GI-Edition

Lecture Notes in Informatics

Volker Wohlgemuth, Stefan Naumann, Hans-Knud Arndt, Grit Behrens, Maximilian Höb (Editors)

EnviroInfo 2022

26.–28. September 2022
Hamburg, Germany

Short- /Work in Progress-Papers

328

Volker Wohlgemuth et al. (Hrsg.) EnviroInfo 2022

GESELLSCHAFT
FÜR INFORMATIK

Volker Wohlgemuth, Stefan Naumann, Hans-Knud Arndt, Grit Behrens, Maximilian Höb (Editors)

EnviroInfo 2022
Short-/Work in Progress-Papers

September 26-28, 2022
Hamburg, Germany

Gesellschaft für Informatik e.V. (GI)

Lecture Notes in Informatics (LNI) - Proceedings
Series of the Gesellschaft für Informatik (GI)

Volume P-328

ISBN 978-3-88579-722-7
ISSN 1617-5468

printed by Köllen Druck+Verlag GmbH, Bonn

Preface

This book presents the main research results of the 36th edition of the long-standing and established international and interdisciplinary conference series on environmental information and communication technologies (EnviroInfo 2022).

The conference was held from 26 to 28 of September 2022. It was organized under the patronage of the Technical Committee on Environmental Informatics of the Gesellschaft für Informatik e.V. (German Informatics Society - GI)

This book presents a selection of peer-reviewed research papers that describe innovative scientific approaches and ongoing research in environmental informatics and the emerging field of environmental sustainability. Combining and shaping national and international activities in the field of applied informatics and Environmental Informatics, the EnviroInfo conference series aims at presenting and discussing the latest state-of-the-art development on information and communication technology (ICT) and environmental-related fields. A special focus of the conference focuses on the question how Environmental Informatics can help to reach the Sustainable Development Goals of the UN and which goals are explicitly addressed by the Environmental Informatics Community.

Fig. 1: Number of contributions in this book that can be assigned to a sustainability development goal of the United Nations (UN)

The respective articles cover a broad range of scientific aspects including advances in core environmental informatics-related technologies, such as Artificial Intelligence Application, Sustainable Mobility, Green Coding, ICT and Circular Economy and other relevant topics in the field of Environmental Informatics.

We would like to thank all contributors for their submissions. Special thanks also go to the members of the program and organizing committees, for reviewing all submissions and Franziska Mai and Till Börsch for layouting and organizing the submissions in this book. In particular, we like to thank our local organizers from University of Hamburg for the local organizing support.

Last but not least, a warm thank you to our sponsors that supported the conference.

Berlin	Volker Wohlgemuth
Magdeburg	Hans-Knud Arndt
Bielefeld	Grit Behrens
Birkenfeld	Stefan Naumann
München	Maximilian Höb

Hamburg, September 2022

Sponsoren

DISY Informationssysteme GmbH

General Chairs

Volker Wohlgemuth	HTW Berlin, University of Applied Sciences, Berlin, Germany
Stefan Naumann	Umwelt-Campus Birkenfeld – Trier University of Applied Sciences, Germany
Hans-Knud Arndt	Otto-von-Guericke-University Magdeburg, Germany
Grit Behrens	University of Applied Sciences Bielefeld, Germany
Maximilian Höb	Leibniz Supercomputing Centre (LRZ), Munich

Programme Committee

Achim Guldner	Umwelt-Campus Birkenfeld – Trier University of Applied Sciences, Germany
Amna Ramzy	German University in Cairo (GUC), Egypt
Anna Kramers	KTH Royal Institute of Technology, Sweden
Andreas Abecker	Disy Informationssysteme GmbH, Germany
Andreas Kamilaris	CYENS Center of Excellence, Cyprus
Andreas Möller	Leuphana University Luneburg, Germany
Andreas Winter	University of Oldenburg, Germany
Barak Fishbain	Technion – Israel Institute of Technology, Israel
Benjamin Wagner vom Berg	University of Applied Sciences Bremerhaven, Germany
Benno Schmidt	Bochum University of Applied Sciences
Benoît Otjacques	Luxemburg Institute of Science and Technology (LIST), Luxembourg
Christian Bunse	Stralsund University of Applied Sciences, Germany
Daniel Eberz-Eder	Dienstleistungszentrum ländlicher Raum (DLR), Rheinhessen-Nahe-Hunsrück, Germany
Dieter Kranzlmüller	Leibniz Supecomputing Centre (LRZ) of the Bavarian Academy of Sciences and Humanities, Germany
Eva Kern	Umwelt-Campus Birkenfeld – Trier University of Applied Sciences, Germany
Frank Fuchs-Kittowski	Hochschule für Technik und Wirtschaft Berlin (HTW

Table of Contents

Applied Environmental Informatics

Artificial Intelligence Applications

Circular Economy and Industrial Symbiosys

Green Coding

Green Data Centres

Sustainable Mobility

APPLIED ENVIRONMENTAL INFORMATICS

Cloud-based Processing of data from Non-Target-Analysis for Tracking Micropollutants in Surface Water

Viktoria Pauw[1], Mohamad Hayek[1], Elham Shojaei[1], Stephan Hachinger[1], Uwe Müller[1] and Tobias Bader[1]

Abstract: Tens of thousands of chemicals used by consumers, agriculture and industry enter the aquatic environment as micropollutants every day. Using targeted analysis we are so far only able to detect a small subset of the chemicals that are present. Therefore so called non-target screening (NTS) using liquid chromatography in combination with high-resolution mass spectrometry (LC-HRMS) is increasingly used by labs to perform more comprehensive monitoring. However, a high degree of variance in measurements and processing workflows results in low comparability of data from separate laboratories. On one hand this is caused by differences in processing techniques which are due to stationary laboratory equipment and on the other hand by differing priorities in the detection strategy and evaluation workflow. The K2I project funded by BMBF aims at fostering collaboration between laboratories by providing a joint platform for uploading and processing LC-HRMS data. A cloud based datalake and processing pipeline is being developed. A standardized processing workflow can then be executed which is enhanced by data mining tools including machine learning techniques. An indexing and searching software is employed to create a web based access to the processed data for participants.

Keywords: Non-target screening; LC-HRMS; Water Safety; AI; Cloud; Environmental Chemistry

Addresses Sustainable Development Goal 6: Clean water and sanitation

1. Motivation

River water is a major source of drinking water, is used in food production and industry and therefore assuring the absence of potentially harmful chemicals is a vital issue for environment and economic sustainability. Pollutants in the aquatic environment can affect aquatic organism, accumulate in sediment soil and potentially pose health risks to humans in case of chronic exposure [Sy21]. Not all chemicals can be removed by standard water treatment processes and only for a small subset, targeted tests can be carried out on a regular basis.

Non-target screening using liquid chromatography coupled to high-resolution mass spectrometry (LC-HRMS) allows the detection of thousands of signals from different substances at the same time (comp. [Ho17]), covering any substance that is successfully ionized and present in a sufficiently large quantity to pass the detection threshold [Ba17].

[1] Leibniz-Rechenzentrum, Boltzmannstraße 1, 85748 Garching, Germany, viktoria.pauw@lrz.de

To identify unknown substances in water samples, the results of a LC-HRMS measurement can be compared to lists of known chemical compounds. The measurements consist of signal peaks that correspond to a specific retention time (RT) in the chromatograph and have a certain mass to charge ratio *(m/z)* which is determined by the mass spectrometer after being ionized. These peaks can be scanned for unknown combinations of RT and m/z, indicating the presence of unregistered substances and for strong or recurring signals that cannot be attributed to a known cause. It can be advantageous to narrow down the source of emission to compare measurements taken at different locations, both in the same water body (different sites at the same river or lake) and in separate waters. Thus, the combination of LC-HRMS data from different sampling locations and laboratories, which could be combined with spatial and temporal coordinates and additional information such as known environmental influences could be used to more effectively notice and track micropollutants in surface water across larger regions.

2. Project Outline

In the project K2I (https://www.k2i-tracker.de), several German research institutions and water providers collaborate to improve the situation for joint water monitoring on a transregional scale. Prior work was carried out at Landeswasserversorgung (LW) in Baden-Würtemberg to test the feasibility of data mining on LC-HRMS measurements [Ba18].

2.1 Partners

Landeswasserversorgung (LW) is one of the oldest and largest water providers of Germany, delivering 100 mio cubic meters of drinking water to three million people in Bavaria and Baden-Würtemberg. Technologiezentrum Wasser (TZW) is an independent, research institute covering all fields of drinking water supply. The Technical University of Munich contributes research on markers of micro-biological contamination (like bacteria or viruses). Associated partners providing data and feedback are laboratories of large public water supply companies such as Hessenwasser (HW), Hamburg Wasser (HH), Westfälische Wasser- und Umweltanalytik (WWU), BodenseeWasserversorgung (BWV). In order to unify data processing between the labs for project purposes and beyond, and make available sufficient computing power for data processing, the Leibniz Supercomputing Centre (LRZ, Garching b.M., Bavaria) has been included in the project consortium.

2.2 ETL Pipeline, Storage and Compute Resource

The unified processing approach mentioned in Section 2.1 requires a modern computing infrastructure. To this purpose, an experimental but reliable Kubernetes cluster at LRZ

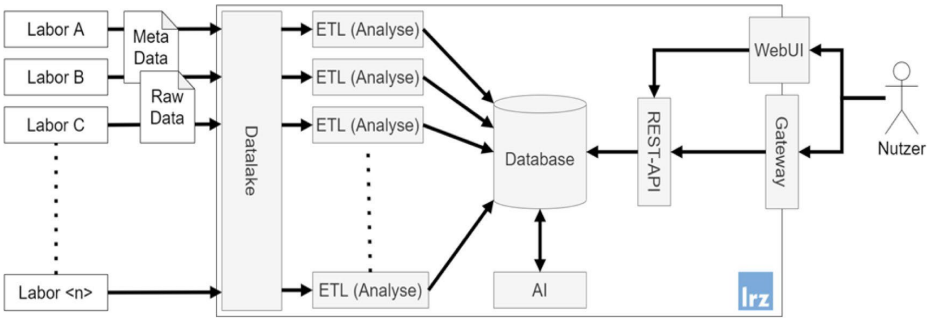


Fig. 1: Sketch of K2I processing set-up including cloud-storage and ETL (Extract Transform Load) pipeline

with about 20 virtual CPU cores (3 worker nodes) is used. This cluster can flexibly be extended as the project progresses. It uses the NFS-based part of the LRZ Cloud Storage as a storage backend, with a current usage in the 10 TB range.

The processing software enviMass v4.4 [Lo19] is used in the workflow for pre- processing steps like peak picking, RT-alignment between samples, homologue series detection. The company envibee (https://www.envibee.ch) which is maintaining and ex- panding the code is a partner in development.

The different labs access the storage resources available via a deployment of MinIO in Gateway mode. The storage space in MinIO is structured into two different entities, one for the upload of samples and their metadata and another for the output produced by enviMass. To automate the processing of the input data, a containerized version of enviMass was developed with envibee and included in the ETL Pipeline (see Fig. 1). In front of enviMass, a RESTful API is provided and used by the pipeline to trigger different actions within the enviMass container. Once the execution of a data project is concluded in enviMass, the output is available through MinIO. All blocks within the pipeline are containerized and run on kubernetes. This allows to have all the services in high availability and scalable mode. The use of containers and an orchestration software aims at providing a portable solution that can be re-deployed on different architectures and within a continuous integration model with little overhead.

2.3 Datalake and Data Mining

The processed data from enviMass output is then fed into a database, which can be accessed by the users through a searching tool with a web interface. The current prototype uses ElasticSearch (https://www.elastic.co) and Grafana (https://grafana.com) as a front end. Automized evaluation scripts using data mining and machine learning techniques will access the database and provide summaries and reports. Obligatory metadata that has to

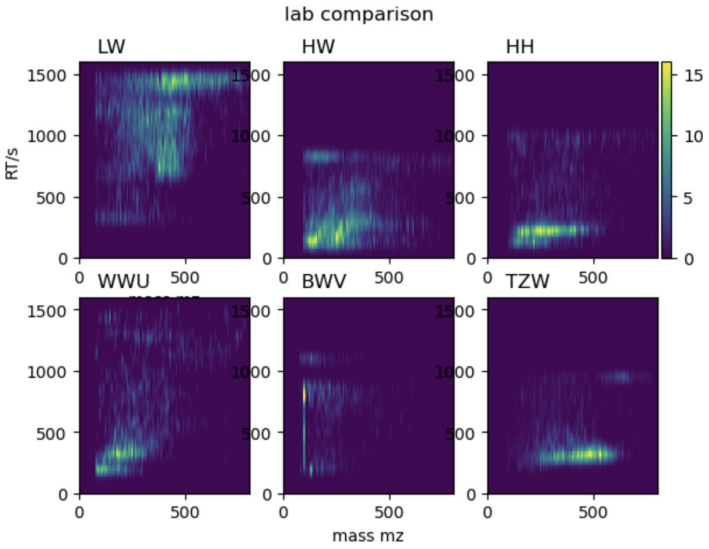


Fig. 2: Different Retention Time scales for different labs

be provided during the upload consists of the name of the laboratory, a sample identifier, resource type (sample or blank), publisher (uploading person), publication year, date (day when sample was taken), geo location (coordinates of sampling location, longitude and latitude), matrix (type of water body), and polarity (positive or negative ionization). Optional specifications to aid in-depth analysis are: Impact source (potential known influences, e.g. agriculture, industry) and discharge (volumetric flow rate of river water).

3. Challenges and Strategies

3.1 Interlab Alignment

A major obstacle in directly comparing and combining data processed in different laboratories are deviations in the measured retention time due to varying chromatographic equipment. While only C18 columns where used for separation before mass spectrometry, manufacturers and modifications differ, resulting in varying offsets in retention timescales and aberrations in precision (cf. Fig 2). [Wa22].

In [Ya18] an alignment algorithm is developed based on Subwindow Factor Analysis and Mass Spectral information (SFA-MS). [Ot18] developed a package called GCalignR based on retention time to align the data in biology. It used GC-MS (Gas Chromatography Mass Spectrometry) and FID (flame ionization detection) to characterize the chemical composition. [Fu17] employed a peak alignment strategy based on the mass spectra and

retention times of the peaks in which the maximum mass spectral correlation coefficient path was extracted using a modified dynamic programming method. It has been reported that due to complexity, all existing developed software runs into difficulties and need to be improved. Researchers are recommended to check the alignment of important biomarkers manually [VL13] and [Ko10]. These algorithms are developed based on data from the same laboratory, which experience fewer harmonization problems (chromatographic characteristics such as column, eluents, gradients are the same).

Our current approach is to use data-sets that were pre-processed by enviMass which executes inta-lab alignment, then identifying common features, like intentionally added compounds called internal standards (ISTDs), or particularly easily discernible known substances and using parameter estimation to achieve a pairwise fit of the retention time values for these supporting points.

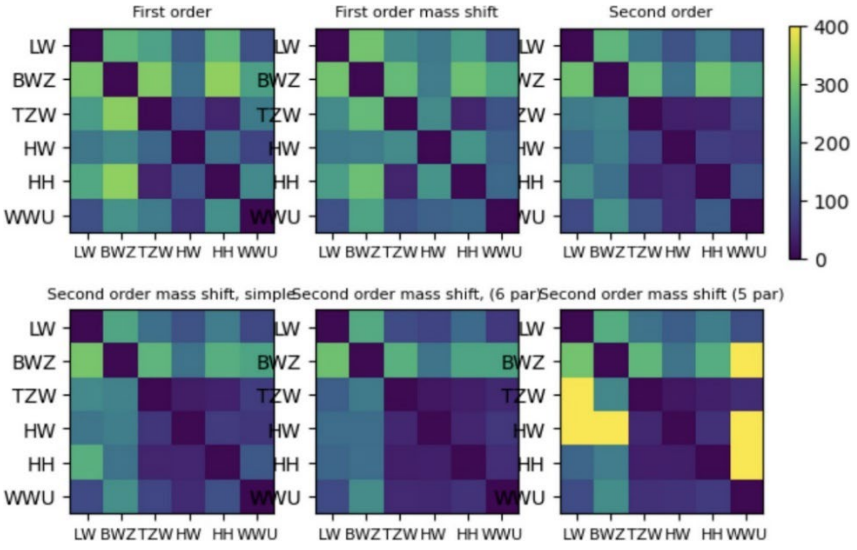


Fig. 3: Heatmap of pairwise alignment fits between different labs using internal standards. Color: Sum of the square error in RT for all ISTDs

Using a polynomial for a one-dimensional fit of the retention times from different labs, a second order fitting function gave best results. The variance among data-sets in RT for the ISTDs could be reduced to be always lower than 70 seconds. For some lab pairings, this could be improved further by making the fitting parameters mass dependent (Fig 3), as different molecular mass ranges showed different trends for the shift in retention times (Fig 4), with very high retention times often being more difficult to fit. Another complicating effect is the so called injection peak, where eluent ions are displayed by the injected sample and many different substances accumulate with the same low retention times. The removal of these fringe areas of the chromatographic measurements is necessary to achieve an accurate RT correlation.

