li G (2006): Residential population estimation using a remote sensing derived impervious surface approach. *International Journal of Remote Sensing*, 27(16): 3553–3570.

[795] Lo CP (1995): Automated population and dwelling unit estimation from high-resolution satellite images: a GIS approach. *International Journal of Remote Sensing*, 16(1): 17–34.

[796] Lo CP (2013): Population Estimation Using Geographically Weighted Regression. *GIScience & Remote Sensing*, 45(2): 131–148.

[797] Tobler WR (1969): Satellite confirmation of settlement size coefficients. *Area*, 1(3): 30–34.

[798] Kraus SP, Senger LW, Ryerson JM (1974): Estimating population from photographically determined residential land use types. *Remote sensing of Environment*, 3(1): 35–42.

[799] Lo CP, Welch R (1977): Chinese Urban Population Estimates. *Annals of the Association of American Geographers*, 67(2): 246–253.

[800] Al-Garni AM (1995): Mathematical predictive models for population estimation in urban areas using space products and GIS technology. *Mathematical and Computer Modelling*, 22(1): 95–107.

[801] Wu C, Murray AT (2007): Population Estimation Using Landsat Enhanced Thematic Mapper Imagery. *Geographical Analysis*, 39(1): 26–43.

[802] Zhan FB, Tapia Silva FO, Santillana M (2010): Estimating small-area population growth using geographic-knowledge-guided cellular automata. *International Journal of Remote Sensing*, 31(21): 5689–5707.

[803] Langford M (2013): An Evaluation of Small Area Population Estimation Techniques Using Open Access Ancillary Data. *Geographical Analysis*, 45(3): 324–344.

[804] Xie Z (2013): A Framework for Interpolating the Population Surface at the Residential-Housing-Unit Level. *GIScience & Remote Sensing*, 43(3): 233–251.

[805] Deng C, Wu C, Wang L (2010): Improving the housing-unit method for small-area population estimation using remote-sensing and GIS information. *International Journal of Remote Sensing*, 31(21): 5673–5688.

[806] Wu C, Murray AT (2005): A cokriging method for estimating population density in urban areas. *Computers, Environment and Urban Systems*, 29(5): 558–579.

[807] Bagan H, Yamagata Y (2012): Landsat analysis of urban growth: How Tokyo became the world's largest megacity during the last 40 years. *Remote sensing of Environment*, 127: 210–222.

[808] Yuan Y, Smith RM, Limp WF (1997): Remodeling census population with spatial information from LandSat TM imagery. *Computers, Environment and Urban Systems*, 21(3-4): 245–258.

[809] Stevens FR, Gaughan AE, Linard C, Tatem AJ (2015): Disaggregating Census Data for Population Mapping Using Random Forests with Remotely-Sensed and Ancillary Data. *PLOS ONE*, 10(2): e0107042.

[810] Gaughan AE, Stevens FR, Linard C, Jia P, Tatem AJ (2013): High Resolution Population Distribution Maps for Southeast Asia in 2010 and 2015. *PLOS ONE*, 8(2): e55882.

[811] Anderson SJ, Tuttle BT, Powell RL, Sutton PC (2010): Characterizing relationships between population density and nighttime imagery for Denver, Colorado: issues of scale and representation. *International Journal of Remote Sensing*, 31(21): 5733–5746.

[812] Sutton P (1997): Modeling population density with night-time satellite imagery and GIS. *Computers, Environment and Urban Systems*, 21(3-4): 227–244.

[813] Doll CNH, Muller JP, Morley JG (2006): Mapping regional economic activity from night-time light satellite imagery. *Ecological Economics*, 57(1): 75–92.

[814] Pozzi F, Small C (2005): Analysis of urban land cover and population density in the United States. *Photogrammetric Engineering and Remote Sensing*, 71(6): 719–726.

[815] Xie Y (1995): The overlaid network algorithms for areal interpolation problem. *Computers, Environment and Urban Systems*, 19(4): 287–306.

[816] Steiger E, Westerholt R, Resch B, Zipf A (2015): Twitter as an indicator for whereabouts of people? Correlating Twitter with UK census data. *Computers, Environment and Urban Systems*, 54: 255–265.

[817] Harvey JT (2010): Estimating census district populations from satellite imagery: Some approaches and limitations. *International Journal of Remote Sensing*, 23(10): 2071–2095.

[818] van Oosterom P, Martinez-Rubi O, Ivanova M, Horhammer M, Geringer D, Ravada S, Tijssen T, Kodde M, Gonçalves R (2015): Massive point cloud data management: Design, implementation and execution of a point cloud benchmark. *Computers and Graphics*, 49: 92–125.

[819] van der Sande C, Soudarissanane S, Khoshelham K (2010): Assessment of Relative Accuracy of AHN-2 Laser Scanning Data Using Planar Features. *Sensors*, 10(9): 8198–8214.

[820] Brinegar SJ, Popick SJ (2013): A Comparative Analysis of Small Area Population Estimation Methods. *Cartography and Geographic Information Science*, 37(4): 273–284.

[821] Mennis J (2009): Dasymetric Mapping for Estimating Population in Small Areas. *Geography Compass*, 3(2): 727–745.

[822] Goodchild MF, Lam NSN (1980): Areal Interpolation - a Variant of the Traditional Spatial Problem. *Geo-Processing*, 1(3): 297–312.

[823] Flowerdew R, Green M, Kehris E (1991): Using areal interpolation methods in geographic information systems. *Papers in Regional Science*, 70(3): 303–315.

[824] Liu XH, P C Kyriakidis (2008): Population-density estimation using regression and area-to-point residual kriging. *International Journal of Geographical Information Science*, 22(4): 431–447.

[825] Langford M (2006): Obtaining population estimates in non-census reporting zones: An evaluation of the 3-class dasymetric method. *Computers, Environment and Urban Systems*, 30(2): 161–180.

[826] Zandbergen PA, Ignizio DA (2013): Comparison of Dasymetric Mapping Techniques for Small-Area Population Estimates. *Cartography and Geographic Information Science*, 37(3): 199–214.

[827] Mennis J (2003): Generating Surface Models of Population Using Dasymetric Mapping. *The Professional Geographer*, 55(1): 31–42.

[828] Wright JK (1936): A Method of Mapping Densities of Population: With Cape Cod as an Example. *Geographical Review*, 26(1): 103.

[829] Maantay JA, Maroko AR, Herrmann C (2007): Mapping Population Distribution in the Urban Environment: The Cadastral-based Expert Dasymetric System (CEDS). *Cartography and Geographic Information Science*, 34(2): 77–102.

[830] Eicher CL, Brewer CA (2001): Dasymetric mapping and areal interpolation: Implementation and evaluation. *Cartography and Geographic Information Science*, 28(2): 125–138.

[831] Mennis J, Hultgren T (2006): Intelligent Dasymetric Mapping and Its Application to Areal Interpolation. *Cartography and Geographic Information Science*, 33(3): 179–194.

[832] Holt JB, Lo CP, Hodler TW (2004): Dasymetric Estimation of Population Density and Areal Interpolation of Census Data. *Cartography and Geographic Information Science*, 31(2): 103–121.

[833] Alahmadi M, Atkinson P, Martin D (2015): Fine spatial resolution residential land-use data for small-area population mapping: a case study in Riyadh, Saudi Arabia. *International Journal of Remote Sensing*, 36(17): 4315–4331.

[834] Liu X, Clarke K (2002): Estimation of Residential Population Using High Resolution Satellite Imagery. *Proceedings of the 3rd Symposium in Remote Sensing of Urban Areas*. Istanbul, Turkey, pp. 153–160.

[835] Greger K (2015): Spatio-Temporal Building Population Estimation for Highly Urbanized Areas Using GIS. *Transactions in GIS*, 19(1): 129–150.

[836] Zoraghein H, Leyk S, Ruther M, Buttenfield BP (2016): Exploiting temporal information in parcel data to refine small area population estimates. *Computers, Environment and Urban Systems*, 58: 19–28.

[837] Hermosilla T, Ruiz LA, Recio JA, Balsa-Barreiro J (2012): Land-use mapping of Valencia city area from aerial images and LiDAR data. *GEOProcessing 2012: The Fourth International Conference on Advanced Geographic Information Systems, Applications, and Services*. Valencia, Spain, pp. 232–237.

[838] Xiong B, Oude Elberink S, Vosselman G (2014): A graph edit dictionary for correcting errors in roof topology graphs reconstructed from point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing*, 93: 227–242.

[839] Swanson DA, Hough Jr GC (2012): An Evaluation of Persons per Household (PPH) Estimates Generated by the American Community Survey: A Demographic Perspective. *Population Research and Policy Review*, 31(2): 235–266.

[840] Fisher PF, Langford M (1996): Modeling Sensitivity to Accuracy in Classified Imagery: A Study of Areal Interpolation by Dasymetric Mapping. *The Professional Geographer*, 48(3): 299–309.

[841] Fisher PF, Langford M (1995): Modelling the Errors in Areal Interpolation between Zonal Systems by Monte Carlo Simulation. *Environment and Planning A*, 27(2): 211–224.

[842] Sadahiro Y (1999): Accuracy of areal interpolation: A comparison of alternative methods. *Journal of Geographical Systems*, 1(4): 323–346.

[843] Zhu L, Lehtomäki M, Hyyppä J, Puttonen E, Krooks A, Hyyppä H (2015): Automated 3D Scene Reconstruction from Open Geospatial Data Sources: Airborne Laser Scanning and a 2D Topographic Database. *Remote Sensing*, 7(6): 6710–6740.

[844] Tatem AJ, Noor AM, von Hagen C, Di Gregorio A, Hay SI (2007): High Resolution Population Maps for Low Income Nations: Combining Land Cover and Census in East Africa. *PLOS ONE*, 2(12): e1298.

[845] Linard C, Gilbert M, Snow RW, Noor AM, Tatem AJ (2012): Population Distribution, Settlement Patterns and Accessibility across Africa in 2010. *PLOS ONE*, 7(2): e31743.

[846] Spoorenberg T (2014): Provisional results of the 2014 census of Myanmar: The surprise that wasn't. *Asian Population Studies*, 11(1): 4–6.

[847] Spoorenberg T (2015): Myanmar's first census in more than 30 years: A radical revision of the official population count. *Population & Societies*, 527: 1–4.

[848] van Winden K, Biljecki F, van der Spek S (2016): Automatic Update of Road Attributes by Mining GPS Tracks. *Transactions in GIS*, 20(5): 664–683.

[849] Stoter J, Ledoux H, Zlatanova S, Biljecki F (2016): Towards sustainable and clean 3D Geoinformation. TH Kolbe, R Bill, A Donaubauer, eds., *Geoinformationssysteme 2016: Beiträge zur 3. Münchner GI-Runde*, Munich, Germany, pp. 100–113.

[850] Chen K (2002): An approach to linking remotely sensed data and areal census data. *International Journal of Remote Sensing*, 23(1): 37–48.

[851] Langford M, Higgs G (2006): Measuring Potential Access to Primary Healthcare Services: The Influence of Alternative Spatial Representations of Population. *The Professional Geographer*, 58(3): 294–306.

[852] Hay SI, Noor AM, Nelson A, Tatem AJ (2005): The accuracy of human population maps for public health application. *Tropical Medicine and International Health*, 10(10): 1073–1086.

[853] Poulsen E, Kennedy LW (2004): Using Dasymetric Mapping for Spatially Aggregated Crime Data. *Journal of Quantitative Criminology*, 20(3): 243–262.

[854] Lin J, Cromley RG (2015): Evaluating geolocated Twitter data as a control layer for areal interpolation of population. *Applied Geography*, 58: 41–47.

[855] Vine MF, Degnan D, Hanchette C (1997): Geographic information systems: their use in environmental epidemiologic research. *Environmental Health Perspectives*, 105(6): 598–605.

[856] Wandl A, Nadin V, Zonneveld W, Rooij R (2014): Beyond urban–rural classifications: Characterising and mapping territories-in-between across Europe. *Landscape and Urban Planning*, 130: 50–63.

[857] Ögren M, Barregard L (2016): Road Traffic Noise Exposure in Gothenburg 1975–2010. *PLOS ONE*, 11(5): e0155328.

[858] Kohler N, Hassler U (2002): The building stock as a research object. *Building Research & Information*, 30(4): 226–236.

[859] Kavgic M, Mavrogianni A, Mumovic D, Summerfield A, Stevanovic Z, Djurovic-Petrovic M (2010): A review of bottom-up building stock models for energy consumption in the residential sector. *Building and Environment*, 45(7): 1683–1697.

[860] Hargreaves AJ (2015): Representing the dwelling stock as 3D generic tiles estimated from average residential density. *Computers, Environment and Urban Systems*, 54: 280–300.

[861] Levy PS, Lemeshow S (2008): Stratification and Stratified Random Sampling. *Sampling of Populations: Methods and Applications*, John Wiley & Sons, Inc., Hoboken, NJ, USA, pp. 121–142.

[862] Erdélyi R, Wang Y, Guo W, Hanna E, Colantuono G (2014): Three-dimensional SOlar RAdiation Model (SORAM) and its application to 3-D urban planning. *Solar Energy*, 101: 63–73.

[863] Burnicki AC, Brown DG, Goovaerts P (2007): Simulating error propagation in land-cover change analysis: The implications of temporal dependence. *Computers, Environment and Urban Systems*, 31(3): 282–302.

[864] Fogl M, Moudrý V (2016): Influence of vegetation canopies on solar potential in urban environments. *Applied Geography*, 66: 73–80.

[865] Bajsanski I, Stojakovic V, Jovanovic M (2016): Effect of tree location on mitigating parking lot insolation. *Computers, Environment and Urban Systems*, 56: 59–67.

[866] Williams L (1978): Casting curved shadows on curved surfaces. *ACM SIGGRAPH Computer Graphics*, 12(3): 270–274.

[867] Woo A, Poulin P, Fournier A (1990): A survey of shadow algorithms. *IEEE Computer Graphics and Applications*, 10(6): 13–32.

[868] Whitted T (1980): An improved illumination model for shaded display. *Communications of the ACM*, 23(6): 343–349.

[869] Möller T, Trumbore B (1997): Fast, minimum storage ray-triangle intersection. *Journal of Graphics Tools*, 2(1): 21–28.

[870] Blinn J (1988): Me and My (Fake) Shadow. *IEEE Computer Graphics and Applications*, 8(1): 82–86.

[871] Huttenlocher DP, Klanderman GA, Rucklidge WJ (1993): Comparing images using the Hausdorff distance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(9): 850–863.

[872] Samal A, Seth S, Cueto K (2004): A feature-based approach to conflation of geospatial sources. *International Journal of Geographical Information Science*, 18(5): 459–489.

[873] Ruiz JJ, Ariza FJ, Ureña MA, Blázquez EB (2011): Digital map conflation: a review of the process and a proposal for classification. *International Journal of Geographical Information Science*, 25(9): 1439–1466.

[874] Goodchild MF, Hunter GJ (1997): A simple positional accuracy measure for linear features. *International Journal of Geographical Information Science*, 11(3): 299–306.

[875] Min D, Zhilin L, Xiaoyong C (2007): Extended Hausdorff distance for spatial objects in GIS. *International Journal of Geographical Information Science*, 21(4): 459–475.

[876] Girres JF, Touya G (2010): Quality Assessment of the French OpenStreetMap Dataset. *Transactions in GIS*, 14(4): 435–459.

[877] Gil de la Vega P, Ariza-López FJ, Mozas-Calvache AT (2016): Models for positional accuracy assessment of linear features: 2D and 3D cases. *Survey Review*, 48(350): 347–360.

[878] Cignoni P, Rocchini C, Scopigno R (1998): Metro: Measuring Error on Simplified Surfaces. *Computer Graphics Forum*, 17(2): 167–174.

[879] Mustiere S, Devogele T (2008): Matching Networks with Different Levels of Detail. *Geoinformatica*, 12(4): 435–453.

[880] Liu L, Qiao S, Zhang Y, Hu J (2012): An efficient outlying trajectories mining approach based on relative distance. *International Journal of Geographical Information Science*, 26(10): 1789–1810.

[881] Agarwal PK, Har-Peled S, Sharir M, Wang Y (2010): Hausdorff distance under translation for points and balls. *ACM Transactions on Algorithms*, 6(4): 1–26.

[882] Aspert N, Santa-Cruz D, Ebrahimi T (2002): MESH: measuring errors between surfaces using the Hausdorff distance. *IEEE International Conference on Multimedia and Expo (ICME)*. IEEE, pp. 705–708.

[883] Bretagnon P, Francou G (1988): Planetary theories in rectangular and spherical variables. VSOP 87 solutions. *Astronomy and Astrophysics*, 202: 309–315.

[884] Meeus JH (1998): *Astronomical Algorithms*. Willmann-Bell, Incorporated, Richmond, VA, 2 edn.

[885] Biljecki F, Arroyo Ohori K (2015): Automatic Semantic-preserving Conversion Between OBJ and CityGML. *Eurographics Workshop on Urban Data Modelling and Visualisation 2015*. Delft, Netherlands, pp. 25–30.

[886] Hobby JD (1999): Practical segment intersection with finite precision output. *Computational Geometry*, 13(4): 199–214.