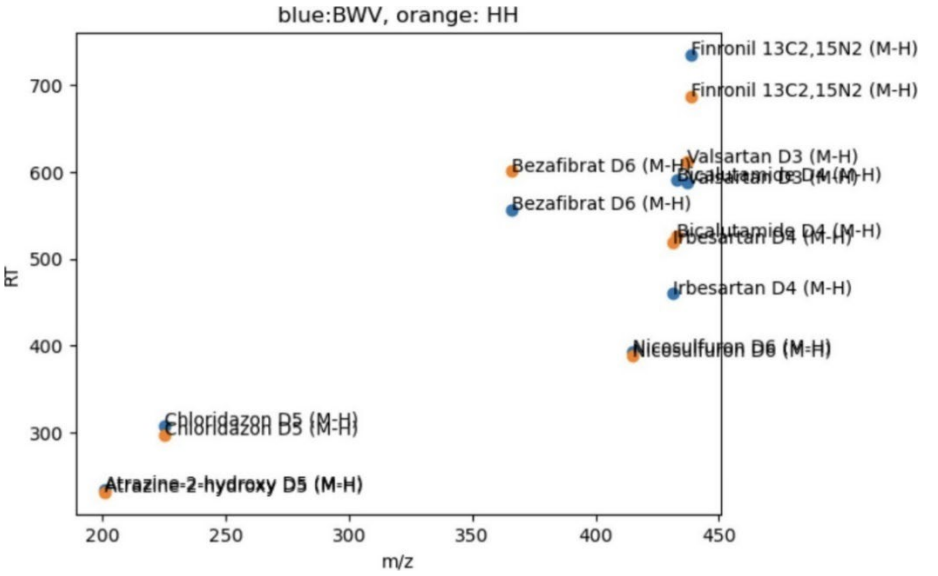


Fig. 4: Internal standards alignment with a simple fit showing a tendency for larger gaps for higher mass to charge ratios for this particular lab.

3.2 Triplicates

Multiple processings (replicas) of the same sample (usually three, called triplicates) can be used as to compare different possible representations of the same chemical properties, distorted or altered through the addition of noise in the different processing steps (comp. Fig 5). The different triplicates can be used to spot fragmented profiles and stray peaks. If an RT-m/z combination appears only in one or two of the three triplicates, but a close value (tolerance of 10 ppm in m/z and approx. 1-10 seconds in RT were used in our tests) with a similar intensity is present, these are combined into one signal. If no such complementary substance profile is found, the value is discarded as noise.

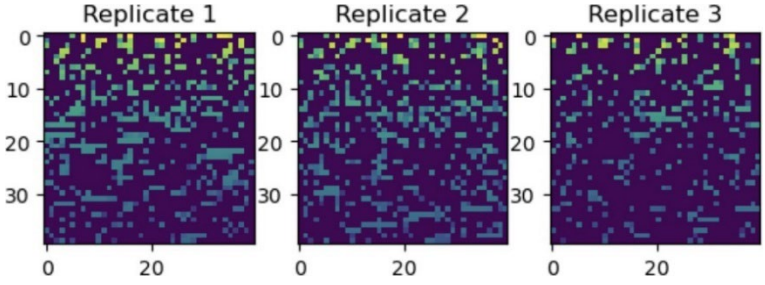


Fig. 5: Graphical representation of varying profile intensity in triplicates, showing deviation in intensity (color) and distribution of signal

3.3 Use Cases and Applicable Evaluation Concepts

Once data standardisation among labs and measurement attempts is achieved, the data can be subject to further analysis. In the following, we sketch some pathways we are currently exploring in the K2I project.

Trend analysis

A surprising jump or decrease in intensity of a feature is a noteworthy observation that typically warrants further investigation. Making statistics over longer time frames of the fluctuations of known (or unknown) substances can also lead to valuable insights about contamination patterns. Comparing intensity between labs is only possible in a semi-quantitative way, needing at least 5 fold changes to reliably identify a meaningful difference, as was found within a proof of concept study where different concentrations of known contaminants were analyzed by different labs (T. Bader and T. Waldmann [Wa22]). Within one lab it is possible to detect more fine-grained (approx. two-fold) intensity changes. Juxtaposing the general direction of trends among labs is nevertheless deemed useful. Therefore different views for intra- and inter-lab quantitative comparison appear necessary.

Pattern recognition and Anomaly Detection

A main goal of the on-going work is to apply analytical methods to the pre-processed data to detect changes that hint at the presence of new substances or that can help to identify significant changes in the mixture of chemicals in the aquatic environment. After successfully recognizing a change it can then be tried to attribute it to particular environmental or anthropogenic influences. To be able to identify such changes, an assessment of known and expected patterns is necessary. Because of the high number of features and fluctuations and the large number of samples necessary to arrive at meaningful conclusions, it is difficult to evaluate this manually. Statistical methods and time series analysis are a step to structure the large influx of data points. In addition we aim to evaluate different approaches to apply machine learning to use the growing database to detect common signatures of the water chemical mixtures that are indicative of, for example, certain seasons, weather conditions, industrial and agricultural activity, traffic and consumer habits. Patterns that deviate from or contrast these known blends of signals can then be assessed as an anomaly. A change in the spectral patterns compared to the already known states measured through different metrics. One approach is using a neural network [Li20] such as an Autoencoder [Kr91], an Encoder-Decoder pair that learns to condense patterns into a reduced representation space (latent space). A convolutional layer can combine close profiles to even out variance in retention time and, so some degree, mass. Such a reduction in effective dimensionality can make the evaluation more manageable but poses the danger of blurring out peaks making it harder to discern between separate signals. Strong intensity variations or accumulation of uncommon features will likely still be noticeable this way.

Up-/Downstream comparison

When an intensity peak or anomaly is detected, comparison with up or downstream measurements can give an indication towards the validity of the detection, and help to pin down the possible sources. To this end clear definitions for anomaly types that translate between different labs is necessary.

As masses are measured with a high degree of accuracy, a previously undetected m/z value that lies outside the tolerance for m/z variation is a contender for an anomaly, as is a commonly seen m/z value, that appears at a retention time outside the determined tolerance.

As was discussed above, quantitative comparison of intensity among different labs is typically not feasible, apart from changes that span orders of magnitude.

An overall pattern of profiles that is very untypical (as determined for example by a neural network as formulated above) would also be considered noteworthy and the difference between several sampling locations on the same body of water can lead to interesting insights.

4. Conclusion and Outlook

Collaboration between different water providers in sharing and comparing their NTS data can be a step towards mor unified workflows and a transregional approach to detect sources of pollution that may have gone unnoticed otherwise. In the further course of the project it is planned to to use our IT and ETL-pipeline infrastructure set up on Kubernetes to incorporate new measurements into the database on a daily or nearly daily basis. This will include automatic re-processing after upload and performing instant anomaly detection and trend analysis. By making the results accessible to laboratory staff, we offer a tool to help researchers search for noteworthy changes faster. The newly generated samples can then at the same time be used to validate out-of-sample performance of the AI (by manually confirming the validity of proposed anomalies) and running a re-training of the AI models. These can then be developed further towards greater detection precision and information content, benefiting from continuous feedback and input from the users and developers. The current project is running up to April 2023. Given the encouraging results, possible extensions will certainly be discussed.

UN sustainable development goals (https://sdgs.un.org/goals) that are connected to the project are 6 (Clean Water), 11 (Sustainable Cities and Communities), 12 (Responsible Consumption) and 17 (Partnership).

Acknowledgement

The authors wish to thank the Federal Ministry of Education and Research (BMBF) for funding the K2I-project under contract number 02WDG1593A-D

5. List of Terms and Abbreviations

Abbreviation/Term	Explanation
NTS	Non Target Screening
LC-HRMS	Liquid Chromatography - High Mass Resolution Spectroscopy
RT	Retention Time (measurement in liquid chromatography)
ETL	Extract-Transform-Load (type of processing pipeline)
ISTD	Internal Standard (deliberately added chemical to calibrate mea- surements)
Triplicate/Replicate	Additional Processing of the same water sample (usually three), to filter out noise
Project (EnviMass)	A complete set of samples, blanks, metadata, target information and parameter sets that can be processed together by enviMass and will generate output files which are the input for further analysis.
Profiles (EnviMass)	Joined peaks corresponding to a certain mass and retention-time range across the samples that are deemed likely to belong to the same substance (through alignment and peak shape evaluation) by the processing algorithm of enviMass.

Bibliography

[Ba17] Bader, T.; Schulz; Kümmerer, W.; Winzenbacher, K.: LC-HRMS Data Processing Strategy for Reliable Sample Comparison Exemplified by the Assessment of Water Treatment Processes. Analytical chemistry 89/24, 2017.

[Ba18] Bader, T.: Mining of LC-HRMS data for the assessment of water treatment processes. Leuphana University, Lueneburg, 2018.

[Lo19] Loos, M., 2019, url: https://doi.org/10.5281/zenodo.1213098, visited on: 05/21/2022.

[Fu17] Fu, H.-Y.; Hu, O.; Zhang, Y.-M.; Zhang, L.; Song, J.-J.; Lu, P.; Zheng, Q.-X.; Liu, P.-P.; Chen, Q.-S.; Wang, B., et al.: Mass-spectra-based peak alignment for automatic nontargeted metabolic profiling analysis for biomarker screen- ing in plant samples. Journal of Chromatography A 1513/, pp. 201–209, 2017.

[Ho17] Hollender, J.; Schymanski, E.; Singer, H.; Ferguson, L.: Nontarget Screening with High Resolution Mass Spectrometry in the Environment: Ready to Go? Environ. Sci. Technol 51/, p. 243, 2017.

[Ko10] Koh, Y.; Pasikanti, K. K.; Yap, C. W.; Chan, E. C. Y.: Comparative evaluation of software for retention time alignment of gas chromatography/time-of- flight mass spectrometry-based metabonomic data. Journal of chromatogra- phy A 1217/52, pp. 8308–8316, 2010.

[Kr91] Kramer, M.: Nonlinear principal component analysis using autoassociative neural networks. AIChE Journal 37/, pp. 233–24, 1991.

[Li20] Liebal, U.; Phan, A.; Sudhakar, M.; Raman, K.; Blank, L.: Nonlinear princi- pal component analysis using autoassociative neural networks. Metabolites 10/, p. 243, 2020.

[Ot18] Ottensmann, M.; Stoffel, M. A.; Nichols, H. J.; Hoffman, J. I.: GCalignR: An R package for aligning gas-chromatography data for ecological and evolutionary studies. PloS one 13/6, e0198311, 2018.

[Sy21] Syafrudin, M.; Kristanti, R. A.; Yuniarto, A.; Hadibarata, T.; Rhee, J. A.-o.; Algarni, W. A.; T. S.; Almarri, A. H.; Al-Mohaimeed, A. M.: Pesticides in Drinking Water-A Review. International Journal of Environmental Research and Public Health. International Journal of Environmental Research and Public Health 18/2, 2021.

[VL13] Vu, T. N.; Laukens, K.: Getting your peaks in line: a review of alignment methods for NMR spectral data. Metabolites 3/2, pp. 259–276, 2013.

[Wa22] Waldmann, T.: Inter Laboratory Data Processing and Data Mining Using LC-HRMS Data. Hochschule Aalen, Aalen, 2022.

[Ya18] Yang, T.-B.; Yan, P.; He, M.; Hong, L.; Pei, R.; Zhang, Z.-M.; Yi, L.-Z.; Yuan, X.-Y.: Application of Subwindow Factor Analysis and Mass Spectral information for accurate alignment of non-targeted metabolic profiling. Journal of Chromatography A 1563/, pp. 162–170, 2018.

Improving Data Quality of Programme of Measures for the Water Framework Directive in Saxony

Friedhelm Hosenfeld [1], Roland Dimmer[2] and Christoph Mattes[3]

Abstract: A web application is presented to support the responsible authorities in the management of measures for the WFD (Water Framework Directive) in Saxony. The web application enables the maintenance of WFD measures data by different authorities in a common database. The central data management supports the tasks of implementing the WFD of the LfULG for the fulfilment of the EU reporting obligations. A key requirement deals with the improvement of data quality implemented by comprehensive consistency and completeness checks, input rules and support functions for geometry creation. The spatial data are verified for consistency with the attribute data during data acquisition.

Keywords: WFD (EC Water Framework Directive), WFD measure management, Disy Cadenza, data quality, quality assurance, web application, improving spatial data

Addresses Sustainable Development Goal 6: Clean water and sanitation

1. Introduction and overview

1.1 Overview

The Water Framework Directive (WFD) of the European Community [Di00], which came into force in 2000, pursues an integrated water protection policy in Europe, which also brings about coordinated management of the waters within the river basins across state and national borders. In order to achieve the central objective, a good status of as many water bodies as possible, plans and programmes of measures were created and will be updated in further management periods.

This paper describes a web application operated by the State Agency for the Environment, Agriculture and Geology (LfULG) for the management of measures planned and implemented under the WFD in the Free State of Saxony and to support reporting to the

[1] DigSyLand - Institute for Digital System Analysis & Landscape Diagnosis, Zum Dorfteich 6, D-24975 Husby, hosenfeld@digsyland.de, https://orcid.org/0000-0002-0538-7684
[2] Sächsisches Landesamt für Umwelt, Landwirtschaft und Geologie, Referat 44, Oberflächenwasser, Wasserrahmenrichtlinie, Zur Wetterwarte 11, D-01109 Dresden-Klotzsche, Roland.Dimmer@smekul.sachsen.de
[3] Disy Informationssysteme GmbH, Ludwig-Erhard-Allee 6, D-76131 Karlsruhe, Deutschland, christoph.mattes@disy.net

EU. One main goal focusses on facilitating central data management improving data quality and quality assurance methods. The application was developed by the companies Disy and DigSyLand and integrated into the Saxon data portal iDA, which provides interdisciplinary access to environmental data and maps on the web.

In the following sections, the tasks of measure management in Saxony, the initial situation and the requirements for the application are discussed. Section 2 deals with the design of the application and explains the components of the software architecture. Section 3 is dedicated to the implementation of data quality improvement methods in the web application, where the special challenges due to the technical requirements are mentioned. The article concludes with a summary and an outlook in section 4.

1.2 Initial situation and general conditions

The LfULG maintains a state-wide database on the implementation of the WFD and already makes much of these data available in a variety of ways (like map services, fact sheets and downloads), because the LfULG is the central authority in Saxony for informing the authorities involved and the public about the state-wide data and activities for the implementation of the Water Framework Directive.

Of particular importance are the measures to improve the status of water bodies, for which the data are maintained locally by other water authorities and about 18 additional institutions according to their responsibilities. In particular, information on the preparation and updating of management plans and on programmes of measures is essential for reporting to the EU, as well as for interim reports on the progress of measure implementation.

The data on WFD measures maintained by other water authorities and further stakeholders were kept by them in different standardized Microsoft Excel tables, which were merged manually at certain dates. The LfULG manually compiled the data in a database for reporting purposes at state level. Working with the Excel tables proved to be error-prone, inflexible and time-consuming.

Especially the non-optimal data storage and management with all consequences resulted in the need to develop a web application that both reduces the workload and supports the optimisation of data quality.

1.3 Requirements for a data management web application

The aim of the data management application for WFD measures was to create a possibility for web-based creation and editing of WFD measure data by the users in the Saxon authorities, whereby all relevant data should be available in a central database.