[887] Ledoux H, Arroyo Ohori K, Meijers M (2014): A triangulation-based approach to automatically repair GIS polygons. *Computers & Geosciences*, 66: 121–131.

[888] Strzalka A, Eicker U, Coors V, Schumacher J (2010): Modeling energy demand for heating at city scale. *Proceedings of SimBuild 2010, 4th National Conference of IBPSA-USA*. New York City, NY, United States, pp. 358–364.

[889] Fisher-Gewirtzman D, Shashkov A, Doytsher Y (2013): Voxel based volumetric visibility analysis of urban environments. *Survey Review*, 45(333): 451–461.

[890] van der Hoeven F, Wandl A (2015): *Hotterdam: How space is making Rotterdam warmer, how this affects the health of its inhabitants, and what can be done about it*. TU Delft.

[891] Fonseca JA, Schlueter A (2015): Integrated model for characterization of spatiotemporal building energy consumption patterns in neighborhoods and city districts. *Applied Energy*, 142: 247–265.

[892] Fonseca JA, Nguyen TA, Schlueter A, Marechal F (2016): City Energy Analyst (CEA): Integrated framework for analysis and optimization of building energy systems in neighborhoods and city districts. *Energy and Buildings*, 113: 202–226.

[893] Kanters J, Wall M, Kjellsson E (2014): The Solar Map as a Knowledge Base for Solar Energy Use. *Energy Procedia*, 48: 1597–1606.

[894] Biljecki F, Heuvelink GBM, Ledoux H, Stoter J (2015): Propagation of positional error in 3D GIS: estimation of the solar irradiation of building roofs. *International Journal of Geographical Information Science*, 29(12): 2269–2294.

[895] Biljecki F, Ledoux H, Stoter J (2014): Error propagation in the computation of volumes in 3D city models with the Monte Carlo method. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, II-2: 31–39.

[896] Heuvelink GBM (2005): Propagation of error in spatial modelling with GIS. PA Longley, MF Goodchild, DJ Maguire, DW Rhind, eds., *Geographical Information Systems. Principles, Techniques, Management and Applications*, Wiley, pp. 207–217.

[897] Xue J, Leung Y, Ma JH (2015): High-order Taylor series expansion methods for error propagation in geographic information systems. *Journal of Geographical Systems*, 17(2): 187–206.

[898] Yeh AGo, Li X (2006): Errors and uncertainties in urban cellular automata. *Computers, Environment and Urban Systems*, 30(1): 10–28.

[899] Taylor JR (1997): *An Introduction to Error Analysis*. The Study of Uncertainties in Physical Measurements. University Science Books, Sausalito, California.

[900] Behrens TEJ, Woolrich MW, Jenkinson M, Johansen Berg H, Nunes RG, Clare S, Matthews PM, Brady JM, Smith SM (2003): Characterization and propagation of uncertainty in diffusion-weighted MR imaging. *Magnetic Resonance in Medicine*, 50(5): 1077–1088.

[901] MacLeod M, Fraser AJ, Mackay D (2002): Evaluating and expressing the propagation of uncertainty in chemical fate and bioaccumulation models. *Environmental Toxicology and Chemistry*, 21(4): 700–709.

[902] Veregin H (2005): Data quality parameters. PA Longley, MF Goodchild, DJ Maguire, DW Rhind, eds., *Geographical Information Systems. Principles, Techniques, Management and Applications*, John Wiley & Sons, Inc, pp. 177–189.

[903] Fisher PF (2005): Models of uncertainty in spatial data. PA Longley, MF Goodchild, DJ Maguire, DW Rhind, eds., *Geographical Information Systems. Principles, Techniques, Management and Applications*, Wiley, pp. 191–205.

[904] Heuvelink GBM, Burrough PA, Stein A (1989): Propagation of errors in spatial modelling with GIS. *International Journal of Geographical Information Systems*, 3(4): 303–322.

[905] Arbia G, Griffith D, Haining R (1998): Error propagation modelling in raster GIS: overlay operations. *International Journal of Geographical Information Science*, 12(2): 145–167.

[906] van Oort P, Stein A, Bregt AK, de Bruin S, Kuipers J (2005): A variance and covariance equation for area estimates with a geographic information system. *Forest Science*, 51(4): 347–356.

[907] Shi W, Li QQ, Zhu C (2005): Estimating the propagation error of DEM from higher-order interpolation algorithms. *International Journal of Remote Sensing*, 26(14): 3069–3084.

[908] de Bruin S, Heuvelink GBM, Brown JD (2008): Propagation of positional measurement errors to agricultural field boundaries and associated costs. *Computers and Electronics in Agriculture*, 63(2): 245–256.

[909] Shi W, Cheung CK, Zhu C (2003): Modelling error propagation in vector-based buffer analysis. *International Journal of Geographical Information Science*, 17(3): 251–271.

[910] Heuvelink GBM, Burrough PA (1993): Error propagation in cartographic modelling using Boolean logic and continuous classification. *International Journal of Geographical Information Systems*, 7(3): 231–246.

[911] Veregin H (1995): Developing and testing of an error propagation model for GIS overlay operations. *International Journal of Geographical Information Systems*, 9(6): 595–619.

[912] Goulden T, Hopkinson C, Jamieson R, Sterling S (2016): Sensitivity of DEM, slope, aspect and watershed attributes to LiDAR measurement uncertainty. *Remote sensing of Environment*, 179: 23–35.

[913] Griffith DA, Millones M, Vincent M, Johnson DL, Hunt A (2007): Impacts of Positional Error on Spatial Regression Analysis: A Case Study of Address Locations in Syracuse, New York. *Transactions in GIS*, 11(5): 655–679.

[914] Zandbergen PA, Hart TC, Lenzer KE, Camponovo ME (2012): Error propagation models to examine the effects of geocoding quality on spatial analysis of individual-level datasets. *Spatial and Spatio-temporal Epidemiology*, 3(1): 69–82.

[915] Shi W (1998): A generic statistical approach for modelling error of geometric features in GIS. *International Journal of Geographical Information Science*, 12(2): 131–143.

[916] Hengl T, Heuvelink GBM, van Loon EE (2010): On the uncertainty of stream networks derived from elevation data: the error propagation approach. *Hydrology and Earth System Sciences*, 14(7): 1153–1165.

[917] Emmi PC, Horton CA (1995): A Monte Carlo simulation of error propagation in a GIS-based assessment of seismic risk. *International Journal of Geographical Information Systems*, 9(4): 447–461.

[918] Qi H, Qi P, Altinakar MS (2013): GIS-Based Spatial Monte Carlo Analysis for Integrated Flood Management with Two Dimensional Flood Simulation. *Water Resources Management*, 27(10): 3631–3645.

[919] Zhou G, Esaki T, Mitani Y, Xie M, Mori J (2003): Spatial probabilistic modeling of slope failure using an integrated GIS Monte Carlo simulation approach. *Engineering Geology*, 68(3-4): 373–386.

[920] Davis TJ, Keller CP (1997): Modelling uncertainty in natural resource analysis using fuzzy sets and Monte Carlo simulation: slope stability prediction. *International Journal of Geographical Information Science*, 11(5): 409–434.

[921] Fan L, Smethurst J, Atkinson P, Powrie W (2014): Propagation of vertical and horizontal source data errors into a TIN with linear interpolation. *International Journal of Geographical Information Science*, 28(7): 1378–1400.

[922] Hong S, Vonderohe A (2014): Uncertainty and Sensitivity Assessments of GPS and GIS Integrated Applications for Transportation. *Sensors*, 14(2): 2683–2702.

[923] Fisher PF (1991): Modelling soil map-unit inclusions by Monte Carlo simulation. *International Journal of Geographical Information Systems*, 5(2): 193–208.

[924] Zhang S, Li X, Chen Y (2015): Error assessment of grid-based direct solar radiation models. *International Journal of Geographical Information Science*, 29(10): 1782–1806.

[925] Nault E, Peronato G, Rey E, Andersen M (2015): Review and critical analysis of early-design phase evaluation metrics for the solar potential of neighborhood designs. *Building and Environment*, 92: 679–691.

[926] Molin A, Schneider S, Rohdin P, Moshfegh B (2016): Assessing a regional building applied PV potential – Spatial and dynamic analysis of supply and load matching. *Renewable Energy*, 91: 261–274.

[927] Yang H, Lu L (2007): The Optimum Tilt Angles and Orientations of PV Claddings for Building-Integrated Photovoltaic (BIPV) Applications. *Journal of Solar Energy Engineering*, 129(2): 253–255.

[928] Wittmann H, Bajons P, Doneus M, Friesinger H (1997): Identification of roof areas suited for solar energy conversion systems. *Renewable Energy*, 11(1): 25–36.

[929] Mainzer K, Fath K, McKenna R, Stengel J, Fichtner W, Schultmann F (2014): A high-resolution determination of the technical potential for residential-roof-mounted photovoltaic systems in Germany. *Solar Energy*, 105: 715–731.

[930] Šúri M, Huld TA, Dunlop ED (2005): PV-GIS: a web-based solar radiation database for the calculation of PV potential in Europe. *International Journal of Sustainable Energy*, 24(2): 55–67.

[931] Šúri M, Huld TA, Dunlop ED, Ossenbrink HA (2007): Potential of solar electricity generation in the European Union member states and candidate countries. *Solar Energy*, 81(10): 1295–1305.

[932] Huld T, Müller R, Gambardella A (2012): A new solar radiation database for estimating PV performance in Europe and Africa. *Solar Energy*, 86(6): 1803–1815.

[933] Lukač N, Seme S, Dežan K, Žalik B, Štumberger G (2016): Economic and environmental assessment of rooftops regarding suitability for photovoltaic systems installation based on remote sensing data. *Energy*, 107: 854–865.

[934] Baumanns K, Löwner MO (2009): Refined estimation of solar energy potential on roof areas using decision trees on CityGML-data. *Geophysical Research Abstracts. EGU General Assembly*. Vienna, Austria, p. 14044.

[935] Gulin M, Vašak M, Perić N (2013): Dynamical optimal positioning of a photovoltaic panel in all weather conditions. *Applied Energy*, 108: 429–438.

[936] Masters GM (2013): *Renewable and Efficient Electric Power Systems*. Wiley-Interscience, Hoboken, 2 edn.

[937] David M, Lauret P, Boland J (2013): Evaluating tilted plane models for solar radiation using comprehensive testing procedures, at a southern hemisphere location. *Renewable Energy*, 51: 124–131.

[938] Demain C, Journée M, Bertrand C (2013): Evaluation of different models to estimate the global solar radiation on inclined surfaces. *Renewable Energy*, 50: 710–721.

[939] Gulin M, Vašak M, Baotić M (2013): Estimation of the global solar irradiance on tilted surfaces. *Proceedings of the 17th International Conference on Electrical Drives and Power Electronics (EDPE 2013)*. Dubrovnik, Croatia, pp. 334–339.

[940] Reda I, Andreas A (2004): Solar position algorithm for solar radiation applications. *Solar Energy*, 76(5): 577–589.

[941] Perez R, Ineichen P, Seals R, Michalsky J, Stewart R (1990): Modeling daylight availability and irradiance components from direct and global irradiance. *Solar Energy*, 44(5): 271–289.

[942] Charles N, Kabalan M, Singh P (2015): Open source photovoltaic system performance modeling with python. *2015 IEEE Canada International Humanitarian Technology Conference (IHTC2015)*. IEEE, pp. 1–4.

[943] Kalogirou SA (2003): Generation of typical meteorological year (TMY-2) for Nicosia, Cyprus. *Renewable Energy*, 28(15): 2317–2334.

[944] Thevenard DJ, Brunger AP (2002): The development of typical weather years for international locations: Part I, algorithms. *ASHRAE Transactions*, 108: 376–383.

[945] Christensen CB, Barker GM (2001): Effects of tilt and azimuth on annual incident solar radiation for United States locations. *Proceedings of Solar Forum 2001: Solar Energy: The Power to Choose*. Washington, DC, United States, pp. 1–8.

[946] Rowlands IH, Kemery BP, Beausoleil-Morrison I (2011): Optimal solar-PV tilt angle and azimuth: An Ontario (Canada) case-study. *Energy Policy*, 39(3): 1397–1409.

[947] Li Y, Brimicombe AJ, Ralphs MP (2000): Spatial data quality and sensitivity analysis in GIS and environmental modelling: the case of coastal oil spills. *Computers, Environment and Urban Systems*, 24(2): 95–108.

[948] Gosselin PH, Cord M (2006): Feature-based approach to semi-supervised similarity learning. *Pattern Recognition*, 39(10): 1839–1851.

[949] Bayod-Rújula AA, Ortego-Bielsa A, Martínez-Gracia A (2011): Photovoltaics on flat roofs: Energy considerations. *Energy*, 36(4): 1996–2010.

[950] Li Y, Liu C (2017): Estimating solar energy potentials on pitched roofs. *Energy and Buildings*, 139: 101–107.

[951] Brown JD, Heuvelink GBM (2007): The Data Uncertainty Engine (DUE): A software tool for assessing and simulating uncertain environmental variables. *Computers & Geosciences*, 33(2): 172–190.

[952] Heuvelink GBM, Brown JD, van Loon EE (2007): A probabilistic framework for representing and simulating uncertain environmental variables. *International Journal of Geographical Information Science*, 21(5): 497–513.

[953] Ben-Haim G, Dalyot S, Doytsher Y (2015): Local Absolute Vertical Accuracy Computation of Wide-Coverage Digital Terrain Models. *Advances in Spatial Data Handling and Analysis*, Springer International Publishing, Cham, Switzerland, pp. 209–225.

[954] Xue S, Dang Y, Liu J, Mi J, Dong C, Cheng Y, Wang X, Wan J (2016): Bias estimation and correction for triangle-based surface area calculations. *International Journal of Geographical Information Science*, 30(11): 2155–2170.

[955] Goodchild MF (1991): Issues of quality and uncertainty. JC Muller, ed., *Advances in cartography*, Elsevier, New York, United States, pp. 113–139.

[956] Caspary W, Scheuring R (1993): Positional accuracy in spatial databases. *Computers, Environment and Urban Systems*, 17(2): 103–110.

[957] Goodchild MF (2004): A general framework for error analysis in measurement-based GIS. *Journal of Geographical Systems*, 6(4): 323–324.

[958] Loch-Dehbi S, Plümer L (2011): Automatic reasoning for geometric constraints in 3D city models with uncertain observations. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66(2): 177–187.

[959] Navratil G, Achatschitz C (2004): Influence of correlation on the quality of area computation. *Proceedings of the ISSDQ '04*. pp. 291–304.

[960] Leica (2008): Leica ALS70-CM City Mapping Airborne LIDAR Product Specifications. Brochure, Leica Geosystems AG, Heerbrugg.

[961] Wang J, Kutterer H, Fang X (2016): External error modelling with combined model in terrestrial laser scanning. *Survey Review*, 48(346): 40–50.

[962] Mårtensson SG, Reshetyuk Y (2016): Height uncertainty in digital terrain modelling with unmanned aircraft systems. *Survey Review*, Advance online publication.

[963] Rottensteiner F, Sohn G, Gerke M, Wegner JD, Breitkopf U, Jung J (2014): Results of the ISPRS benchmark on urban object detection and 3D building reconstruction. *ISPRS Journal of Photogrammetry and Remote Sensing*, 93: 256–271.

[964] Kabolizade M, Ebadi H, Mohammadzadeh A (2012): Design and implementation of an algorithm for automatic 3D reconstruction of building models using genetic algorithm. *International Journal of Applied Earth Observation and Geoinformation*, 19: 104–114.

[965] D'Agostino RB (1971): An omnibus test of normality for moderate and large size samples. *Biometrika*, 58(2): 341–348.

[966] Mardaljevic J, Roy N (2016): The sunlight beam index. *Lighting Research and Technology*, 48(1): 55–69.

[967] López-Fernández L, Lagüela S, Picón I, González-Aguilera D (2015): Large Scale Automatic Analysis and Classification of Roof Surfaces for the Installation of Solar Panels Using a Multi-Sensor Aerial Platform. *Remote Sensing*, 7(9): 11226–11248.

[968] Chaturvedi K, Kolbe TH (2015): Dynamizers - Modeling and Implementing Dynamic Properties for Semantic 3D City Models. *Eurographics Workshop on Urban Data Modelling and Visualisation*. The Eurographics Association, Delft, Netherlands, pp. 43–48.

[969] Biljecki F, Heuvelink GBM, Ledoux H, Stoter J (2017): The effect of acquisition error and level of detail on the accuracy of spatial analyses. *Cartography and Geographic Information Science*, Advance online publication.

[970] Burrough PA, McDonnell RA (1998): Errors and Quality Control. *Principles of Geographical Information Systems*, Oxford University Press, pp. 220–240.

[971] Goodchild MF, Li L (2012): Assuring the quality of volunteered geographic information. *Spatial Statistics*, 1: 110–120.

[972] Camboim S, Bravo J, Sluter C (2015): An Investigation into the Completeness of, and the Updates to, OpenStreetMap Data in a Heterogeneous Area in Brazil. *ISPRS International Journal of Geo-Information*, 4(3): 1366–1388.

[973] Senaratne H, Mobasheri A, Ali AL, Capineri C, Haklay MM (2017): A review of volunteered geographic information quality assessment methods. *International Journal of Geographical Information Science*, 31(1): 139–167.