The LfULG developed a technical concept that formed the basis for the conception and

development of the web application, which mentioned the following requirements, among others:

- Web-based processing of the WFD measure data, so that in particular the creation, editing, deletion, display, research and export of these data are possible.
- In order to continue to support also local data management, import of measure records into the central database should be offered in addition to interactive editing.
- With a suitable user management, the technical and spatial responsibilities should be appropriately mapped within the application.
- The ability to integrate the solutions into the existing IT infrastructure had to be guaranteed. This included in particular the integration into the iDA data portal, which is based on the Disy Cadenza[4] evaluation and GIS platform.
- The technical requirements included extensive consistency conditions, which are to be ensured by the data management, as well as conditions for the analysis and optimisation of WFD measure geometries in relation to the water bodies important for the WFD and their catchment areas.

2. Design of the application

2.1 Software architecture

The software architecture of the web application complies to the general conditions of the LfULG and uses the components that are already in use in the data and analysis platform iDA (see section 2.2):

- Data storage in the central Oracle database (section 2.3)
- Analysis, export and GIS functionality in Cadenza Web (section 2.4)
- PHP for implementation of data processing and management functions (section 2.5).

2.2 Saxony's data portal iDA - interdisciplinary data and analysis

Saxony's web portal iDA (interdisciplinary data and analysis) offers a comprehensive access to environmental data and maps, which originate from measuring and research programmes of the LfULG and further information systems of the Free State of Saxony[5].

[4] Disy Cadenza: Software for data analysis and business intelligence (https://www.disy.net/en/products/disy-cadenza/overview/, accessed: 10/07/2022)
[5] iDA: https://www.umwelt.sachsen.de/umwelt/infosysteme/ida/, accessed: 10/07/2022

The uniform portal platform is implemented with Cadenza Web and combines the structured access to data of the different departments of the state. In addition to public access, further access options are offered for registered users, who thus can also access non-public thematic data according to their area of responsibility.

2.3 Database design and data management

The thematic data available in iDA are managed in the Oracle RDBMS, supplemented by spatial data that are integrated via services such as ArcGIS Server REST services.

Concerning the WFD measures management, data modelling in the Oracle database included the representation of geometries. This means that the geometries can be used directly by common GIS software with Oracle support without additional interfaces, especially by Cadenza but also by PHP applications.

The data model requires that the WFD measures are edited exclusively using the data management application in order to guarantee consistency and access rights. Due to complexity and specific requirements, as well as better configurability, most consistency checks have to be handled by the data management application, except for referential integrity constraints on the database level (see section 3.2). By default, each WFD measure data set has exactly one geometry (polygon, line or point), except for conceptual measures which have no concrete geometry.

Special challenges for the data modelling included the mapping of multiple assignments of properties to WFD measures, which cause both content-related and technical consequences. For example, several water bodies and also several entries of the nationally defined LAWA[6] catalogue of measures [LA20] can be assigned to a WFD measure. This catalogue controls the validity of geometry types, but also of water body categories. In addition, it must be possible for analyses to unambiguously evaluate these multiple links, so that, for example, WFD measures are not considered multiple times in calculations and map representations.

Other important technical and content-related aspects that had to be appropriately considered in the data modelling include:

- Ownership (right to change) of a WFD measure, which can also be passed on to other institutions.
- Flags for technical correctness and completeness (factual data and geometries)
- Different statuses with different consistency requirements, also ensuring that legacy data are also available for evaluation in the same system.

[6] LAWA: Bund/Länder-Arbeitsgemeinschaft Wasser: German Working Group on water issues of the Federal States and the Federal Government represented by the Federal Environment Ministry (https://www.lawa.de/English-About-LAWA.html, accessed: 21/05/2022)

2.4 Cadenza Web as Evaluation and GIS Component

The evaluation and GIS platform Disy Cadenza forms the main component of the iDA web portal, which provides access to environmental data and map inventories in Saxony for the public and other user groups.

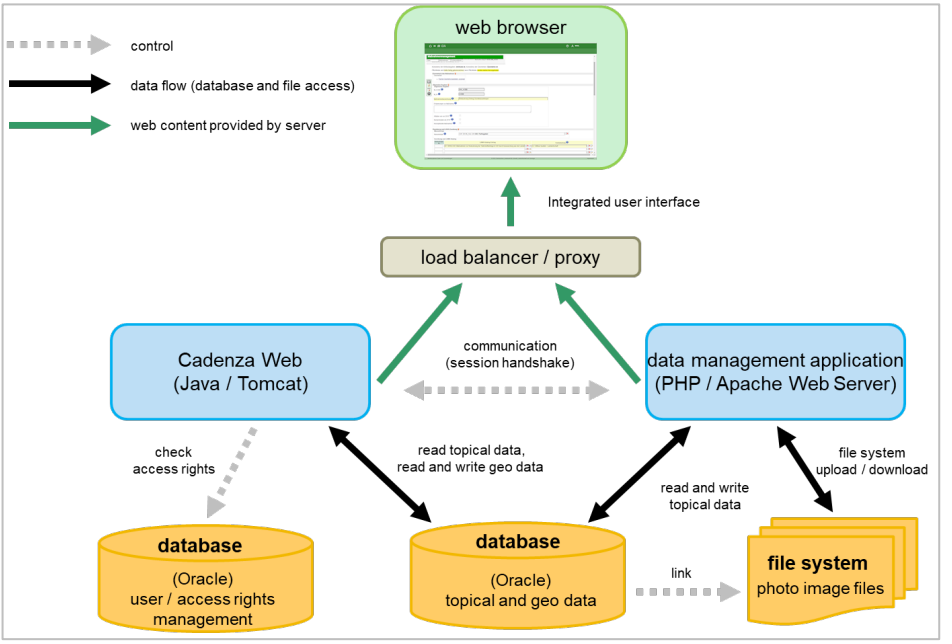


Fig. 1: Overview of the interaction of data sources, Cadenza Web and PHP data management application

The Cadenza Web variant includes numerous evaluation functionalities of topical and spatial data as well as the acquisition and maintenance of geometries.

Cadenza is continuously developed by the company Disy in close cooperation with several federal and state authorities, especially from the environmental sector, and is therefore tailored to the current requirements of these cooperation partners (for example applications, also for WFD applications, see e.g. [HTT16, Ho12b, Ho21a, Ho21b]).The so-called Cadenza repository forms an intermediate layer for the individual application areas as an integrating view of the underlying data sources, which can be database-based as well as services and geodata files (for details see e.g. [Ho21a] p16f.).

With filter forms Cadenza offers the combined query of topical data and spatial data according to all required criteria with presentation of the results in exportable tables, but also as reports and interactive maps.

While the focus of the Cadenza component is on data analysis, reporting and GIS, for the realisation of thematic data management systems in which data are edited and processed, it offers the integration of data management applications via a programming application interface with which, for example, data acquisition functions can be added [Ho12a], which can be tailored to the respective application purpose. In this way, the generic, configurable standard functionalities can be supplemented with customised acquisition and data management functionalities.

The companies Disy and DigSyLand have not only developed data management information systems following this architecture in Saxony on behalf of the LfULG, but also for the environmental state authorities in Schleswig-Holstein [Ho12b, Ho21a, Ho21b]. While earlier developments of WFD supporting web applications in Schleswig-Holstein were successfully realized only based on PHP [Ho08, HLT11], the additional standard functions offered by Cadenza, especially concerning spatial data turned out to be essential.

Therefore, it was obvious to use this combination of components of Cadenza Web and PHP also for the realisation of the WFD measures management.

2.5 PHP management application for data editing and maintenance

For the management of the WFD measures in Saxony a data management application was designed to fulfil the requirements of the technical concept regarding the acquisition and maintenance of the measures.

This component was realised using the scripting language PHP[7] and seamlessly integrated into the Cadenza Web environment via its application programming interface (see Fig. 1).

Both Cadenza Web and the PHP application access the Oracle database directly, while Cadenza Web only performs read accesses except for editing geometries of the WFD measures, while the management application is responsible for editing the topical data (see Fig. 1).

3. Implementation of data quality improvement methods in the web application

3.1 PHP-Framework as the basis for the data management application

The data management application is based on a PHP-framework developed by DigSyLand, which was optimised for the integration into the Cadenza Web application programming framework, so that a mapping of the application-related access functions to

[7] PHP stands for “PHP: Hypertext Preprocessor” (https://www.php.net/)

the user groups of the Cadenza user administration is possible (see also [Ho12b, Ho21a, Ho21b]).

3.2 Geometry adjustment and consistency checks

While the WFD measure input form could be designed largely by configuration on the basis of the PHP framework, some special functions, which particularly concern the complex consistency checks and also the checking, adjustment and intersection of the WFD measure geometries, had to be programmed separately.

The important technical requirement that linear measures should be adapted to the actual course of the waterbody route was implemented in PHP based on Oracle geometry processing functions: After a line is drawn on the river water using the Cadenza geometry acquisition function and then transferred to the data management application, the latter forms a new line geometry from the line segments of the watercourse from the starting point to the drawn end point of the WFD measure, which lies exactly on the watercourse line.

For all measure geometries (points, lines, or polygons) it is checked in which water bodies or catchment areas and in which municipalities they are located. These assignments are automatically added as attributes in the attribute data for evaluations and consistency checks. In addition, it is checked whether the measure lies within the user's area of responsibility. Thus, the lower water authorities can only create geometries in their district area, while other authorities can edit WFD measures Saxony-wide.

Invalid measure geometries are marked accordingly, but are initially kept in the system to facilitate subsequent corrections without having to completely create the geometry again.

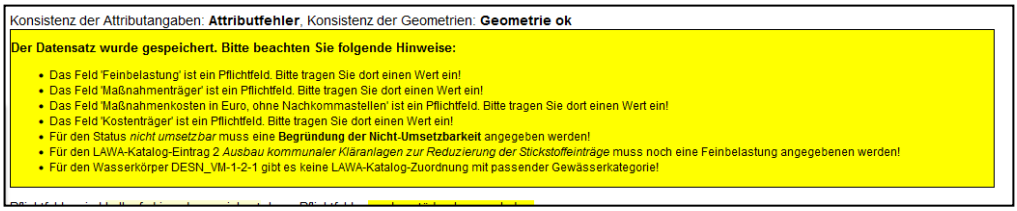


Fig. 2: Results of comprehensive consistency checks are displayed after saving a record

This also applies to WFD measures whose attribute data were entered inconsistently or incompletely: These are marked as incorrect with regard to the attributes, so that a subsequent completion and correction can be made (see Fig. 2).

In addition to the obligatory filling of mandatory fields, the consistency checks include the checks of content-related correlations, such as the LAWA catalogue entry [LA20] matching the water body category and the geometry type with the specification of the associated pressures and other specifications (see Fig.2).

As far as possible, the data are already checked during input or controlled by limiting selection options. Due to the dependencies of different attributes, the data can only be analysed completely after saving, so corresponding hints based on the consistency checks are displayed (see Fig. 2).

For measures marked as legacy data, which were initially transferred to the Oracle database, slightly relaxed consistency conditions apply to enable inclusion in the application.

3.3 Import interface

The institutions responsible for WFD measure management are subject to different local general conditions. While the central data management in the newly developed WFD measure management application is favoured by the LfULG, local measure management solutions are also still in use. In order to also be able to transfer the measures maintained locally into the central database, a standardized import format based on Excel files was specified, which is based on the original Excel exchange format.

The data management application offers an import interface for the transfer of externally maintained data. Both new measures can be introduced and existing measures can be updated.

During the import, the same consistency conditions are checked as during the interactive input. If an automatic correction is possible, it is carried out while otherwise incorrect data records are rejected during the import. After the import, a detailed log is generated with notes on errors and corrections. Point and line geometries are defined by specifying the coordinates of the (end) points. Recently an additional import option was implemented for polygon geometries based on shape files.

3.4 Evaluations and map representations

Using the Cadenza standard functions the search and evaluation of the managed measures were developed and integrated into a new Cadenza repository, so that all attributes of WFD measures are available as criteria for the definition of the search. The results can be displayed and exported as a list, but can also be displayed on the map. Several map display options of the measures were implemented, symbolizing the measures according to different properties like status, responsibilities, waterbody category and more.

4. Summary and outlook

With the web application the central data management of the WFD measures was implemented which offers access for all stakeholders for interactive processing and evaluation of the common data pool. Additionally external third-party data management solutions were connected via an import interface. Using the data management application all organisations responsible for WFD measures in Saxony have a common up-to-date view on all relevant data. Especially in areas where different organisations are planning WFD measures a better exchange is facilitated.

While formerly the spatial representations of the measures in the manually compiled central data pool were very limited, the new web application offers suitable map representations of the WFD measures according to different criteria which can be combined with several related map layers which were centrally managed by the LfULG.

The realisation of comprehensive consistency checks and completeness rules help to ensure the data quality, thus achieving a key requirement which is especially important for EU reporting. The improved data quality also enables better evaluations on the state level and reduces the amount of effort to work efficiently with the data.

In autumn 2021 a pilot phase started when interested participants of different organisations tested the application and exchanged their experience with the LfULG.

As a result of the pilot phase, it was found that the application improves the management of measure data in the desired way, but that there was still a need for supplementing certain functionalities.

While these optimizations are currently being added which include among others the extension and performance improvements of the import interface, archiving of WFD measure records, more evaluation options, additional data fields and improvements of usability, users can test the application on a test system and deliver their feedback.

In autumn and winter 2022 the final data migration into the new system is planned so that productive operation of the WFD measure management web application can start at the end of 2022. Support for filling EU reporting templates is planned for future development phases.

Bibliography

[Di00] Directive 2000/60/EC of the European parliament and of the council of 23 October 2000 establishing a framework for Community action in the field of water policy (EU Water Framework Directive). In: Official Journal of the European Communities, 22/12/2000 pp. 0001-0073, 2000, https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32000L0060&qid=1653152811078&from=EN, accessed: 21/05/2022.

[Ho08] Hosenfeld, F. et al.: Database of measures supporting creation of management plan and programme of measures for implementing Water Framework Directive. In: Möller, A.; Page, B.; Schreiber, M. (eds): Environmental Informatics and Industrial Ecology. 22th International Conference on Informatics for Environmental Protection, pp. 541 – 548, 2008.

[Ho12a] Hofmann, C. et al: Cadenza - Raumbezogene Berichtssysteme und Fachanwendungen auf der Basis von Cadenza. In: Weissenbach, K.; Ebel, R.; Weidemann, R. (eds.): UIS BW Umweltinformationssystem Baden-Württemberg, F+E-Vorhaben MAF-UIS, Phase I 2011/2012, Report-Nr.. KIT-SR 76616, Karlsruhe, pp. 81-92, https://pudi.lubw.de/detailseite/-/publication/38498-Cadenza_-_Raumbezogene_Berichtssysteme_und_Fachanwendungen_auf_der_Basis_von_Cadenza.pdf, accessed: 21/05/2022.

[Ho12b] Hosenfeld, F. et al.: Development of a Database Supporting the Management of Nature Conservation Measures in Schleswig-Holstein. In: Arndt, H.-K.; Knetsch, G.; Pillmann, W. (eds.), Light up the Ideas of Environmental Informatics, Dessau, Shaker Verlag Aachen, pp. 623–629, 2021, http://enviroinfo.eu/sites/default/files/pdfs/vol7793/0623.pdf, accessed: 21/05/2022.

[Ho21a] Hosenfeld, F. et al.: Neuentwicklung einer Intranet-Web-Anwendung für das Geotopkataster Schleswig-Holstein. In: Freitag, F. et al. (eds.): Umweltinformationssysteme - Wie verändert die Digitalisierung unsere Gesellschaft? Tagungsband des 27. Workshops „Umweltinformationssysteme (UIS2020)“ des Arbeitskreises „Umweltinformationssysteme“ der Fachgruppe „Informatik im Umweltschutz‘‘ der Gesellschaft für Informatik (GI). Springer Vieweg, Wiesbaden, pp. 13-27, 2021, https://doi.org/10.1007/978-3-658-30889-6_2, accessed: 21/05/2022.

[Ho21b] Hosenfeld, F. et al.: Governmental Information Systems Supporting Nature Conservation in Schleswig-Holstein. In: Wohlgemuth, V. et al. (eds.): Environmental Informatics. A bogeyman or saviour to achieve the UN Sustainable Development Goals? Adjunct Proceedings of the 35th edition of the EnviroInfo – the long standing and established international and interdisciplinary conference series on leading environmental information and communication technologies Berlin, Germany, September 27-29, pp. 133-140, 2021.

[HLT11] Hosenfeld, F.; Lietz, J.; Trepel, M.: Assessment Database Supporting Implementation of Water Framework Directive in Schleswig-Holstein. In: Pillmann, W.; Schade, S.; Smits, P. (eds.): Innovations in Sharing - Environmental Observations and Information, Ispra, pp. 460 – 466, 2011.