[974] Lunetta RS, Congalton RG, Fenstermaker LK, Jensen JR, McGwire KC, Tinney LR (1991): Remote sensing and geographic information systems: Error sources and research issues. *Photogrammetric Engineering and Remote Sensing*, 57(6): 677–687.

[975] Shi W, Ehlers M, Tempfli K (1999): Analytical Modelling of Positional and Thematic Uncertainties in the Integration of Remote Sensing and Geographical Information Systems. *Transactions in GIS*, 3(2): 119–136.

[976] Couturier S, Mas JF, Cuevas G, Benitez J, Vega-Guzman A, Coria-Tapia V (2009): An Accuracy Index with Positional and Thematic Fuzzy Bounds for Land-use/Land-cover Maps. *Photogrammetric Engineering and Remote Sensing*, 75(7): 789–805.

[977] Tayyebi AH, Tayyebi A, Khanna N (2013): Assessing uncertainty dimensions in land-use change models: using swap and multiplicative error models for injecting attribute and positional errors in spatial data. *International Journal of Remote Sensing*, 35(1): 149–170.

[978] Rios JF, Renschler CS (2016): A New Combined Assessment of Mixed Uncertainty in Spatial Models: Conceptualization and Implementation. *Transactions in GIS*, Advance online publication.

[979] Lee M, Chun Y, Griffith D (2015): Uncertainties of spatial data analysis introduced by selected sources of error. *Proceedings of the 13th International Conference on GeoComputation*. Richardson, Texas, USA, pp. 18–24.

[980] Leao SZ (2017): Assessing the trade-off between data quality and spatial resolution for the Thornthwaite Moisture Index mapping. *Journal of Spatial Science*, 62(1): 85–102.

[981] Swan LG, Ugursal VI (2009): Modeling of end-use energy consumption in the residential sector: A review of modeling techniques. *Renewable and Sustainable Energy Reviews*, 13(8): 1819–1835.

[982] Ioannou A, Itard LCM (2015): Energy performance and comfort in residential buildings: Sensitivity for building parameters and occupancy. *Energy and Buildings*, 92: 216–233.

[983] Guerra-Santin O, Itard L (2010): Occupants' behaviour: determinants and effects on residential heating consumption. *Building Research & Information*, 38(3): 318–338.

[984] Majcen D, Itard L, Visscher H (2013): Actual and theoretical gas consumption in Dutch dwellings: What causes the differences? *Energy Policy*, 61: 460–471.

[985] Heuvelink GBM, Brown JD (2016): Uncertain Environmental Variables in GIS. *Encyclopedia of GIS*, Springer International Publishing, Cham, pp. 1–9.

[986] Lukač N, Seme S, Žlaus D, Štumberger G, Žalik B (2014): Buildings roofs photovoltaic potential assessment based on LiDAR (Light Detection And Ranging) data. *Energy*, 66: 598–609.

[987] Brito MC, Amaro e Silva R, Cristovão AR, Teixeira Freitas SR (2016): Obstruction Surveying Methods for PV Application in Urban Environments. *32nd European Photovoltaic Solar Energy Conference and Exhibition*, pp. 2759–2764.

[988] Berk S, Ferlan M (2016): Accurate area determination in the cadaster: case study of Slovenia. *Cartography and Geographic Information Science*, Advance online publication.

[989] Van Renterghem T, Botteldooren D (2010): The importance of roof shape for road traffic noise shielding in the urban environment. *Journal of Sound and Vibration*, 329(9): 1422–1434.

[990] Girres JF (2011): A model to estimate length measurements uncertainty in vector databases. *7th International Symposium on Spatial Data Quality (ISSDQ'11)*. pp. 83–88.

[991] Chrisman NR (2017): Calculating on a round planet. *International Journal of Geographical Information Science*, 31(4): 637–657.

[992] Gahegan M, Ehlers M (2000): A framework for the modelling of uncertainty between remote sensing and geographic information systems. *ISPRS Journal of Photogrammetry and Remote Sensing*, 55(3): 176–188.

[993] Demir N (2015): Use of Airborne Laser Scanning Data and Image-based Three-dimensional (3-D) Edges for Automated Planar Roof Reconstruction. *Lasers in Engineering*, 32: 173–205.

[994] Ariza-López FJ, Rodríguez-Avi J (2015): Estimating the count of completeness errors in geographic data sets by means of a generalized Waring regression model. *International Journal of Geographical Information Science*, 29(8): 1394–1418.

[995] Peronato G, Rey E, Andersen M (2016): 3D-modeling of vegetation from LiDAR point clouds and assessment of its impact on façade solar irradiation. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLII-2/W2: 67–70.

[996] Sun Y, Shahzad M, Zhu X (2016): First prismatic building model reconstruction from tomosar point clouds. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLI-B3: 381–386.

[997] Willenborg B, Sindram M, Kolbe TH (2016): Semantic 3D City Models Serving as Information Hub for 3D Field Based Simulations. TP Kersten, ed., *Lösungen für eine Welt im Wandel, Dreiländertagung der SGPF, DGPF und OVG*. pp. 54–65.

[998] Chaturvedi K, Kolbe TH (2016): Integrating dynamic data and sensors with semantic 3D city models in the context of smart cities. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, IV-2/W1: 31–38.

[999] Krämer M, Haist J, Reitz T (2007): Methods for spatial data quality of 3D city models. *Eurographics Italian Chapter Conference*.

[1000] Akca D, Freeman M, Sargent I, Gruen A (2010): Quality assessment of 3D building data. *The Photogrammetric Record*, 25(132): 339–355.

[1001] Qin R, Tian J, Reinartz P (2016): 3D change detection – Approaches and applications. *ISPRS Journal of Photogrammetry and Remote Sensing*, 122: 41–56.

[1002] Steinhage V, Behley J, Meisel S, Cremers AB (2013): Reconstruction by components for automated updating of 3D city models. *Applied Geomatics*, 5(4): 285–298.

[1003] Iwaszczuk D, Stilla U (2016): Quality assessment of building textures extracted from oblique airborne thermal imagery. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, III-1: 3–8.

[1004] Rumpler M, Irschara A, Wendel A, Bischof H (2012): Rapid 3D City Model Approximation from Publicly Available Geographic Data Sources and Georeferenced Aerial Images. *Proceedings of the 17th Computer Vision Winter Workshop*. Mala Nedelja, Slovenia, pp. 1–8.

[1005] Khoshelham K, Zlatanova S (2016): Sensors for Indoor Mapping and Navigation. *Sensors*, 16(5): 655.

[1006] Díaz-Vilariño L, Boguslawski P, Khoshelham K, Lorenzo H, Mahdjoubi L (2016): Indoor navigation from point clouds: 3D modelling and obstacle detection. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLI-B4: 275–281.

[1007] Chen C, Yang BS, Song S (2016): Low cost and efficient 3D indoor mapping using multiple consumer RGB-D cameras. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLI-B1: 169–174.

[1008] Babacan K, Jung J, Wichmann A, Jahromi BA, Shahbazi M, Sohn G, Kada M (2016): Towards object driven floor plan extraction from laser point cloud. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLI-B3: 3–10.

[1009] Diakité AA, Zlatanova S (2016): First experiments with the Tango tablet for indoor scanning. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, III-4: 67–72.

[1010] Gunduz M, Isikdag U, Basaraner M (2016): A review of recent research in indoor modelling and mapping. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLI-B4: 289–294.

[1011] Amirebrahimi S (2016): *A Framework for Micro Level Assessment and 3D Visualisation of Flood Damage to a Building*. Ph.D. thesis, University of Melbourne, Melbourne, Victoria, Australia.

[1012] Amirebrahimi S, Rajabifard A, Mendis P, Ngo T (2016): A BIM-GIS integration method in support of the assessment and 3D visualisation of flood damage to a building. *Journal of Spatial Science*, 61(2): 317–350.

[1013] Besuievsky G, Beckers B, Patow G (2016): Skyline-control Based LoD Generation for Solar Analysis in 3D Cities. *Proceedings of FICUP – An international conference on urban physics*. Quito, Ecuador, pp. 1–13.

[1014] Matikainen L, Lehtomäki M, Ahokas E, Hyyppä J, Karjalainen M, Jaakkola A, Kukko A, Heinonen T (2016): Remote sensing methods for power line corridor surveys. *ISPRS Journal of Photogrammetry and Remote Sensing*, 119: 10–31.

[1015] Li Y, Fan J, Huang Y, Chen Z (2016): LiDAR-incorporated traffic sign detection from video log images of mobile mapping system. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLI-B1: 661–668.

[1016] Yao Q, Tan B, Huang Y (2016): Fast drawing of traffic sign using mobile mapping system. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLI-B3: 937–944.

[1017] Yu Y, Li J, Wen C, Guan H, Luo H, Wang C (2016): Bag-of-visual-phrases and hierarchical deep models for traffic sign detection and recognition in mobile laser scanning data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 113: 106–123.

[1018] Hosseinyalmdary S, Yilmaz A (2016): Traffic light detection using conic section geometry. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, III-1: 191–200.

[1019] Hui Z, Hu Y, Jin S, Yevenyo YZ (2016): Road centerline extraction from airborne LiDAR point cloud based on hierarchical fusion and optimization. *ISPRS Journal of Photogrammetry and Remote Sensing*, 118: 22–36.

[1020] Liu L, Lim S (2016): A framework of road extraction from airborne lidar data and aerial imagery. *Journal of Spatial Science*, 61(2): 263–281.

[1021] Elberink S, Khoshelham K (2015): Automatic Extraction of Railroad Centerlines from Mobile Laser Scanning Data. *Remote Sensing*, 7(5): 5565–5583.

[1022] Fang F, Im J, Lee J, Kim K (2016): An improved tree crown delineation method based on live crown ratios from airborne LiDAR data. *GIScience & Remote Sensing*, 53(3): 402–419.

[1023] Wu B, Yu B, Wu Q, Huang Y, Chen Z, Wu J (2016): Individual tree crown delineation using localized contour tree method and airborne LiDAR data in coniferous forests. *International Journal of Applied Earth Observation and Geoinformation*, 52: 82–94.

[1024] Kong F, Yan W, Zheng G, Yin H, Cavan G, Zhan W, Zhang N, Cheng L (2016): Retrieval of three-dimensional tree canopy and shade using terrestrial laser scanning (TLS) data to analyze the cooling effect of vegetation. *Agricultural and Forest Meteorology*, 217: 22–34.

[1025] Kutzner T, Kolbe TH (2016): Extending Semantic 3D City Models by Supply and Disposal Networks for Analysing the Urban Supply Situation. TP Kersten, ed., *Lösungen für eine Welt im Wandel, Dreiländertagung der SGPF, DGPF und OVG*. pp. 382–394.

[1026] Kumar K, Ledoux H, Stoter J (2016): A CityGML extension for handling very large TINs. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, IV-2-W1: 137–143.

ABSTRACT

The concept of level of detail (LOD) describes the content of 3D city models and it plays an essential role during their life cycle. On one hand it comes akin to the concepts of scale in cartography and LOD in computer graphics, on the other hand it is a standalone concept that requires attention. LOD has an influence on tendering and acquisition, and it has a hand in storage, maintenance, and application aspects. However, it has not been significantly researched, and this PhD thesis fills this void.

This thesis reviews dozens of current LOD standards, revealing that most practitioners consider the LOD to be comprised solely of the geometric detail of data and there are disparate views on the concept as a whole. However, the research suggests that the LOD encompasses additional metrics, such as semantics and texture. The thesis formalises the concept, enabling integration and comparison of current LOD standards. The established framework may be applied to cartography and to different forms of 3D geoinformation such as point clouds.

Following the formalised concept, a new LOD specification is presented improving the LOD concept in the current OGC CityGML 2.0 standard, a prominent norm in the 3D GIS industry. The specification introduces 16 LODs for buildings that are shaped after analysing the capabilities of acquisition techniques and a large number of real-world datasets. The improved LOD specification may be integrated in product portfolios and tenders, preventing misunderstandings between stakeholders, and as a better language for communicating the specifics of a dataset to be acquired. The specification also considers different approaches to realise the data. Such geometric references result in dozens of different variants of the same LOD.

3D data according to the LOD specification was generated using a procedural modelling engine that was developed over the course of the research. The engine is capable of producing 3D city models in a large number of different variants and according to the CityGML standard.

The thesis also catalogues the many different ways to create 3D city models. A prominent technique for producing data in a different LOD is generalisation, i.e. simplifying a 3D city model. The inverse—augmenting the LOD of a dataset—has not been researched to a great extent, and this thesis gives an overview of the topic. This research demonstrates that it is possible to generate 3D city models without

elevation measurements, inherently augmenting the LOD of coarser data (2D footprints). The method relies on machine learning: several attributes found in 2D datasets may hint at the height of a building, thus enabling extrusion and creating 3D city models suited for several applications.

Some acquisition techniques may result in multi-LOD datasets, and nowadays there are some regions represented in different, independent datasets. However, it was found that possibilities to link such data are deficient. The lack of linking mechanisms inhibits acquisition, storage, and maintenance of multi-LOD data. Two methods for linking features across two or more LODs have been developed resulting in an increased consistency of multi-LOD datasets. The first method links matching geometries across multiple LODs, while the second method establishes a 4D data structure in which the LOD is modelled as the fourth (spatial) dimension.

It is often believed that the more detailed 3D data the better. However, similarly as in computer graphics, dealing with data at fine LODs comes at a cost: such datasets are harder to obtain, their storage footprint is large, and their usage within a spatial analysis may be slow. Scarce research has been dedicated to investigating whether an increase in the LOD of the data brings a comparably significant increase in benefits when the data is used in a spatial analysis.

First, an analysis using real-world multi-LOD data was carried out. Different LODs of spatial data covering the Netherlands was used in a spatial analysis to refine population maps, obtaining different results for each LOD. However, several problems are exposed, revealing that using real data for such investigations is not optimal.

The remainder of the research focuses on using procedurally generated data for such experiments. Synthetic data in several different LODs has been generated and employed for four spatial analyses (estimation of the building shadow, envelope area, volume, and solar irradiation). The experiments result in different conclusions. Finer LODs usually bring some improvement to the quality of the spatial analysis, but not always and such may be negligible. The results of the experiments ultimately depend on the spatial analysis that is considered. The varying results between different spatial analyses make each of them unique. Furthermore, the benefit a finer LOD brings to a spatial analysis is not always clear and easily measurable. In short, striving to produce data at finer LODs may please the eye, but this is not always counter-balanced in the benefit it brings to a spatial analysis.

A further addition to the equation above is that when realised, 3D city models are unavoidably burdened with acquisition errors. An error propagation analysis was performed by disturbing the procedurally generated datasets with a range of simulated positional errors. Comparisons have been made between the intentionally degraded datasets and their error-free counterparts, thus obtaining the magnitude of uncertainty the positional errors cause in a spatial analysis. Based on these ex-

periments, several findings are discovered, most importantly:

- How the LODs are realised (which geometric references are used) has a larger influence than the LOD. A coarse LOD produced with a favourable geometric reference may yield better results than a finer LOD realised with an unfavourable reference.
- Positional errors considerably affect spatial analyses. The effect is comparable across similar LODs. Simpler LODs are sligthly less affected by positional errors, but they may contain a large systematic error.
- Errors induced in the acquisition process generally cancel out the improvement provided by finer LODs. The main conclusion is that in the considered spatial analyses the positional error has a significantly higher impact than the LOD. As a consequence, it is suggested that it is pointless to acquire geoinformation at a fine LOD if the acquisition method is not accurate, and instead it is advised to focus on the improvement of accuracy of the data.

The thesis proposes additional research for future work. For example, since this research focuses specifically on 3D building models, it would be worth extending the research to other urban features such as roads and vegetation. Furthermore, quality control in 3D GIS does not encompass the evaluation of the LOD of data. Hence integration of the LOD in quality standards should be a priority for future work.

SAMENVATTING

'Level of Detail' (LOD), oftewel het detailniveau van 3D data, beschrijft de inhoud van 3D stadsmodellen en speelt een belangrijke rol bij de verschillende stappen van de 3D informatieketen (van inwinning tot gebruik). Het LOD van 3D stadsmodellen heeft impact op de aanbesteding en inwinning ervan, alsmede op de opslag, bijhouding en het gebruik van de 3D data. Aan de ene kant is het concept LOD verwant met schaal in de cartografie en LOD in computer graphics, aan de andere kant is het een op zichzelf staand concept dat afzonderlijke aandacht verdient. Desondanks is het concept LOD van 3D stadsmodellen nog nooit goed bestudeerd en dit PhD onderzoek vult dit gat.

Dit onderzoek analyseert als eerste bestaande LOD standaarden. Dit levert het inzicht dat de meeste professionals LOD beschouwen als een concept dat alleen het geometrische detailniveau beschrijft, terwijl er uiteenlopende meningen zijn over de definitie van het LOD concept als geheel. Dit onderzoek laat ook zien dat het LOD aanvullende aspecten beschrijft dan alleen het geometrisch detailniveau, zoals semantiek en textuur. Dit onderzoekt formuleert vervolgens formele definities voor het LOD concept van 3D data waardoor het mogelijk is verschillende LOD standaarden te integreren en vergelijken. Het resulterende formele kader kan worden toegepast op cartografische representatie maar ook op verschillende vormen van 3D geoinformatie zoals puntenwolken.