[HTT16] Hosenfeld, F.; Tiffert, J.; Trepel, M.: Waterbody and Nutrient Information System Schleswig-Holstein. In: Wohlgemuth, V.; Fuchs-Kittowski, F.; Wittmann J. (eds.): Environmental Informatics – Stability, Continuity, Innovation. Adjunct Proceedings of the 30th edition of the Enviroinfo. Berlin, pp. 407-412, 2016.

[LA20] LAWA (Bund/Länder-Arbeitsgemeinschaft Wasser): LAWA-BLANO Maßnahmenkatalog (WRRL, HWRMRL, MSRL), 60pp., 2020, https://www.lawa.de/documents/lawa-blano-massnahmenkatalog_1594133389.pdf, accessed: 21/05/2022.

Towards extended reality soundwalks as community noise communication tool

Iwer Petersen [1]

Abstract: Noise is getting increasing attention as environmental factor. Communication of potential impact to affected citizens is part of a participatory approach to many projects like e.g., in wind-energy projects. Sound, however, is hard to grasp by non-specialists looking at commonly used metrics. Noise maps from sound simulations are often used to communicate expected noise levels from certain sources but only provide abstract insights that still can be hard to transfer for non-experts. Extended reality (XR) in combination with real-time spatial audio allows to explore the impact of future projects in the built environment as a direct experience. While the generation of a visual environment poses no greater challenge by using geospatial data, realistic modelling of a spatial soundscape from multiple point-sound-sources is not trivial. This paper reflects on the challenges to present an audio-visual windfarm in XR, outlines established techniques that have proven useful as well as challenges and opportunities in creating, delivering, and evaluating virtual soundscapes.

Keywords: spatial audio, immersive audio, soundscape

Addresses Sustainable Development Goal 7: Affordable and clean energy

1. Motivation

Audio in immersive environments has received increasing attention over the past years as sound, synchronized with the visual environment, increases the feeling of "being there"[Hr19]. With the emergence of spatial, geometry-aware audio engines (e.g. [Me13], [RGR20]) the possibilities for immersive audio greatly increased. Instead of only controlling volume and panning of individual sounds to create the ear-signal of the listener, individual sounds are placed in a three-dimensional scene and the ear-signal is generated at run-time based on the position of the listener and other run-time parameters. This sound propagation model enables more physically correct sound simulation and seems like a perfect addition to immersive virtual-reality applications. While the audio engines are already in use in computer gaming, they are mainly used in a creative manner. Given the physical motivated nature of the audio engines however, virtual replications of realistic acoustic scenarios seem feasible. Given the complex, often frequency dependent nature of sound propagation, the definition of a realistic acoustic environment with its

[1] Hamburg University of Applied Sciences, Department of Computer Science, Berliner Tor 7, Hamburg, 20099, iwer.petersen@haw-hamburg.de, https://orcid.org/0000-0003-3803-062X

numerous variables for each individual sound source is no trivial task. At the same time there are only subjective evaluation methods available to evaluate the quality of a virtual soundscape that are rather time intensive.

The following sections will outline the process to create a realistic environment in a reasonable amount of time. After a brief look into the creation of the visual environment, the challenges in creating the acoustic environment are explained in more detail. The problems to evaluate the acoustic environment objectively are explained before the paper closes with an outlook into future work on the subject.

2. Visual Environment

The visual environment has different challenges depending on the use case for virtual reality (VR) or for augmented reality (AR). In AR most of the visual environment is the real world, seen either through a head-mounted AR-Display or through a camera image rendered on a screen. Tracking accuracy ensures that virtual content is fixed within the reference frame. However, with objects in increasing distances the tracking accuracy is reducing logarithmically, making large scale AR experiences require manual calibration of the location, orientation, and elevation of the user for accurate representation (e.g., [KHL19], [BF22]).

In VR the everything that can be seen is part of the visual environment and must be created. Using techniques and data from the context of Geo-information systems (GIS) modelling of a realistic virtual environment for immersive applications is very approachable (e.g. [Hr19], [Ma14]). Geo-data standardization enables the development of ready-to-use pipelines to generate 3D content for the landscape, water areas, roads, and buildings. This requires an established spatial reference between Geo-coordinates and a Cartesian coordinate system. Geo-tiff elevation maps for example carry all the information to create the shape of a landscape in a realistic scale. A raster resolution of 1 square meter per pixel with at least 16 bit gray-scale depth proved suitable for virtual-reality (VR) landscape generation. Land-cover information can be used to determine the texture of the landscape and vegetation types. ESRI Shapefiles or Openstreetmaps data can be used to generate polygonal geometry for water surfaces or simple building models extruded from their footprint. Road networks from Openstreetmaps can also be used to animate vehicles along roads.

While there is still no ideal pipeline for the direct use of arbitrary Geo-data in immersive virtual environments, the approach is clear and extensible. Several commercial services are already providing e.g., complete cities as 3D content for immersive environments.

2.1 Acoustic Simulation

The production of a realistic virtual soundscape, however, is a much more complicated

task. After a sound is produced by a source, it is influenced in various ways while it travels to listener. All sounds arriving at the listener with a relevant level compose a three-dimensional sound-field that preserves the direction of the individual sounds. The two-channelled ear-signal for the listener then must be generated by convoluting the sound-field using a head-related transfer function (HRTF).

Spatial audio technologies already provide a framework to simulate a realistic soundscape, but the number of variables involved make it tedious to find the correct settings. Many acoustic effects can be explained and modelled using physical principals, but some propagation effects depend heavily on the sound source. The complex, often frequency-dependent behaviour of sound waves is sometimes difficult to analyse and reproduce.

Moreover, is it hard to objectively evaluate a 3D acoustic simulation, because generally this is done by subjective listening experiments. It is impossible for researchers to assess, whether a negative evaluation of a participant results from an incorrect simulation or e.g., from the participants experiences in the past.

2.2 The spatial audio pipeline

In the spatial audio pipeline a sound-field around the listener is generated at runtime from individual point sound sources at specific locations in 3D space. Each point sound source is subject to different acoustical phenomena such as directional emission characteristics, distance-based attenuation, or interaction with geometry like reflection and diffraction. Depending on the path a sound travels, these effects can occur independently and simultaneously. By utilizing the geometry, that is used to render the visual content in VR, sound paths for each sound source can be determined and the acoustic effects can be applied and mixed at run-time.

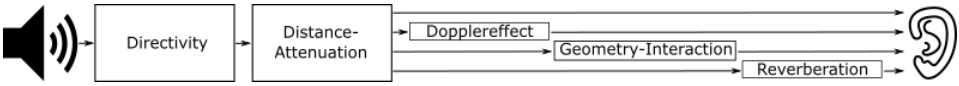


Fig. 1: An example of a spatial audio pipeline for a single sound source.

Fig. 1 shows an example of such an effect pipeline for a single sound source. A source sound is successively processed by several effects, that modify the sound according to parameters related to source-directivity, distance, and movement speed. Objects in the path between source and listener can segment the sound path and apply reflection or diffraction effects. In VR this is possible, because the objects that the sound should interact with, usually are present for visual purposes. Spatial audio pipelines leverage visual geometry in different ways to calculate acoustic effects at run-time. In Augmented-Reality (AR) the virtual scene usually does not have any knowledge about the real world, and in consequence can only consider virtual objects for acoustic interaction. In part these problems are common to the visual realm e.g., for occluding virtual objects behind far away real objects and approaches using vision-based information retrieval may assist

spatial sound rendering in AR in the future [Ko21].

On the concrete example of a windfarm a wide variety of distances, sounds, sound levels and sound source movement speed can occur. To provide a rich and natural acoustic background, sound sources from wind turbines, natural sources, distant streets or urban areas have to be considered. This involves rather quite sources like wind noises in vegetation or animals as well as louder, farther reaching sounds like the aerodynamic noise of the wind turbine blades or distant cars or e.g., church bells. Through the large diameter of the wind turbine rotors, sound sources can reach significant speeds, making it important to accurately model the doppler effect, that is responsible for the frequency modulation of the turbine blade noise.

While the inter-operation of visual and acoustical content is beneficial sometimes, it is notable that the rate of the visual content generation of ~30 - 120 frames per second is not sufficient for audio modulation. The Doppler-effect for example is easily modeled with a distance-based delay effect, that considers the speed of sound and the distance between source and receiver. If the fast-moving source object would update the distance only each visual frame, the modulation steps would be clearly audible. For a smooth transition of a fast-passing car for example, the distance between source and listener must be taken at audio framerate.

Some of the effects, like the Doppler-effect, are easily generalizable with knowledge of physical sound propagation. Sound dissipation and air absorption can be modeled using physical parameters, adapting e.g., the speed of sound to the atmospheric temperature, humidity, and pressure. The general rules for sound reflection and diffraction can be augmented using acoustic materials, that attenuate specific parts of the acoustic spectrum when a sound hits an object. Different 3D objects then can be tagged with different acoustic materials for different reflection and diffraction properties.

Other properties however are completely subject of the specific sound that should be reproduced. On the one hand there is the obvious audio signal that will be emitted, on the other hand there are source specific directivity properties, that influence different attenuation stages.

To represent the source signal, anechoic samples would be the ideal source material for spatial audio engines, because they contain ideally no noise besides the intended signal. A good alternative seems to be near-field recordings with a high signal to noise ratio regarding the intended signal. In case of fast-moving sound sources like rotor blades of a wind-energy turbine however, near-field recordings seem impossible. The alternative is to approximate the source sound with a synthesized noise signal, that can be subsequently refined to get better results.

The directivity of a sound source defines not only how much of the signal is emitted in which direction but also how the sound dissipates with increasing distance. If the energy of the audio signal is radiated with a spherical sound source, it dissipates with increasing distance according to the inverse square law. With a hypothetical ideal directional source,

it would dissipate linearly. The directivity of a sound source therefore has direct impact on the perceived loudness at a specific spatial relation between source and receiver. Consequently, it is not trivial to "reverse-engineer" a source signal by analyzing recordings taken from a distance. Specialized acoustic cameras (see e.g., [Me21]) can help locate the sound sources and even give some clues regarding directivity, but nevertheless significant uncertainties regarding the sound source properties can remain.

From this point it would be beneficial to aim towards a perceptual similarity of the virtual soundscape with the real counterpart. Given the usual soundwalk approach however seems an impractical evaluation tool for iterative refinement as will be discussed in section 2.4.

2.3 Last mile of acoustic AR

While the delivery of a virtual soundscape to a listener via headphones in VR only must deal with minor challenges, which mostly revolve around personalized HRTFs, acoustic AR is presented with the much greater challenge of achieving acoustic transparency for headphones. Acoustic transparency means, that listeners can hear virtual sounds without limiting their ability to hear and locate real sound sources. There are two distinguishable approaches to acoustic transparent headphone designs. Closed-ear approaches record the real-world using microphones and mix the signal with the virtual soundscape while open-ear approaches deliver the virtual soundscape in the near-field of the ear, where it mixes physically with the surrounding real-world sounds. While open-ear approaches to acoustic transparency often must deal with distortion caused by the loudspeaker assembly in the vicinity of the ear [SNM21], the challenge in closed-ear approaches is often recording and reproducing the real, outside soundscape in a way that preserves e.g., directional cues. Several approaches try to use microphone arrays to preserve directional cues while recording the real sound-field surrounding the headphone user ([En21], [PCF19], [Be21], [Br21]). Practical acoustic AR overall seems to be a few steps away from what is already possible in VR.

2.4 Evaluation of Acoustic Virtual Environments

When acoustic XR is directed towards public participation processes in terms of noise, ecological validity of the virtual soundscape is of great concern. Ecological validity is used to quantify how well a laboratory experiment applies to the real world. For acoustic immersive environments a high ecological validity would mean, that participants of the virtual environment would perceive the virtual soundscape likely as realistic. The methodology to investigate ecological validity in virtual acoustic environments is partly shared by the ISO-12913 ([DIN14], [DIN20], [DIN21]) series about soundscape measurement and evaluation. An evaluation method often found in soundscape research is the soundwalk: An in-situ acoustic assessment of several pre-selected sites, each accompanied by a structured questionnaire about sounds at that location. The questionnaire captures the participants most prominently perceived sound sources and a

qualitative assessment of the surrounding acoustic environment. This perception-based assessment method allows the quantification of subjective, perceived soundscape quality. While the approach generally also works in virtual environments, it is not practical to aid development and repeated tuning of the virtual soundscape to account for the uncertainties mentioned at the end of section 2.2.

In recent years the idea of objective soundscape metrics to evaluate soundscape quality became more popular. The idea is to take psycho-acoustic and contextual factors into account to enhance or replace decibel-based sound metrics [Ka19]. Different soundscapes would become comparable by their sets of metrics (e.g. [Wi21.1], [Wi21.2]). A recent study has shown that answers to soundwalk questionnaires can be correlated to a set of soundscape metrics [BC21]. With these approaches it seems feasible to 1) compare virtual soundscapes to recordings of real soundscapes objectively, and 2) predict the soundwalk assessment automatically and aid the development and tuning process of a simulated, virtual acoustic environment.

3. Outlook

The approach to practical, realistic soundscapes in VR or AR applications consists of three separate challenges. The first challenge is the actual source signal of all the sounds that are represented. While in many cases anechoic or nearfield recordings are possible, some sound sources only occur under circumstances that do not allow direct recording of the signal. Synthesis of the desired signal can be arbitrarily complex and introduces uncertainties. Secondly a physical motivated sound propagation model must be designed around the used spatial audio engine that is used. Spatial sound engines often use simplifications that distort physical sound propagation models for the sake of performance or simplicity. This also introduces a significant uncertainty in the ecological validity of the acoustic simulation. Thirdly the synthesized, virtual sound field must be delivered to a human's ear in a way, that preserves locatability of individual sound sources and, in case of AR, mixes seamlessly with naturally occurring sound of the real world. With all the uncertainties introduced, the validation of the sound scape could easily become very tedious. A developer tool, that can predict soundwalk questionnaire results and therefore calculate an approximation of a human's reaction to the soundscape would reduce developer bias and introduce objective quality metrics that could help approximating the virtual soundscape to its real counterpart.

While a virtual counterpart to an existing windfarm was developed, in-situ soundwalks were performed at the location of the real windfarm, collecting psycho-acoustic evaluations together with audio recordings for later evaluation. In the future these soundwalks will be replicated in the VR version of the windfarm to compare human reactions to the real and the virtual windfarm. Due to the described uncertainties, it can be expected that the evaluation of real and virtual soundscapes would probably differ notably. In the next step psycho-acoustic metrics will be calculated from the real soundwalk

recordings and in the simulation at runtime to verify the predictability of soundwalk questionnaires. Given a sufficient predictability, the next step would be to refine the acoustic simulation to minimize the divergence in psycho-acoustic metrics, before conducting a final soundwalk evaluation of the virtual soundscape. When the method proves viable, it could lay the foundation for reinforcement learning based algorithmic refinement of virtual soundscapes on the basis of psycho-acoustic metrics.

Bibliography

[Be21] Ben-Hur, Z., Alon, D. L., Mehra, R., & Rafaely, B. (2021). Binaural Reproduction Based on Bilateral Ambisonics and Ear-Aligned HRTFs. *IEEE/ACM Transactions on Audio, Speech, and Language Processing, 29*, 901–913. doi:10.1109/TASLP.2021.3055038

[Br21] Berebi, O., Rafaely, B., Ben-Hur, Z., & Alon, D. L. (September 2021). Enabling Head-Tracking for Binaural Sound Reproduction Based on Bilateral Ambisonics. *2021 Immersive and 3D Audio: from Architecture to Automotive (I3DA)*, (S. 1–7). doi:10.1109/I3DA48870.2021.9610936

[BC21] Bilen, A. O., & Can, Z. Y. (2021). An applied soundscape approach for acoustic evaluation – compatibility with ISO 12913. *Applied Acoustics, 180*, 108–112. doi:https://doi.org/10.1016/j.apacoust.2021.108112

[BF22] Burkard, S. & Fuchs-Kittowski, F. (2022). Mobile Outdoor AR Application for Precise Visualization of Wind Turbines using Digital Surface Models. *GISTAM Geographical Information Systems Theory, Applications and Management 8.* 15–24.

[Co21] Colombo, M. (October 2021). Vision-based Acoustic Information Retrieval for Interactive Sound Rendering. *2021 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, (S. 487–490). doi:10.1109/ISMAR-Adjunct54149.2021.00115

[En21] Engel, I. (May 2021). *Improving Binaural Audio Techniques for Augmented Reality.* Ph.D. dissertation. doi:10.25560/92700

[DIN14] D. I. für Normung e.V. (Hrsg.)(February 2018). Acoustics - Soundscape - Part 1: Definition and conceptual framework (ISO 12913-1:2014). *Acoustics - Soundscape - Part 1: Definition and conceptual framework (ISO 12913-1:2014)*. doi:10.31030/2784629

[DIN20] D. I. für Normung e.V. (Hrsg.)(November 2020). Acoustics - Soundscape - Part 2: Data collection and reporting requirements (ISO/TS 12913-2:2018). *Acoustics - Soundscape - Part 2: Data collection and reporting requirements (ISO/TS 12913-2:2018)*. doi:10.31030/3162349

[DIN21] D. I. für Normung e.V. (Hrsg.)(June 2021). Acoustics - Soundscape - Part 3: Data analysis (ISO/TS 12913-3:2021-06). *Acoustics - Soundscape - Part 3: Data analysis (ISO/TS 12913-3:2021-06)*. doi:10.31030/3247926

[Hr19] Hruby, F. (January 2019). The Sound of Being There: Audiovisual Cartography with Immersive Virtual Environments. *KN - Journal of Cartography and Geographic Information : Kartographische Nachrichten, 69*, 19. doi:10.1007/s42489-019-00003-5

[Ka19] Kang, J., Aletta, F., Oberman, T., Erfanian, M., Kachlicka, M., Lionello, M., & Mitchell,

A. (2019). Towards soundscape indices. *Proceedings of the International Congress on Acoustics*, *2019-September*, pp. 2488-2495. doi:10.18154/RWTH-CONV-239249

[KHL19] Kilimann, J.-E., Heitkamp, D., Lensing, P. (2019). An Augmented Reality Application for Mobile Visualization of GIS-Referenced Landscape Planning Projects. *VRCAI Virtual-Reality Continuum and its Applications in Industry 17*. 1–5.

[Me21] Meller, S., Petersen, I., Rokita, D., Ueberle, F., Wendholt, B (2021). Charakterisierung von gemessenen Geräuschen einer Windenergieanlage für die Auralisation in einem virtuellen Windpark. *DAGA Fortschritte der Akustik Vol. 47.* 1253–1256.

[Ma14] 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*, 29–48.

[Me13] Mehra, R., Raghuvanshi, N., Antani, L., Chandak, A., Curtis, S., & Manocha, D. (April 2013). Wave-Based Sound Propagation in Large Open Scenes Using an Equivalent Source Formulation. *ACM Trans. Graph., 32*. doi:10.1145/2451236.2451245

[PCF19] Patel, V., Cheer, J., & Fontana, S. (February 2020). Design and Implementation of an Active Noise Control Headphone With Directional Hear-Through Capability. *IEEE Transactions on Consumer Electronics, 66*, 32–40. doi:10.1109/TCE.2019.2956634

[RGR20] Rosen, M., Godin, K., & Raghuvanshi, N. (2020, August). Interactive sound propagation for dynamic scenes using 2D wave simulation. *Computer Graphics Forum (Symposium on Computer Animation), 39*. Retrieved October 23, 2020, from https://www.microsoft.com/en-us/research/publication/interactive-sound-propagation-for-dynamic-scenes-using-2d-wave-simulation/

[SNM21] Schneiderwind, C., Neidhardt, A., & Meyer, D. (2021, May). Comparing the effect of different open headphone models on the perception of a real sound source. *150.* Audio Engineering Society. Retrieved February 28, 2022, from https://www.aes.org/e-lib/browse.cfm?elib=21082

[Wi21.1] Wilford, D. C., Miksis-Olds, J. L., Martin, S. B., Howard, D. R., Lowell, K., Lyons, A. P., & Smith, M. J. (2021). Quantitative Soundscape Analysis to Understand Multidimensional Features. *Frontiers in Marine Science, 8*. doi:10.3389/fmars.2021.672336

[Wi21.2] Wilford, D., Miksis-Olds, J., Martin, B., & Lowell, K. (2021). Introduction and application of a proposed method for quantitative soundscape analysis: The soundscape code. *The Journal of the Acoustical Society of America, 149*, A72-A72. doi:10.1121/10.0004555

Relation Extraction from Environmental Law Text Using Natural Language Understanding

Heiko Thimm [1] and Phil Schneider[2]

Abstract: In the last decades the highly active area of environmental legislation has produced a vast amount of text documents that contain laws and regulations enacted by various types of rule setters. This large body of legal text documents is still growing with an increasing speed. In order to assure compliance with the regulations, today, corporate specialists spend a lot of time with the reviewing and assessment of these documents. It seems that through the use of text processing assistance tools these important corporate environmental compliance management tasks can be completed in less time. To develop corresponding assistance tools has been the broader goal of this work in which initial text processing experiments with a common Natural Language Understanding pipeline are described. The obtained results confirm that in order to extract meaningful relations from text documents of the environmental legislation area, domain-specific processing techniques that are tailored to the specific language and format of legal text are required.

Keywords: environmental legislation, legal text, Natural Language Processing, Natural Language Understanding, Relation Extraction, text processing pipeline.

Addresses Sustainable Development Goal 9: Industry, innovation and infrastructure

1. Introduction

Due to recent technology advancements in numerous domains, text processing applications that are based on Natural Language Understanding (NLU) techniques are being increasingly used for a variety of tasks [AHN19]. These applications, for example, support human users in administrative tasks, information search and acquisition tasks, judgement tasks, and decision tasks. Obviously, at the forefront of the use of NLU techniques are domains were large amounts of text documents are at the center of the core business processes. This characteristic is particularly true for the legal domain in general. Not only the daily tasks of law firms require to deal with large amounts of legal text documents. Also, corporate business processes require various legal specialist to frequently (often even daily) review large amounts of text documents. This is in particular a duty of environmental compliance management specialists. As part of their common daily duties [Th15] they need to check environmental regulations described in text

[1] Pforzheim University, School of Engineering, Tiefenbronner Str. 65, Pforzheim, 75175, heiko.thimm@hs-pforzheim.de, https://orcid.org/0000-0002-6200-9655
[2] Pforzheim University, School of Engineering, Tiefenbronner Str. 65, Pforzheim, 75175, Phil.Schneider@web.de

documents published by environmental rule setters of various levels (community level, state level, country level, international level, supra-national level). The documents either describe a new environmental regulation (e.g. law, directive, mandatory standard) or a revision of an already existing regulation (or revision). According to a rough estimate of environmental compliance management practitioners, a typical mid-sized globally acting production company with multiple production sites in different world regions, needs to check on a monthly basis several douzens of new English and Non-English environmental regulations (or revisions). Many of them can be filtered out right away because eligibility criteria of the regulation are not fulfilled. For the set of remaining regulations more extensive investigations are necessary possibly requiring group decisions. The investigations typically yield about 2-3 regulations which require compliance enforcement measures such as information measures, training measures, equipment/infrasturcture measures, or product revision or production process revision measures. Note that also any revision of products or/and production processes may require to review text documents of environmental legislation in order to check and assure that the change is in compliance with the relevant environmental legislation [Th15].

Our long term research targets to develop tools that assist environmental compliance specialists in their duty to review and analyse legal text documents for judgement and decision tasks. In particular the tools are intended to enable companies to complete core environmental compliance management tasks in less time such as the relevance assessment task, the measure determination task, and the maintenance of a regulation registry [Th15]. Inspired by recent advancements in the area of LegalTech [DA19], [Ha19] and LegalAI [Zh20], in the initial phase of our research, we explore potential possibilities of NLU approaches. The results will be used to develop and test assistance systems for corporate compliance management tasks in particular the task to review an analyse legal text documents. Based on a review of NLU methods and Natural Language Processing (NLP) methods and techniques, a NLU pipeline was implemented which is able to extract relations from documents. The pipleline was tested with various text documents from the environmental legislation domain. In the further continuation of this ongoing research this initial NLU pipeline will be optimized and also other pipelines will be developed in order to test other NLU approaches including approaches that use machine learning techniques.

The following Section 2 gives a general overview of NLP and NLU and corresponding common main methods. Our initial experiments with a NLU pipeline for relation extraction and an outlook on forthcomming further experiments are described in Section 3. Section 4 contains our conclusions.

2. NLP and NLU – Overview and Methods

Both NLP and NLU focus on making sense of unstructured text data, but there is a difference between the two. NLP is primarily concerned with how computers are programmed to process language and to enable `natural´ communication between

computers and humans. NLP processes are more of a statistical or pattern matching process to derive information from unstructured text data. NLU, on the other hand, prioritizes the ability to understand human language and, thus, refers to how unstructured data is reorganized for machines to be able to `understand´ and analyse it [Ru06]. Initial NLU approaches analyse rules and grammatical characteristics to understand language. More recent approaches make use of Machine Learning techniques [Le22]. NLP and NLU often supply complementary solutions to a problem since they have different theoretical backgrounds, one statistical and one rule-based or Machine Learning-based. But some researchers suggest to view NLU as a subarea of NLP. Fig. 1 adopted from MacCartney's presentation on `Understanding Natural Language Understanding´ at the Inaugural Meeting of the ACM Special Interest Group on AI of the Bay Area Chapter in 2014 contains a corresponding visualization of this view on the two disciplines and identifies for each discipline common problems and application areas [Ma14].

So-called `NLU-pipelines´ refer to a series of steps in which complementary NLP/NLU processing tasks are performed on a given input text or a text corpus in order to provide particular application results. Typical results are a summary, the overall topic, the category, the sentiment of the input text or information being extracted such as named entities and relations among entities.

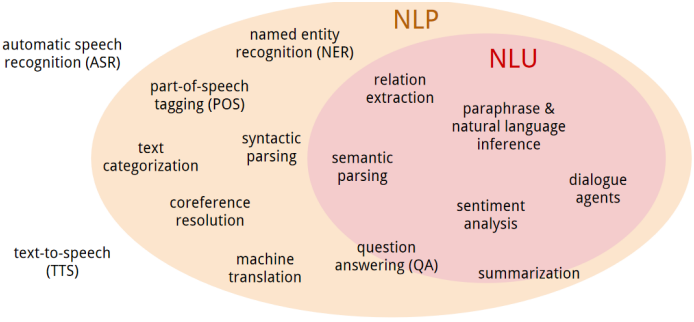


Fig. 1: Terminology NLP vs. NLU including typical applications (copied from [Ma14])

Typically, the first steps of an NLU pipeline perform pre-processing tasks to prepare the input text and, henceforth, enable that the application-specific goal can be achieved by the further processing steps. A common step that follows the text pre-processing is Part-Of-Speech Tagging (POS) which attempts to associate words and symbols in a text with word categories. A brief overview of common pre-processing techniques and POS techniques are contained in the next two paragraphs. This is followed by overviews of embedding-based methods and symbol-based methods. These two types of methods are primarily applied in the legal domain [Zh20]which subsumes the particular application domain of this research.

Common Pre-processing methods. The common pre-processing methods of most NLP tasks are tokenization, stemming, and lemmatization [AHN19], [Ma14]. The purpose of

tokenization is to break chunks of language input into sets of tokens that correspond to paragraphs, sentences, and words. After the tokenization step, the text is converted to lower-case followed by an elimination of numbers, punctuations, and stop words such as `and´, `the´, `a´, `an´ and similar words [Na18]. That is, basically everything which is redundant and does not convey any meaningful insight for the data gets eliminated [AHN19]. Stemming is the process of reducing infected or derived words to their word stem, base or root form. It basically affixes to suffixes and prefixes or to the roots of words known as a lemma. A stemmer removes the endings of many words, e.g., `consolidate´, `consolidated´ and `consolidating´ would be converted to `consolid´ [AHN19]. Lemmatization is the process of reducing inflected forms of a word while still ensuring that the reduced form belongs to the language. This reduced form or root word is called a lemma. For example, `organizes´, `organized´ and `organizing´ are all forms of the lemma `organize´. The inflection of a word allows to express different grammatical categories like tense (`organized´ vs. `organize´), number (`trains´ vs. `train´). Lemmatization is necessary because it helps to reduce the inflected forms of a word and enables to analyse them as a single item. It can also help to normalize the text. As a result, the content becomes more understandable and obtains a clear meaning [AHN19], [Gh20].

Part-Of-Speech Tagging (POS). POS is a common NLU technique which explores the role of a particular word in a sentence. The technique uses eight so-called `parts of speech´: noun, pronoun, adjective, verb, adverb, preposition, conjunction, and interjection. Single words get mark-ups in the text with pre-defined tags such as `N´ for noun, and `V´ for verb. The POS tags, for example, enable keyword extraction based on filtering nouns that typically carry the most significant information [Gh20].

Embedding-based methods aka representation learning. In the legal domain these methods focus on the representation of legal facts and knowledge in an embedding space. In order to solve application-specific tasks they may use deep learning methods. However, these methods require a large amount of prepared training data. Word Embedding methods can fill the gap between text and vectors and convert text into a format that can be processed by AI techniques [Zh20]. Word embeddings are words which are represented by a vector or an array of real numbers. Through the embedding process sentences are transcribed into an array of words and each word of the sentence is transformed into an embedding. Semantic similarity measurement methods can explore the connections between the words and compute similarities like the human memory. Words represented by a vector offer many advantages. They facilitate designing and training deep neural networks since the input consists of machine processible vectors instead of words. Several mathematical techniques for processing numbers are available and can be applied to perform classification, feature extraction, etc.

Symbol-based methods aka structured prediction methods. These methods use knowledge from the respective application domain to solve the NLU task. For example, in the legal domain, especially parts of the symbolic legal knowledge, such as events and relationships, can provide interpretability for lawyers [Zh20]. Two main approaches that use symbol-based methods are information extraction and relation extraction. Information

extraction addresses the general problem of detecting entities referred in texts, the relations between them and the events they participate in. Informally, the goal is to detect elements such as `who´ did `what´ to `whom´, `when´ and `where´. It is the general primary goal to convert a large amount of text into a formal representation of specific fine-grained facts. The structured data obtained from the input text refer to events, entities, facts or relationships between entities presented in the text. This structured information allows computers to perform logic inference or computation on the data, which is challenging if only raw text representation is used [TNS16]. Named Entity Recognition (NER) is a method often used to analyse the text for specific information like names, places, etc. [Gh20]. Relation Extraction focuses on discovering the semantic relations among entities in a text. Various Relation Extraction techniques extract relationship instances that belong to a set of relationship types. These techniques can be grouped into rule-based approaches, supervised approaches, and semi-supervised approaches [Ba16]. Rule-based approaches extract pre-defined relationship types from manually-crafted rules. Supervised approaches use manually labelled documents where the labels describe the type of relationship between each recognized pair of entities. A manually labelled collection of documents is used to train classifiers which, henceforth, are capable to identify all trained relationship types in any dataset. Semi-supervised approaches make use of known relations to recognize new relationships. From the textual contexts of the established relationships, the method derives new approaches and patterns, which in turn are used to derive new relationships.

3. Experiments with an NLU Pipeline for Relation Extraction

The goal of this research is to evaluate the information extraction capabilities of NLU techniques in the domain of corporate environmental compliance management. Through respective experiments we seek to answer two questions: 1. to which extent can useful information be extracted from text documents that are typically reviewed by corporate compliance managers and 2. what specific NLU pipeline is capable to perform this task. Of the planned series of experiments with different NLU pipelines, in a first experiment we used the common NLU pipeline displayed in Fig. 2 which extracts relations from text documents. The raw text of the document is in the initial steps of the pipeline parsed into sentences through a sentence segmenter and further split into words using a tokenizer. Then, each sentence is tagged with POS tags. In the next step, entities are extracted from the text. Finally, rule-based relation extraction is applied to identify relations between different entities in the text and to provide detected relations as tuples [BKL09]. The tuples consist of three elements referred to as triples that are visualized in the form of Knowledge Graphs [Ke22]. Consider for example the triple `(Berlin, capital, Germany)´ that might be extracted from a short text about Germany. Typically, the first element of a triple corresponds to the subject, the second corresponds to the relation, and the third corresponds to the object [Ma14].

The pipeline was implemented based on the programming language Python resulting a

first version program referred in the following as `Compliance Management Information Extractor´ abbreviated CMI Extractor. Various general-purpose NLP/NLU packages and other common Python packages were used for specific steps of the pipeline.