Op basis van het geformaliseerde LOD concept, wordt een nieuwe LOD specificatie voorgesteld. Dit verbetert het LOD concept in de huidige OGC CityGML 2.0 standaard, een prominente standaard in de 3D GIS wereld. De specificatie introduceert 16 LODs voor gebouwen welke zijn vastgesteld na het analyseren van zowel de mogelijkheden van 3D data inwinning als een groot aantal bestaande 3D datasets. De verbeterde LOD specificatie kan worden gebruikt voor productspecificaties en in aanbestedingsteksten zodat misverstanden tussen de verschillende betrokkenen (zoals opdrachtgevers en uitvoerders) wordt voorkomen. Zo biedt de LOD specificatie een betere manier om te communiceren over de precieze inhoud van de 3D data welke moet worden ingewonnen.

De specificaties beschrijft ook hoe 3D data kan worden gegenereerd, waarbij eenzelfde LOD verschillende varianten kan hebben (geometrische referenties). Om de

3D data vervolgens als zodanig te genereren is een procedurele modelleeromgeving ontwikkeld. Deze engine kan 3D stadsmodellen genereren in een groot aantal LOD varianten die voldoen aan de CityGML standaard.

Dit proefschrift beschrijft ook de vele verschillende manieren om 3D stadsmodellen te genereren. Een prominente techniek om data in verschillende LODs te genereren is generalisatie, dat wil zeggen het versimpelen van een 3D stadsmodel. Het omgekeerde—het toevoegen van meer detail zodat een meer gedetailleerd LOD wordt verkregen—is nog nauwelijks bestudeerd en dit onderzoek geeft een overzicht van dit onderwerp. Het onderzoek laat ook zien dat het mogelijk is om 3D stadsmodellen te genereren door 3D detail toe te voegen aan 2D footprints van gebouwen, zonder dat daarbij hoogte gegevens worden gebruikt. De methode is gebaseerd op "machine-learning" technieken: er zijn vaak meerdere attributen in een 2D dataset die een indicatie geven over de hoogte van een gebouw. Hiermee kunnen de footprints van gebouwen worden opgetrokken tot 3D stadsmodellen.

Verschillende technieken om 3D stadsmodellen te genereren leiden tot afzonderlijke datasets op verschillende detailniveau en er zijn dan ook gebieden waar meerdere 3D datasets van beschikbaar zijn op verschillende LODs die verder geen verband met elkaar hebben. Dit geeft problemen. Omdat er geen mechanisme bestaat om het verband vast te leggen tussen de verschillende LODs, is het inwinnen, de opslag en het beheer van consistente multi-LOD data immers niet mogelijk. In dit onderzoek zijn daarom twee methoden ontwikkeld om objecten op twee of meerdere LODs met elkaar te verbinden ten einde consistentie tussen verschillende LODs te bewerkstelligen. De eerste methode verbindt met elkaar corresponderende geometrieën; de tweede methode realiseert een 4D datastructuur waarbij het LOD wordt gemodelleerd als vierde dimensie.

Vaak wordt gedacht dat hoe gedetailleerder het LOD van 3D data, hoe beter de data. Maar net zoals in computer graphics heeft het omgaan met gedetailleerde LODs ook een keerzijde: het is moeilijker de data in te winnen, het vereist veel data opslag en het gebruik van gedetailleerde LOD data in ruimtelijke analyses kan traag zijn. Aan de andere kant is er nog maar weinig onderzoek gedaan of een gedetailleerder LOD daadwerkelijk betere resultaten oplevert als het wordt gebruikt in ruimtelijke analyses.

Daarom is er eerst een analyse uitgevoerd met bestaande landsdekkende multi-LOD data sets van Nederland. Deze zijn gebruikt in een ruimtelijke analyse om bevolkingskaarten te verfijnen waarbij verschillende LODs verschillende resultaten opleverden. De bestaande 3D data sets waren echter niet geschikt om precies inzicht te geven in deze verschillen. Daarom is in een volgende fase van het onderzoek gebruik gemaakt van procedureel gegenereerde 3D data. Syntactische data op verschillende LODs zijn gegenereerd en gebruikt voor vier ruimtelijke analyses: bepaling van schaduw, de omtrek en het volume van gebouwen en van zonne-instraling.

Uit deze experimenten kunnen verschillende conclusies worden getrokken. Meer gedetailleerde LODs geven vaak betere resultaten van ruimtelijke analyses. Maar dit is lang niet altijd zo en de verbeteringen zijn soms minimaal. De resultaten van de experimenten zijn uiteindelijk afhankelijk van de specifieke ruimtelijke analyse. De variatie in resultaten is voor iedere ruimtelijke analyse weer anders. Bovendien is de meerwaarde van een gedetailleerder LOD bij een specifieke analyse niet altijd duidelijk en meetbaar. Kortom: het streven naar een zo hoog mogelijk detailniveau kan daarom visueel aantrekkelijk zijn maar levert niet altijd significant betere resultaten op van ruimtelijke analyses.

Naast dat het LOD de kwaliteit van de resultaten van ruimtelijke analyses beïnvloedt, wordt deze kwaliteit ook beïnvloed door fouten geïntroduceerd bij de inwinning. Daarom is een foutenvoortplanting-analyse uitgevoerd waarbij verschillende test datasets procedureel zijn gegenereerd op basis van gesimuleerde positionele fouten. Vervolgens zijn de uitkomsten van ruimtelijke analyses van deze met opzet verslechterde 3D datasets vergeleken met de uitkomsten op basis van de oorspronkelijke data. Dit gaf inzicht in de mate van onzekerheid die positionele fouten geven in ruimtelijke analyses. De belangrijkste inzichten hier zijn:

- Welke geometrische referentie wordt gekozen voor een specifiek LOD heeft een grotere invloed dan het gekozen LOD zelf. Een minder gedetailleerd LOD met een meer geschikte geometrische referentie kan betere resultaten opleveren dan een meer gedetailleerd LOD met minder geschikte geometrische referentie.

- Positionele fouten hebben een aanzienlijk effect op de resultaten van ruimtelijke analyses. Dit effect is vergelijkbaar voor varianten van dezelfde LODs (bijvoorbeeld LOD2.1 en LOD2.2). Het effect van positionele fouten is kleiner bij minder gedetailleerde LODs, maar deze kunnen weer een grote systematisch fout hebben.

- Fouten geïntroduceerd bij de inwinning doen over het algemeen de verbeteringen die een meer gedetailleerde LOD geeft teniet. Een positionele fout heeft namelijk een veel grotere invloed op de kwaliteit dan het gekozen LOD. Daarom is het zinloos om geoinformatie op een hoger detailniveau in te winnen als de inwinmethode niet precies is. Het wordt daarentegen geadviseerd om te focussen op een hogere precisie van de data.

Deze thesis beschrijft ook toekomstig onderzoek dat nodig is om de resultaten verder te verbeteren. Zoals het bestuderen van het LOD van andere objecten dan gebouwen (in dit onderzoek lag de focus op gebouwen). Een ander onderwerp dat meer onderzoek behoeft is de integratie van LOD in kwaliteitsstandaarden omdat

huidige kwaliteitscontrole in 3D GIS geen onderscheid maakt tussen verschillende LODs van de data.

CURRICULUM VITÆ

Filip Biljecki (1987) was born in Zagreb, Croatia, and spent his childhood in Lugano, Switzerland. He obtained a BSc in Geodesy and Geoinformatics from the University of Zagreb, and an MSc in Geomatics from the Delft University of Technology. During his studies he was presented with prizes such as the Dean's and Rector's awards. He graduated with an MSc thesis that introduced a method to automatically segment and classify GPS tracks according to the used transportation mode, consisting of a fuzzy expert system that relies on geographic information. The results of the thesis were published in the *International Journal of Geographical Information Science*.

During the studies, Filip was fortunate enough to gain valuable industry experience with internships in Hansa Luftbild (Germany) and Blom (United Kingdom). Following his graduation, he continued to expand his industry knowledge while working for Geofoto in Croatia as an associate in business development.

In 2012 he was appointed as a PhD candidate at the Delft University of Technology under the supervision of Prof. dr. Jantien Stoter and Dr. Hugo Ledoux working on a project funded by the Netherlands Organisation for Scientific Research (NWO). The research resulted in software and multiple papers in international journals and conferences (pp. 350–353), with an outreach spanning several countries. During the course of his research work he has also been involved in teaching MSc courses, supervising MSc theses, and participation in international collaborations, such as in the OGC CityGML SWG, EuroSDR 3D SIG, and ISPRS Commission IV. In 2016, he carried out a research visit at the Singapore Land Authority, focusing on the quality control of 3D city models. In the same year he was awarded the Young Researcher Award in GIScience by the Austrian Academy of Sciences.