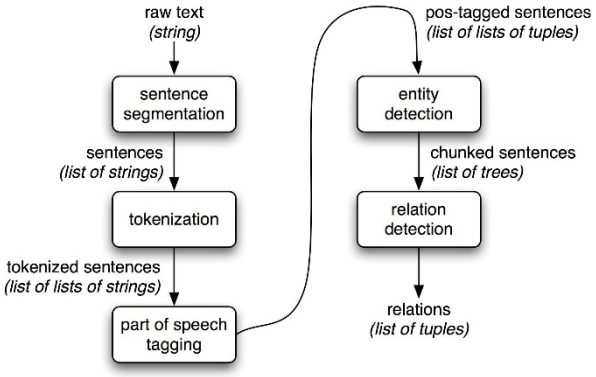


Fig. 2: Relation Extraction pipeline of the CMI Extractor (copied from [Le19])

In order to test and demonstrate the capabilities of the CMI Extractor a test run with the following six relatively short and easy to understand sentences about Elon Musk were performed: *Musk married Riley. Musk is the CEO of Tesla. Musk has a net worth of US$245 billion. Musk is the wealthiest person in the world. Musk founded The Boring Company. Musk graduated in 1995 with a Bachelor of Arts degree in economics.*

The CMI Extractor was able to correctly extract the relations contained in each of the six sentences. The resulting knowledge graph is displayed in Fig. 3. However, it is crucial to bear in mind that the test document consists of simple sentences which represent an ideal input that does not raise complicated processing challenges for the pipeline.

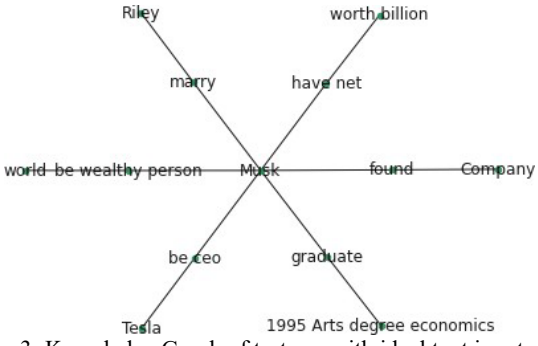


Fig. 3: Knowledge Graph of test run with ideal text input

Obviously, it is much harder for the CMI Extractor to obtain meaningful relations from legal text documents of today's environmental legislation. In order to obtain corresponding insights, the experiments described in Tab. 1 were performed with text

segments of two specific valid directives of the European environmental legislation.

Run	Text	Results / Comments
1	ideal text comprising six simple short declarative sentences with facts about Elon Musk; 42 words	Six meaningful relations extracted; see Fig. 3
2	List item (27) of EU Directive 2020/2184; 264 words; text consists of two complete sentences and a numbered list of 11 list items; list items are comma-separated descriptions containing expert terminology	Four complete relations extracted which require further investigations in order to obtain useful benefits for compliance specialists; see Fig. 4
3	*`Article 2 – Scope´* of EU Directive 2011/65; 261 words; about 80% of the text consists of a numbered list; list items are separated by semicolons; list items consist of several lines of comma separated descriptions containing expert terminology	One complete strange and meaningless relation extracted
4	*`Article 7 - Obligations of manufacturers´* of EU Directive 2011/65; 519 words; same characteristics as described for run 3	Two strange and meaningless relations extracted
5	*`Article 8 - Obligations of authorized representatives´* of EU Directive 2011/65; 162 words; same characteristics as described for run 3	Three strange and meaningless relations extracted
6	*`Article 9 - Obligations of importers´* of EU Directive 2011/65; 432 words; same characteristics as described for run 3	Four strange and meaningless relations extracted
7	*`Article 10 - Obligations of distributors´* of EU Directive 2011/65; 308 words; same characteristics as described for run 3	Five strange and meaningless relations extracted

Tab. 1: Experiments with the CMI Extractor

In the second test run the CMI Extractor was tested with a text fragment of the European Parliament's directive 2020/2184 which targets `the quality of water intended for human consumption´ [Eu20]. Through a random choice from page 8 the list item with number 27 was chosen that consists of 8 sentences and 264 words. In this experiment the CMI Extractor extracted four complete and one incomplete relation (i.e. triples) that are displayed in the knowledge graph of Fig. 4. As opposed to the above experiment the meanings of the extracted relations are less obvious to understand. Using the evaluation framework of a recent Japanese study [TNS16], the extracted relations are to be judged as both *incoherent relation*s and *uninformative relations*. In fact, further investigations are required in order to obtain benefits for compliance specialists from this extraction result.

Fig. 4: Knowledge Graph of test run with a text segment of an EU directive

For the test runs 3 to 7, we used randomly chosen articles of the EU Directive 2011/65 that focusses on `restriction of the use of certain hazardous substances in electrical and electronic equipment´ [Eu11]. In every of these five input texts the CMI Extractor was only able to find some strange and even `more incoherent´ and `more uninformative´ relations as in the second run. It can be assumed that compliance management specialists will not be able to draw any helpful information from these relations. One of the possible reasons for this result is that the input text segments (i.e. articles of the directive) are significantly different from the above ideal text input about Elon Musk and also most common text documents in general. Each of the five test text segments consists to a large degree (~80%) of a numbered list with semicolons separating the different list items. The list items themselves consist of several lines of comma separated enumerations of terms. Many of these terms belong to the expert terminology of the domain of environmental legislation. Even with a revised NLU pipeline that was able to treat semicolons similar to end of sentence points, the extraction performance of the CMI Extractor did not improve.

The above described experimental results confirm the expectation that an NLU pipeline of general-purpose text processing components will only have very limited capabilities to extract useful information from domain-specific text documents. Hence, in the ongoing phase of this research we are exploring approaches to develop a next CMI Extractor that is capable to deal with both 1. the specific language style of environmental legislation documents and 2. the specific terminology of environmental legislation. Consequently, a systematic study of the different types of environmental legislation documents is being prepared and existing dictionaries for the work of corporate environmental compliance management (e.g., [DHS19]) are being evaluated and possibly extended.

One of the promising options to achieve the targeted improvement of the CMI Extractor is to make use of the open source Python package LexNLP which is focused on natural language processing and machine learning for legal and regulatory text [BKD21]. LexNLP supports, among others, the building of unsupervised and supervised models such as word embedding models and tagging models. The package also [BKD21] ´[…] includes pre-trained models based on thousands of unit tests drawn from real documents available from the SEC EDGAR database as well as various judicial and regulatory proceedings´. We also intend to experiment with the popular language processing tool GTP-3 [ZL21] from

OpenAI that is considered to be a foundation model. Foundation models can even be trained on multiple forms of data at the same time. Experiments of our future research will also include to train GTP-3 by use of hand crafted/curated domain knowledge together with industry partners in order to develop a CMI Extractor version with advanced information extraction capabilities. The future refinement of the CMI Extractor will also address the capability to deal with multiple languages.

4. Conclusions

In recent years the business world has been paying increasing attention to the new possibilities that the latest advancements of text processing technologies provide for the digitalization of corporate processes. However, relatively little work of researchers and AI-based software start-ups is devoted to the new possibilities that these advancements offer to the work field of corporate environmental compliance management. Our research attempts to fill this gap by developing and evaluating NLU/NLP-based assistance tools which ultimately extract important information from environmental legislation documents. The extracted information, for example, can be displayed in Knowledge Graphs, thus enabling compliance managers to make fast assessments about the relevance of the document. Furthermore, the extracted information can also be used to generate recommendations for the users based on AI techniques such as machine learning methods. For example, when a legislation document is of relevance for a company and compliance measures are required then recommendations for these measures can be generated from the extracted information and further domain specific knowledge. With respect to this goal, our research is still in an infancy state and will therefore in the future address further experiments with different NLU pipelines. It is expected that from the corresponding experimental results useful recommendations can be obtained for law making regarding meta data and syntactic rules for the legal documents to enable proper NLU support.

Bibliography

[AHN19] Anandarajan, M.; Hill, C.; Nolan, T.: Practical Text Analytics. Maximizing the value of text data. Springer-Verlag, Cham, 2019.

[Ba16] Batista, D. S.: Large-scale semantic relationship extraction for information discovery. Doctoral Dissertation, Liboa, 2016.

[BKD21] Bommarito II, M. J.; Katz, D. M.; Detterman, E. M.: LexNLP: Natural language processing and information extraction for legal and regulatory texts. In (Vogl, R. Ed.): Research handbook on big data law. Edward Elgar Publishing, Northampton, Massachusetts, USA, pp. 216–227, 2021.

[BKL09] Bird, S.; Klein, E.; Loper, E.: Natural language processing with Python. O'Reilly, Sebastopol, California, 2009.

[DA19] DALE, R.: Law and Word Order: NLP in Legal Tech. Natural Language Engineering 1/25, pp. 211–217, 2019.

[DHS19] Deng, Q.; Hine, M., Shaobo, J.; Sujit, S.: Inside the Black Box of Dictionary Building for Text Analytics: A Design Science Approach. J. of Int. Technology and Information Management 3/27, pp. 119–159, 2019.

[Eu11] European Parliament: DIRECTIVE 2011/65/EU on the restriction of the use of certain hazardous substances in electrical and electronic equipment, 2011.

[Eu20] European Parliament: DIRECTIVE 2020/2184 on the quality of water intended for human consumption, 2020.

[Gh20] Ghavami, P.: Big data analytics methods. Analytics techniques in data mining, deep learning and natural language processing. De Gruyter, Boston, Berlin, 2020.

[Ha19] Haney, B. S.: Applied Natural Language Processing for Law Practice. SSRN Electronic Journal, pp. 1–44, 2019.

[Ke22] Kejriwal, M.: Knowledge Graphs: A Practical Review of the Research Landscape. Information 4/13, p. 161, 2022.

[Le19] Learntek: Named Entity Recognition with NLTK. https://www.learntek.org/blog/named-entity-recognition-with-nltk/, accessed April 14th, 2022.

[Le22] Lexalytics: Machine Learning for Natural Language Processing and Text Analytics. https://www.lexalytics.com/resources/wp-content/uploads/sites/3/2019/02/Lexalytics_Machine_Learning_Natural_Language_Processing_Whitepaper.pdf, accessed 22 Apr 2022.

[Ma14] MacCartney: Understanding Natural Language Understanding. ACM SIGAI Bay Area Chapter Inaugural Meeting, 2014.

[Na18] Nay, J.: Natural Language Processing and Machine Learning for Law and Policy Texts. SSRN Electronic Journal, pp. 1–35, 2018.

[Ru06] Rune Sætre: GeneTUC: Natural Language Understanding in Medical Text. Doctoral thesis for the degree of doktor ingeniør, Trondheim, 2006.

[Th15] Thimm, H.: IT-Supported Assurance of Environmental Law Compliance in Small and Medium Sized Enterprises. Int. Journal of Computer and Information Technology 2/4, 297-305, 2015.

[TNS16] Tran, C.-X.; Nguyen, M.-L.; Satoh, K.: A study of open information extraction from legal texts: Proc. First Int. Workshop on Scientific Document Analysis (SCIDOCA), p. 7, 2016.

[Zh20] Zhong, H. et al.: How Does NLP Benefit Legal System: A Summary of Legal Artificial Intelligence. In (Jurafsky, D. et al. Eds.): Proc. 58th Meeting of the Assoc. f. Comp. Linguistics. Assoc. f. Comp. Linguistics, Stroudsburg, PA, pp. 5218–5230, 2020.

[ZL21] Zhang, M.; Li, J.: A commentary of GPT-3 in MIT Technology Review 2021.

Volker Wohlgemuth et al. (Hrsg.): EnviroInfo 2022,
Lecture Notes in Informatics (LNI), Gesellschaft für Informatik, Bonn 2022 53

OData - Usage of a REST Based API Standard in Web based Environmental Information Systems

Désirée Hilbring[1], Kevin Becker[2], Markus Eisenla[3], Katharina Emde[4], Hylke van der Schaaf[5], Horst Spandl[6] and Martina Tauber[7]

Abstract: Environmental information in existing custom-build environmental information systems is manifold. Sharing this information via Web APIs for diverse client types to fulfil the needs of ongoing digitisation efforts is still a challenge. This paper analyses the open standard OData (Open Data Protocol) as a possible communication protocol between independent servers and clients. Of interest is also the question, if the protocol is not only capable of sharing environmental data between independent systems but also if the information provided via OData is sufficient for directly creating a web-based end user client. The developed prototypical implementation is tested in two environmental applications from different domains: a small data overview for decision makers and the integration of information in an environmental platform.

Keywords: OData, Environmental Information Systems, REST, WaterFrame®

Addresses Sustainable Development Goal 9: Industry, innovation and infrastructure

1. Introduction

In recent years many environmental information systems (EIS) have evolved for different environmental domains supporting specified purposes. Taylor analysed a typical environmental agency's products and services and found 63 distinct internet offerings [P16]. They range from providing complex environmental information applications for domain experts through web services providing geoinformation to simple web sites sharing environmental information to the general public. However, achieving interoperability is still a struggle for many projects that try to fulfill the need for digitisation and environmental information sharing.

[1] Fraunhofer IOSB, ILT, Fraunhoferstr. 1, 76131 Karlsruhe, Deutschland, desiree.hilbring@iosb.fraunhofer.de
[2] Fraunhofer IOSB, ILT, Fraunhoferstr. 1, 76131 Karlsruhe, Deutschland, kevin.becker@iosb.fraunhofer.de
[3] Fraunhofer IOSB, ILT, Fraunhoferstr. 1, 76131 Karlsruhe, Deutschland, markus.eisenla@iosb.fraunhofer.de
[4] Fraunhofer IOSB, ILT, Fraunhoferstr. 1, 76131 Karlsruhe, Deutschland, katharina.emde@iosb.fraunhofer.de
[5] Fraunhofer IOSB, ILT, Fraunhoferstr. 1, 76131 Karlsruhe, Deutschland, hylke.vanderschaaf@iosb.fraunhofer.de
[6] LUBW Landesanstalt für Umwelt Baden-Württemberg, Postfach 100163, 76231 Karlsruhe, Deutschland horst.spandl@lubw.bwl.de
[7] LUBW Landesanstalt für Umwelt Baden-Württemberg, Postfach 100163, 76231 Karlsruhe, Deutschland martina.tauber@lubw.bwl.de

This paper analyses the use of OData (Open Data Protocol, an open protocol standard for realizing RESTful APIs) for sharing environmental information. It addresses the question how environmental data can be shared and accessed easily in a generic way and which main characteristics of the data can be shared via the protocol. Furthermore it discusses first ideas how this information can be used by web clients.

A commonly used IT architecture for web applications is a client server architecture with communication via REST based interfaces. OData is a standard collocating a set of best practices for such RESTful APIs.[8] It is an OASIS standard and ISO/IES approved. It iuses the HTTP Protocol to provide CRUD (create, read, upload, delete) operations and supports the JSON Format. It provides a common semantic for data and metadata exchange. The standard has been originally launched by Microsoft and is also used by IBM and SAP[9].

The goal for this paper is to analyse the usage of this standard in the scope of sharing environmental information and data. The following analysis questions have been asked:

- What generic requirements for sharing environmental information exist and can OData fulfil them?
- Is the standard capable of sharing information between independent environmental information systems?
- Is the standard capable for sharing information ready for clients creating an environmental web-based application?

2. Related Work

An overview about Environmental information systems gives Fischer-Stabel in "Umweltdaten und Umweltinformationssysteme" [F21]. They cover a wide field of applications, can be categorized into different system categories (national/international EIS, statewide EIS or regional EIS and operational EIS. They all manage multimodal data about water, air, soil and organisms and try to monitor and control environmental impact.

The Open Geospatial Consortium (OGC)[10] defines many widely accepted standards for modelling and sharing geospatial information. They are designed for many specific use cases, like sharing maps via Web Map Service (WMS) [B06], sharing geographic information via Web Feature Service (WFS) [P09] or sharing IoT sensor data, basically observation data via the Sensor Things API (STA) [LHK15]. Other well-known OGC standards exist. Additionally, many of these standards are used as INSPIRE network services for spatial data[11]. Environmental Information Systems often deal with geospatial

[8] https://www.odata.org, last access 22 June 2022
[9] Overview of the OData Standard | Nordic APIs |, last access 22 June 2022
[10] The Home of Location Technology Innovation and Collaboration | OGC, last access 22 June 2022
[11] Network Services | INSPIRE (europa.eu), last access 22 June 2022

information and for sharing such information picking the appropriate service and providing the geospatial information is advisable. However, besides the spatial information further information exists, which needs to be shared in environmental information systems. The information is often domain specific and modelled in relational databases. Therefore, for data where the semantic data model is crucial and geospatial information is less important other standards need to be considered.