During his PhD research, Filip has served on the programme committees of several international conferences, and assisted as a reviewer for leading international GIS journals such as CEUS, TGIS, and IJGIS, as well as journals in other disciplines (e.g. *Scientometrics* and *Automation in Construction*). He has organised the Eurographics Workshop on Urban Data Modelling and Visualisation in the Netherlands, and co-chaired its subsequent edition in Belgium.

LIST OF PUBLICATIONS

INTERNATIONAL JOURNAL PAPERS (ISI)

1. **Biljecki F**, Ledoux H, Van Oosterom P (2013): Transportation mode-based segmentation and classification of movement trajectories. *International Journal of Geographical Information Science*, 27(2): 385–407. doi: 10.1080/13658816.2012.692791

2. **Biljecki F**, Ledoux H, Stoter J, Zhao J (2014): Formalisation of the level of detail in 3D city modelling. *Computers, Environment and Urban Systems*, 48: 1–15. doi: 10.1016/j.compenvurbsys.2014.05.004

3. Arroyo Ohori K, Ledoux H, **Biljecki F**, Stoter J (2015): Modeling a 3D City Model and Its Levels of Detail as a True 4D Model. *ISPRS International Journal of Geo-Information*, 4(3): 1055–1075. doi: 10.3390/ijgi4031055

4. Boeters R, Arroyo Ohori K, **Biljecki F**, Zlatanova S (2015): Automatically enhancing CityGML LOD2 models with a corresponding indoor geometry. *International Journal of Geographical Information Science*, 29(12): 2248–2268. doi: 10.1080/13658816.2015.1072201

5. **Biljecki F**, Heuvelink GBM, Ledoux H, Stoter J (2015): Propagation of positional error in 3D GIS: estimation of the solar irradiation of building roofs. *International Journal of Geographical Information Science*, 29(12): 2269–2294. doi: 10.1080/13658816.2015.1073292

6. **Biljecki F**, Stoter J, Ledoux H, Zlatanova S, Çöltekin, A (2015): Applications of 3D City Models: State of the Art Review. *ISPRS International Journal of Geo-Information*, 4(4): 2842–2889. doi: 10.3390/ijgi4042842

7. Van Winden K, **Biljecki F**, Van der Spek, S (2016): Automatic Update of Road Attributes by Mining GPS Tracks. *Transactions in GIS*, 20(5): 664–683. doi: 10.1111/tgis.12186

8. **Biljecki F** (2016): A scientometric analysis of selected GIScience journals. *International Journal of Geographical Information Science*, 30(7): 1302–1335. doi: 10.1080/13658816.2015.1130831

9. **Biljecki F**, Ledoux H, Stoter J, Vosselman G (2016): The variants of an LOD of a 3D building model and their influence on spatial analyses. *ISPRS Journal of Photogrammetry and Remote Sensing*, 116: 42–54. doi: 10.1016/j.isprsjprs.2016.03.003

10. **Biljecki F**, Ledoux H, Stoter J (2016): An improved LOD specification for 3D building models. *Computers, Environment and Urban Systems*, 59: 25–37. doi: 10.1016/j.compenvurbsys.2016.04.005

11. **Biljecki F**, Arroyo Ohori K, Ledoux H, Peters R, Stoter J (2016): Population Estimation Using a 3D City Model: A Multi-Scale Country-Wide Study in the Netherlands. *PLOS ONE*, 11(6): e0156808. doi: 10.1371/journal.pone.0156808

12. **Biljecki F**, Ledoux H, Stoter J (2017): Generating 3D city models without elevation data. *Computers, Environment and Urban Systems*, 64: 1–18. doi: 10.1016/j.compenvurbsys.2017.01.001

13. **Biljecki F**, Heuvelink GBM, Ledoux H, Stoter J (2017): The effect of acquisition error and level of detail on the accuracy of spatial analyses. *Cartography and Geographic Information Science*, advance online publication. doi: 10.1080/15230406.2017.1279986

PEER-REVIEWED BOOK CHAPTERS

14. **Biljecki F**, Ledoux H, Stoter J (2015): Improving the consistency of multi-LOD CityGML datasets by removing redundancy. M Breunig, AD Mulhim, E Butwilowski, PV Kuper, J Benner, KH Häfele, eds., *3D Geoinformation Science*, Springer, pp. 1–17. doi: 10.1007/978-3-319-12181-9_1

15. Stoter J, Ledoux H, Zlatanova S, **Biljecki F** (2016): Towards sustainable and clean 3D Geoinformation. TH Kolbe, R Bill, A Donaubauer, eds., *Geoinformationssysteme 2016: Beiträge zur 3. Münchner GI-Runde*, Wichmann Herbert, pp. 100–113.

16. **Biljecki F**, Ledoux H, Stoter J (2017): Does a Finer Level of Detail of a 3D City Model Bring an Improvement for Estimating Shadows? A Abdul-Rahman, ed., *Advances in 3D Geoinformation*, Springer, pp. 31–47. doi: 10.1007/978-3-319-25691-7_2

INTERNATIONAL REFEREED CONFERENCES

17. Arroyo Ohori K, **Biljecki F**, Stoter J, Ledoux H (2013): Manipulating higher dimensional spatial information. *16th AGILE Conference on Geographic Information Science*, pp. 1–7.
doi: 10.4233/uuid:c4751ad5-e008-48a6-806b-af38099e0b3a

18. **Biljecki F**, Zhao J, Stoter J, Ledoux H (2013): Revisiting the concept of level of detail in 3D city modelling. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, II-2/W1: 63–74. doi: 10.5194/isprsannals-II-2-W1-63-2013

19. **Biljecki F**, Ledoux H, Stoter J (2014): Error propagation in the computation of volumes in 3D city models with the Monte Carlo method. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, II-2: 31–39.
doi: 10.5194/isprsannals-II-2-31-2014

20. **Biljecki F**, Ledoux H, Stoter J (2014): Height references of CityGML LOD1 buildings and their influence on applications. *9th 3D GeoInfo Conference 2014*. doi: 10.4233/uuid:09d030b5-67d3-467b-babb-5e5ec10f1b38

21. Aarsen R, Janssen M, Ramkisoen M, **Biljecki F**, Quak W, Verbree E (2015): Installed base registration of decentralised solar panels with applications in crisis management. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XL-3/W3: 219–223. doi: 10.5194/isprsarchives-XL-3-W3-219-2015

22. Peters R, Ledoux H, **Biljecki F** (2015): Visibility Analysis in a Point Cloud Based on the Medial Axis Transform. *Third Eurographics Workshop on Urban Data Modelling and Visualisation*, pp. 7–12. doi: 10.2312/udmv.20151342

23. **Biljecki F**, Arroyo Ohori K (2015): Automatic Semantic-preserving Conversion Between OBJ and CityGML. *Third Eurographics Workshop on Urban Data Modelling and Visualisation*, pp. 25–30. doi: 10.2312/udmv.20151345

24. **Biljecki F**, Ledoux H, Stoter J (2016): Generation of multi-LOD 3D city models in CityGML with the procedural modelling engine Random3Dcity. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, IV-4/W1: 51–59.
doi: 10.5194/isprs-annals-IV-4-W1-51-2016

25. Löwner MO, Gröger G, Benner J, **Biljecki F**, Nagel C (2016): Proposal for a new LOD and multi-representation concept for CityGML. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, IV-2/W1: 3–12.
doi: 10.5194/isprs-annals-IV-2-W1-3-2016

26. Rook M, **Biljecki F**, Diakité AA (2016): Towards automatic semantic labelling of 3D city models. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, IV-2/W1: 23–30. doi: 10.5194/isprs-annals-IV-2-W1-23-2016

27. **Biljecki F**, Ledoux H, Du X, Stoter J, Soon KH, Khoo VHS (2016): The most common geometric and semantic errors in CityGML datasets. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, IV-2/W1: 13–22. doi: 10.5194/isprs-annals-IV-2-W1-13-2016

28. Stoter J, Ploeger H, Roes R, Van der Riet E, **Biljecki F**, Ledoux H (2016): First 3D Cadastral Registration of Multi-level Ownerships Rights in the Netherlands. *5th International FIG Workshop on 3D Cadastres*, pp. 491–504.

MISCELLANEOUS

29. **Biljecki F**, Ledoux H, Stoter J (2014): Redefining the Level of Detail for 3D models. *GIM International*, 28(11): 21–23.

30. OGC (2016): OGC CityGML Quality Interoperability Experiment (OGC 16-064). H Ledoux, D Wagner, eds. (**contributor**)


9 789461 868008

TU Delft