In this paper we are focussing on the question how data, which has often been collected for years and stored in relational databases of environmental agencies can be provided by Web APIs. Since the OData protocol is easily capable of modelling information from relational databases, we analyse this approach. The article "Web APIs for environmental data – state of the art investigation" gives an overview about existing practises for the publication of environmental data using Web APIs focussing on APIs used by organisations handling data similar to meteorological data [P16], which is a sub-set of environmental data. Taylor identifies the following challenges for environmental data: variety, volume, velocity and veracity. OData is only mentioned in context with the SensorThing API. In "Open Data!" Hübener discusses the possibilities for Open Data starting with data available at the German Environmental Agency (Umweltbundesamt – UBA) and considers OData as part of the proposed IT infrastructure [H12]. Wohlgemuth et al. discuss similar ideas in their paper "Entwicklung eines Open-Source basierten Baukastens zur Unterstützung und Etablierung der Ressourceneffizienz in KMU" focusing on SMEs [VZT14]. The client sever communication on the proposed architecture uses the OData-Protocol while Boß and Wohlgemuth uses OData in an operational EIS for energy management [BW15].

3. Functional Requirements for web based Environmental Information Systems

This section lists functional requirements which - according to our expertise - are relevant for sharing environmental information from existing environmental systems. These requirements are the basis for the analysis of the OData protocol implementation in our examples. Requirement for sharing environmental information are similar and manifold in comparison to sharing any other type of information. Main requirements may differ depending on the client using the interface: is it another system or an end user? Is information only presented or is write functionality needed? We marked the lists of common functional requirements with #keyword. This is used to show which requirements help to answer the main questions of the paper from section Introduction.:

- #environmental: a concept, which is widely used and especially important for environmental application.
- #geospatial: a concept, which includes handling of geospatial information.
- #system: a concept, which is needed to share information with another system.

- #user interface: a concept, needed when information shall be presented to end users.

Requirement 1: Environmental data bases are often home-grown relational databases using domain specific information models using strong specific schemas. Relationally structured information, needs to be shared differently compared to non-relational data stored in documents (#environmental). The content of relational databases can often be visualized using tables. For management and sharing of the strongly structured data a list of sub requirement exists:

1. The usual CRUD (create, read, update and delete) operations are necessary (#system, #user interface).
2. Support for common data types like Integer, String, etc. (#system, #user interface).
3. Since environmental information often contains data collected over decades search functions are essential (#system, #user interface).
4. For sharing environmental information, often commonly used key lists are defined for an easy exchange of specific information: Examples are manifold: a taxonomy with biological water organisms, a list of chemical parameters defining common identifiers like CAS or UBA-Code. (#environmental, #system, #user interface)
5. Handling of mandatory and optional attributes. (#system, #user interface)
6. Defining default values for attributes. (#system, #user interface)
7. Handling of valid value ranges. (#system, #user interface)
8. Defining rules for formatting (#system, #user interface)

Requirement 2: Authentication and authorisation play a key role. Managing complex domain information need role concepts (#environmental).

Requirement 3: Error Handling providing information about system or domain errors is necessary (#system, #user interface).

Requirement 4: Presentation of domain specific standard forms, as defined and widely used by domain experts play a key role in environmental information systems (#environmental, #user interface)

Requirement 5: Handling of geospatial information is also common in environmental information systems. The difficulty is to decide, how important the geospatial information is for the application. Does it only support the main goals of the application, or is it so important that a specific OGC standard is obvious to be used (#geospatial).

4. Implementation

EIS address a wide range of specific domains, realized by manifold systems. To test our OData requirements for sharing environmental data between independent systems we extended an existing platform (WaterFrame®) for developing EIS with OData.

4.1 Environmental Applications with WaterFrame®

The WaterFrame® platform from Fraunhofer IOSB provides a framework with basic functionality to develop easy configurable low-code environmental applications based on relational data models: XCNF (Extensible Database Application Configurator)[12]. The platform is used to develop EIS regarding ground water, surface water and water protection, but also for other domains like soil, wind energy or nature protection.

With XCNF thematic views can be easily con^^^^^^d by developers or XCNF experienced users. These views are essentially SQL queries that gather data from one or more database tables, combined with additional meta-data defining how the data should be visualised and edited. The XCNF-Views provide the CRUD functions for the data of the application. For each attribute of a view functional access rights based on role concepts can be configured. Furthermore, analysis code evaluating the attribute data can be integrated. For each view generic import and export functionality can be provided. A combination of these views builds the environmental application.

4.2 OData Concepts

Information to be shared via the OData protocol needs to be mapped to the *Entity Data Model (EDM)* which is the abstract data model of OData and can be identified and manipulated via URLs with simple HTTP messages. Central concepts of the EDM are entities, relationships, entity sets, actions and functions.

To describe the information for clients the OData protocol provides a metadata document in JSON or XML format describing all available entities. It contains all *entity types* along with their properties with name and type, as well as the structure of the entity types via navigation links. It is possible to enhance the instance elements contained in the meta data document with so called *Annotations*, an interesting extensibility point for OData. With metadata annotations characteristics of properties can be further described.

4.3 Extension of an existing environmental framework with OData

XCNF as part of the WaterFrame® platform provides the possibility to create

[12] WaterFrame® - Gewässerinformationssysteme - Fraunhofer IOSB, last access 22 June 2022

environmental applications domain independent. Therefore we extended XCNF with an OData based REST Server. Fig. 1 shows how the existing architecture of XCNF has been extended to integrate an OData Server. The XCNF-Core capable of handling Oracle or H2 databases has been extended with an OData REST API interface of version 4.01 [HP20a] integrated in a jetty web server.

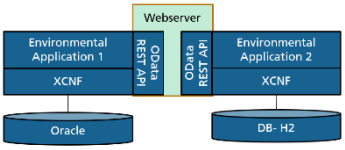


Fig.1: Integration of OData in XNCF architecture

To provide the information from the data model of the environmental application to OData a mapping to the OData EDM needed. In the mapping thematic structured *XCNF views* are automatically mapped to OData *entity types* (see Fig. 2). Thus, for all environmental applications based on XCNF existing thematic views can be adapted or new views can be created from which the content is provided via the OData REST API.

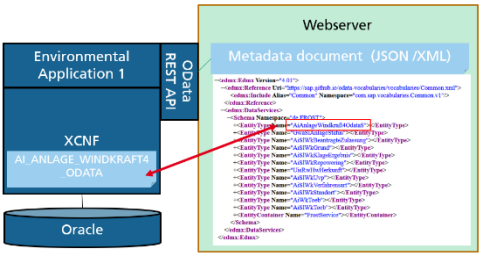


Fig.2: Mapping of XCNF Views to OData Entity Types

With that the common CRUD operations are available to share and manipulate information form the underlying XCNF based application over the internet.

4.4 OData Clients

Information provided by the OData Server can be consumed by different client types.

Generic clients are able to interpret the standardized protocol information in the metadata document and are able to connect directly to different OData Servers:

- One well known example is **Excel**, that can access data from OData 4.0 Servers since Excel version 2019.
- Another generic client has been developed in our project. It is able to interpret the metadata documents of different OData Servers to visualise the information in a

web client and provides the user with CRUD operations.

Specific clients have different sub types. They can directly access information from a specific OData Server for their own tasks (e.g. exchange a specific environmental information piece from and with an external system). Furthermore, specific clients can be designed using the standard OData protocol functions and additionally interpret *Annotations* from a specific OData Server to provide more functionality for their users.

4.5 Addressing the Requirements via OData

In this section we compare the requirements for environmental information applications with the capabilities of the OData Protocol [HP20a].

Requirement 1: Environmental information structured in a relational way can be usually mapped directly to the *Entity Data Model* of OData and therefore the domain information can be provided via the basic functionality of the protocol (#environmental) and fulfil the sub requirements #system and #user interface:

1. CRUD operations are provided by the basic functionality of OData.
2. Common data types are provided by the basic functionality of OData.
3. OData defines a number of query options in section 11.2.1 of the OData protocol as basic functionality. Our implementation supports the $filter functionality and further functions to control the result sets: $count (number of results), $orderBy (result order), $select (limit the result set to specific properties).
4. With the $expand function OData provides a possibility to directly expand navigation links in a query as basic functionality. This function can be used to directly integrate the values of key value pairs of keylists into an OData result set.

Fig.3: The *StaatNr* from Keylist *UISStaatSet* is expanded directly to *Deutschland*.

5. The handling of mandatory and optional attributes can be addressed using additional annotation functionality in OData.
6. Default values can be defined using annotations.
7. The handling of valid value ranges can also be addressed via annotations. The following examples may be used for this: Defining the maximal length of texts

or defining the precision of attributes.

8. More complex formatting rules for attributes can be derived via annotations using regular expressions.

Requirement 2: Authentication can be handled in OData just like in any way a generic RESTful API can be secured. In our test implementation the goal is to reuse the already existing underlying authentication and authorization concept of XCNF (#environmental).

Requirement 3: Error Handling can be realized via using common HTTPs Status Codes[13]. An example maybe using 401: "Unauthorized" for a missing authorization for writing or reading specific attributes (#system, #user interface).

Requirement 4: Another aspect is the question how hints for the layout of forms via the OData protocol can be delivered. Using *Annotations* like *"@inTab("Name of Tab")* grouping of elements belonging together is possible (#environmental, #user interface).

Requirement 5: To decide which standard fits best OData or a geospatial one, one needs to weight the importance of the geospatial information. Consider the following two examples: firstly, showing a nitrate distribution of an area in a map, a typical use case for an OGC WMS. Secondly, information needed for the status overview of complex permission processes of wind generators. That information is not structured in a widely used data model. Only the places of the generators are of interest. The coordinates can be defined using the geometrical types of OData. They are parts of the OData Primitive Types following the OGC Simple Feature Specification [HP20b], which contains widely used types like points or polygons (#geospatial).

Thus, as generic protocol for client server communication of web applications OData provides well known working concepts and functions for exchanging common information types. This applies also for generic concepts with important roles in many environmental applications like a structured data model, a decent role concept or key lists. Using OData gets more specific when trying to provide highly specified forms information with specific layout requirements for end users. The Annotation concept can deliver layout hints for clients. Additionally, providing geospatial information is possible, but depending on the use case other suitable geospatial standards may be preferable.

5. Example Applications

We started to test the prototypical implemented REST-based OData interface for environmental XCNF applications of the WaterFrame® platform in two use cases from different domains, wind energy and ground water. They use different client types: generic and specific.

[13] https://de.wikipedia.org/wiki/HTTP-Statuscode, last access 22 June 2022

5.1 Usage in a Wind Energy Application

The LUBW[14] develops in the domain of trade supervision an environmental application based in XCNF for the data management of the permission process of wind energy generators. The permission process for wind energy generators is complex. Legal impacts in the domains of pollution control, building legislation, nature protection, air transportation, and mobility need to be considered. Therefor the permission process for a wind energy generator might need several years. The goal of the expert users of the environmental application is to manage this permission process, analyse the different states of the processes and identify critical processes. An overview view containing this information uses the basic functions of the OData protocol and can therefore be visualised using generic clients. Fig. 4 shows the data of the corresponding XCNF View provided via the OData Server visualised in the generic client and as result of the OData feed in Excel.

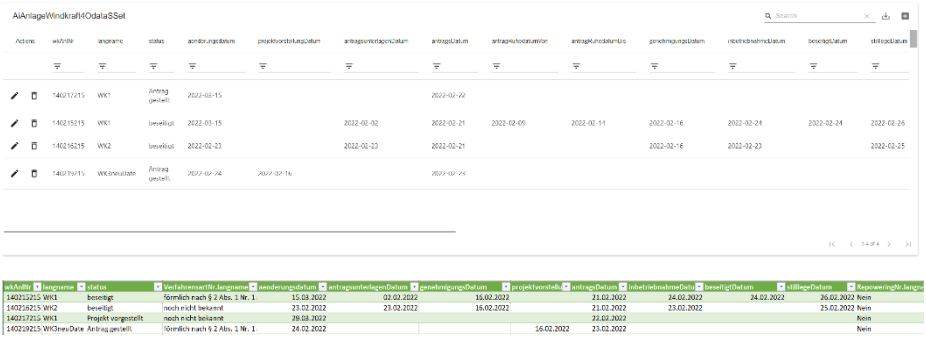


Fig.4: Approval process of wind energy generators available in two generic OData clients.

5.2 Usage in NiMo 4.0

The project NiMo 4.0 is funded by Federal Ministry for the Environment, Nature Conservation and Nuclear Safety (BMU) in the support programme AI Lighthouses for Environment, Climate, Nature and Resources. The main goal of the project is to develop ML algorithms supporting nitrate research in ground water. The results of these algorithms are integrated in the NiMo platform for end users. One example is a time series nitrate forecasting at a point of measurement. In this project on the one side a classic geoinformation standard is used. The OGC SensorThings API standards manages the handling of time series for nitrate and other chemical quality observation parameters [LHK15]. Furthermore, the end users of the NiMo platform are interested in additional hydrogeological metainformation of the survey stations. This information was already available in an existing information system for domain experts: an XCNF based

[14] Landesanstalt für Umwelt Baden-Württemberg: https://www.lubw.baden-wuerttemberg.de/, last access 22 June 2022

WaterFrame® environmental information systems managing ground water measurements and analysis of TLUBN15. The new OData interface of XCNF has now been used to configure an XCNF-View with hydrogeological information and provide it via the REST based OData interface. This information can be now accessed and integrated into the NiMo platform. The NiMo platform on the one side is independent of WaterFrame® and it shows end user information. Thus, the NiMo platform is in context of this paper a specific client. Fig. 6 shows the XCNF-View collecting hydrogeological information and the result of the GET request reading this information:

http://localhost:8680/FROST-Server/ODATA_4.01/GwMstDatenloggerHydroNimoSet?$expand=TeilraumNr .

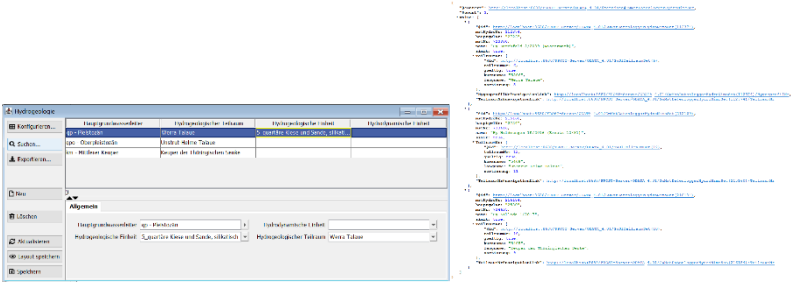


Fig. 5: Hydrogeological Information via OData GET request

6. Conclusion and Outlook

The analyses of the OData prototype showed that providing commonly used environmental data to web applications via the OData protocol is possible. OData is capable of sharing environmental information for different use cases, but depending on the structure of the information to be shared (relationally structured environmental information or geospatial information) other well know open standards should be considered. Since this paper describes work in progress and just presented first results, we are positive about the already reached generic possibilities for sharing environmental data from all WaterFrame® platform applications, not only limited to our example test cases. Adaption of existing or creation of new views sharing data via the OData protocol is configurable and does not need extra implementation effort. In future we plan to investigate and realize the proposed usage of OData Annotations to enhance the information about delivered entity types for our environmental clients.

[15] Thüringer Landesamt für Umwelt, Bergbau und Naturschutz: https://tlubn.thueringen.de/, last access 22 June 2022

Bibliography

[B06] Beaujardiere, J.; OpenGIS Web Map Server Implementation Specification, Version 1.3.0, 06-042, 2006

[BW15] Boß, J., Wohlgemuth, V.:Integration und Weiterentwicklung bestehender Energiemanagement-Applikationen mit dem OpenResKit Framework, INFORMATIK 2015, Bonn 2015

[F21] Fischer-Stabel, P, Umweltdaten und Umweltinformationssysteme (UIS) in Umweltinformationssysteme, ISBN 978-3-87907-700-7

[H12] Hübener, J: Open Data! EnviroInfo, Dessau, Part 2: Open Data and Industrial Ecological Management, Shaker Verlag Aachen, ISBN: 978-3-8440-1248-4, 2012

[HP20a] Handl, R., Pizzo,M (ed): OData Version 4.01. Part 1:Protocol, OASIS Standards, 2020, http://docs.oasis-open.org/odata/odata/v4.01/odata-v4.01-part1-protocol.html, last access 22 June 2022

[HP20b] Handl, R., Pizzo,M (ed): OData Version 4.01. Part 1:Protocol, OASIS Standards, 2020, Section 3.3 Primitive Types, https://docs.oasis-open.org/odata/odata-csdl-json/v4.01/odata-csdl-json-v4.01.html#sec_PrimitiveTypes, last access 22 June 2022

[LHK15] Liang, S, Huang C.-Y., Khalafbeigi, T (ed): OGC SensorThingsAPI Part 1:Sensing Version 1:0, 15-078r6, 2015

[P09] Panagiotis, P: Web Feature Service 2.0 Interface Standard – With Corrigendum, Version 2.0.2, 09-025r2, 2009

[P16] Tayler, P: Web APIs for environmental data – state of the art investigation, Technical Report, 2016, DOI: 10.13140/RG.2.2.17817.83047

[VZT14] Wohlgemuth, V, Ziep, T, Schiemann, L: Entwicklung eines Open-Source basierten Baukastens zur Unterstützung und Etablierung der Ressourceneffizient in produzierenden KMU, 6. BUIS Tage, 16. Tagung der Fachgruppe Betriebliche Umweltinformationssysteme der Gesellschaft für Informatik e.V., Berlin, 2014

Volker Wohlgemuth et al. (Hrsg.): EnviroInfo 2022,
Lecture Notes in Informatics (LNI), Gesellschaft für Informatik, Bonn 2022 65

Smart Citizen Science in pluvial flood disaster risk reduction: Building a mobile application as one tool for drain path identification (Work in progress)

Katharina Haupenthal, Peter Fischer-Stabel[1]

Abstract: Over the last years, flood risk has increased, and the threat of flooding has caused severe damages for economy, society and infrastructure. Hence, the project Urban Flood Resilience- Smart Tools (FloReST) was initiated by six partners from the field of civil engineering, informatics and hydrology to research on tools for high-resolution drain path identification and risk mapping. In this context, a mobile application shall be developed for crowdsourced data collection in civil society. Still being in the early stages of development, a first requirement catalogue for the application is presented and discussed, showing that especially data control is a problematic issue in Citizen Science.

Keywords: Citizen Science, mobile application development, crowdsourced data collection, flood risk management, pluvial flooding, surface runoff, drain path

Addresses Sustainable Development Goal 11: Sustainable cities and communities

1. Introduction

The threat of flooding has increased during the last years and poses not only economic damages but also social and infrastructural ones. Therefore, various projects are being carried out that aim to address the topic of flooding from different perspectives. Whereas in the past, research was mostly focused on *fluvial* flooding, there is still a lack of research carried out on *pluvial* flooding. Fluvial flooding describes floods that result from raising water levels, pluvial flooding is caused by heavy rainfalls. Especially the latter often leads to so-called flash floods, floods that emerge from heavy rainfalls in a short time. In July 2021, large areas of North Rhine Westphalia and Rhineland Palatinate in the western part of Germany have been hit by such flash floods, leading to enormous damages and loss of lives.

Since various research focuses on the observation of water levels, little one is carried out on drain paths. The latter describes water runoffs[2], mostly flash floods, that run naturally and above ground. However, due to increasing urbanization, the natural streams are

[1] University of Applied Science Trier, Umwelt-Campus Birkenfeld, Institute for Software Systems, Campusallee 9925, P.O. Box 1380, D - 55761 Birkenfeld / Germany; k.haupenthal@umwelt-campus.de, p.fischer-stabel@umwelt-campus.de

[2] In this research proposal the terms water runoff and drain path are similarly used.

blocked and thus turn into cellars, houses, etc. which causes severe damages. This problem is tackled by the project *Urban Flood Resilience – Smart Tools (FloReST)*, funded by the German Federal Ministry of Education and Research. A group of six partners from the fields of civil engineering, IT and hydrology is working together to develop tools for high-resolution drain path identification and risk mapping in five different pilot communes in Rhineland Palatinate, Western Germany. One of these tools is planned to be a mobile application that allows citizens of the different pilot communes to actively engage in data collection. Their local expertise and knowledge about past flood events and potential places at risk during floods shall help to develop a high-resolution mapping for local drain paths.

This approach of including citizens into the research process is known as *Citizen Science* and mostly common in environmental monitoring, as bird observation. Wisner, O'keefe and Westgate [WOW1977] already mentioned the early concept of people`s science, in which communities are encouraged to contribute to the observation and documentation of a certain phenomenon of scientific interest. In brief, Citizen Science describes the involvement of non-scientists into a scientific process, in which the level of involvement can vary from data collection to a full involvement into the design process [Se19][Bo09]. In literature, there is also the expression *crowdsourcing*, which is often used in such context, meaning "the outsourcing of tasks to a crowd that would otherwise be too large to accomplish by a single organization" [Se19,2].

The following research statement is developed by the Institute for Software Systems of Umwelt-Campus Birkenfeld, part of University of Applied Science Trier, in the context of the FloReST project, mainly addressing the issue of building a mobile application for the usage of collecting crowdsourced data for drain path mapping. As the project is still in its early stage, the application has not yet been developed. Rather, the team of scientists has started with scanning different mobile applications in the field of early warning and flood detection as well as diving deeper into the topic of Citizen Science and related application development. Based upon the collected information together with survey and interview data from the pilot communes, a requirement analysis for the mobile application shall be carried out in the future process.

As this research proposal is a work-in-progress-statement, the authors aim at providing information about the status quo on the project work as well as receiving qualified input that can be included in the current research. Thus, the research question of this paper is the following: *Which preliminary features should be included in the FloReST Citizen Science mobile application?*

In the following, the case study will be outlined (section2), literature about Citizen Science and surface runoff will be reviewed (section 3), as well as an overview of several pre-selected apps will be presented (section 4.1). Further, application requirements will be analysed, and a first requirement catalogue will be drafted (section 4.2). Finally, limitations and an outlook of will be presented (section 5).

2. Case Study

Five pilot communes in Rhineland Palatinate were selected as target areas based upon historic flood experiences: Mendig, Altenahr, Trier, Linz (Rhine), Herrstein-Rhaunen. Within these areas, the app shall be tested and further developed. The citizens are requested to take part as data-collectors about local drain paths. The underlaying idea is to have a mobile application that allows the citizens to mark spots at risk, as potential drainage blockage in rainy situations, etc. within an interactive map, similarly to a risk detector. Pictures and a categorization of the risk shall be asked in addition. The data will be stored in a data warehouse developed by one of the project partners (Disy).

By sending such information citizens support the data collection of drain path identification within the target areas and add important information to the final risk mapping. An interactive map shall facilitate the positioning of the data collector as well as the handling for potential users. Due to the expected high average age in some of the target areas, the app-usage shall be easy, intuitive and without any greater obstacles. An easy-to-understand layout and contextual description are therefore of great importance.

3. Literature Review

3.1 State of the art: surface runoff

Literature about surface runoff in general is rooted in the field of environmental engineering, as the research of green roof performance in water runoff management [BBJ09][Be10][BNR09][MCR13] or research about harvesting rainwater [St20]. Moreover, research was undertaken in the field of hydrology/ hydromechanics, mostly measuring the quality of water runoff dependent on different surface types [Bu11][CCO21][GDC07][PLC00]. However, neither a human-geographical perspective of water runoff pathways is existing nor participatory approaches for drain path identification.

3.2 (Smart) Citizen Science

In literature, Citizen Science faces great popularity due to its awareness creating approach as well as its impact on the improvement of science-society dialog [Go21][Ke21][Vi21][Wo21]. As Benjamin Franklin stated, "*Tell Me and I Forget; Teach Me and I May Remember; Involve Me and I Learn*" [Fr22], people's involvement in research creates awareness leading to effective public learning and thus ends up in a call for action and resilience building [Be19][We16]. Further, citizens receive the chance to be actively engaged with research, as they might be involved in the design, elaboration and implementation of the latter. In such scenario, citizens act as "collaborators" within

the research process. Haklay [Ha13] distinguishes between “collaborators” and “sensors” (p.116). In contrast to the former, the latter describes a rather passive research behavior of the local community. By using a crowd-sourcing approach, for example, citizens act as “sensors” as they take a contributory role. Citizen Science is a broad terminology; therefore, it is important to define the approach in an early stage related to the research context. In the context of the FloReST project, citizens shall assist in data collection about drain paths in the local communities. Thus, citizens will act as “sensors”.

Especially digitalization has become an engine for Citizen Science as it has improved the process of collecting and accessing citizen-generated data [Ka20][Wo21]. Because of increasing internet access, mobile applications as means for data collection have gained popularity. Examples are Open Street Map[3], ornitho[4] or Scent[5]. Especially in the field of environmental monitoring, Citizen Science is a well-known and popular approach, even before digital transformation and thus mainly experiences a revival as smart Citizen Science [DBZ10].

However, scholars argue that the most critical aspect in Citizen Science is the insurance of data quality [Co08][DBZ10][Di12][Ge21][Ko16]. It is debated in how far a non-scientific data base can be compared to the standards of a scientific one [RP13]. Scholars question if Citizen Science can be considered as a reliable approach to science, arguing that lay scientists are lacking training and knowledge in scientific data management or research methodology and thus might not understand how to properly collect and record data [Co08][DBZ10][Di12]. Crall, Newman and Stohlgren [CNS11] found out, that Citizen Scientists showed a lower ability in correctly identifying invasive plant species compared to professional scientists. Due to an increasing facilitated access to Citizen Science data collection tools, as web and mobile applications, it can be assumed that the risk of unreliable data entries is raising.

Nonetheless, according to Geyer [Ge21], Citizen Science does not aim at competing with science, it rather aims at extending scientific data bases. Citizen Science can be seen as a tool to collect data in research projects, in which a scientific data base might not be representative or hardly to exercise. Kosmala et al. [Ko16] argue that data sets produced by Citizen Scientists show relatively high-quality data compared to scientific research as well as similar biases to professional data collection.

The quality of the data can be validated through different control mechanisms, as community control, through Artificial Intelligence as well as manually. To sum up, although the issue of data validation, Citizen Science faces great interest as it can be considered as a cheap mean to gather on-the-ground-data and to include local knowledge and expertise [Ha15][NLR22].

[3] https://www.openstreetmap.de/
[4] https://www.ornitho.de/
[5] https://scent-project.eu/about-scent

3.3 Smart Citizen Science for urban flood risk management

Citizen Science is considered promising within the field of flood risk assessment [Go21][Ke21][Sc21][Vi21]. Due to the active inclusion of local knowledge and expertise, Citizen Science has shifted the approach in flood risk reduction from a top-down to bottom-up [Pa18]. According to See [Se19], smart Citizen Science has the potential to contribute to the development of early warning systems. In this context, social media plays an increasing role in flood risk assessment. Arthur et al. [Ar18] analyzed Twitter as source for real-time flood mapping. Twitter is perceived by various scholars as instrument to detect, cluster, categorize and map flood events based on the crowd-sourced information [AS17][FS18][Li18][PN16].

Most of the mobile applications found for Citizen Science in flood risk management are such reporting on fluvial flooding. Examples are the application *CrowdWater*[6] by the University of Zurich or a do-it-yourself water level measurement tool developed by the Umwelt-Campus Birkenfeld [Um22]. Citizens are encouraged to involve in water level management by building their own measurement tool using the instructions online.

Although Citizen Science is considered promising within flood risk management it is still limited to simple applications, as the validation of prediction models or emergency assessments [Sc21][Wo21]. Thus, more opportunities for Citizen Science in flood risk management need to be developed. Until today, numerous applications and platforms have been developed for citizens to report in fluvial flood events [Se19]. Research was thus undertaken in the fields of water level measurement ([Fa18][We19]) or water quality monitoring ([Fa17][He21][Le17][Th17]). Considering pluvial flood events, research is mostly limited to Citizen Science in rainfall monitoring ([ACC15][Sh20][We19]). According to See [Se19], little research has been carried out for Citizen Science activities in pluvial flooding. No research was found for Citizen Science in the field of drain path identification.

4. Methodology and results

The following section will provide an overview of the methodological approach used in the current research. In a first step, existing flood warning and Citizen Science mobile applications in flood risk management were identified and analyzed based upon their goals and features [Ap22]. The latter were retrieved from the development intention described in section 2. Drawn from this evaluation, it was assessed in how far the applications fit to the described research aim of mobile application development in the context of the FloReST project. In a second step, Lemmens et al.'s [Le21] criteria for mobile application development in Citizen Science projects were reflected and applied to the current research case.

[6] Described in section 4.1 State of the art: mobile applications

4.1 State of the art: mobile applications

Twelve applications treating flooding or rainfall warning were identified[7] and characterized, from which three were selected for a closer examination, as they appeared the most suitable for our research case. As described in section 1 and 2, the FloReST mobile application aims at crowdsourced data collection by the citizens of the pilot communes, similarly to a risk detector. Tab. 1 provides an overview of the twelve mobile applications identified. The applications nbs. 3, 5 and 6 appeared to be of great interest for the current research, as they are dealing with data crowdsourcing. Hence, they were given a deeper investigation. No mobile application could be identified that deals with drain paths.

Nb.	Name of application	Editor	Goal
1	Alertswiss[8]	Federal Office for Civil Protection Switzerland	Warning
2	BIWAPP[9]	Marktplatz GmbH	Warning
3	Crowdwater[10]	SPOTTERON/University of Zurich	Water level measurement
4	Disaster Alert[11]	Pacific Disaster Centre	Warning
5	Floodcheck[12]	Credia Communications GmbH	Pre-Warning
6	FloodCitiSense[13]	JPI Urban Europe	Pre-Warning
7	MeinPegel[14]	Hochwasserzentralen	Water level measurement
8	MeteoSwiss[15]	Federal Office for Meteorology and climatology	WeatherService
9	NINA[16]	Federal Office for Civil Protection and Disaster Assistance Germany	Warning
10	PegelAlarm[17]	SOBOS	Water level measurement
11	RiverApp[18]	Florian Bessière	Water level measurement
12	WarnWetter[19]	German Weather Service	Weather Service

Tab. 1: Overview mobile applications

[7] The applications were identified based on geographical requirements (only Europe) and based on accessibility.
[8] https://www.alert.swiss/de/app.html
[9] https://www.biwapp.de/
[10] https://crowdwater.ch/de/start-2/
[11] https://disasteralert.pdc.org/disasteralert/
[12] https://www.eglv.de/
[13] http://www.floodcitisense.eu/
[14] https://www.hochwasserzentralen.info/meinepegel/
[15] https://www.meteoswiss.admin.ch/home/services-and-publications/beratung-und-service/meteoswiss-app.html
[16] https://www.bbk.bund.de/DE/Warnung-Vorsorge/Warn-App-NINA/warn-app-nina_node.html
[17] https://pegelalarm.com/
[18] https://www.riverapp.net/de
[19] https://www.dwd.de/DE/leistungen/warnwetterapp/warnwetterapp.html

CrowdWater (3)

Crowdwater works twofold to gather data about fluvial flooding: observing water levels and indicating rainfall intensity. Concerning the first, users are requested to take photos of different water areas (e.g. a river) and demonstrate the water level by adding a measurement scale into the photo. Repeating this process at different points in time, the differences in water levels can be examined. For transparency and visuality, the different spots are marked within a map. As far as the rainfall intensity is concerned the application provides icons showing different rainfall scenarios. By clicking the accurate scenario, citizens provide data about rainfall intensity at a specific location. For data quality control, a gamification approach is used. Every month, users can compete in a community championship with the aim to control and correct as many data entries as possible. Points are distributed for each comment and/or correction and the three first placed are rewarded. Though, it is not clear how Crowdwater verifies the corrected data within the championship. Crowdwater is functioning in the city of Zurich solely.

FloodCitiSense (5)

The mobile application FloodCitiSense aims at using Citizen Science to develop an exact and well-functioning early warning system for pluvial flooding. Citizens are able to provide information about the rainfall intensity via a mobile application. In addition, low-cost rainfall sensors were distributed to citizens to report on rain gauges and complement the official rainfall data collected by the authorities. Nonetheless, the application is only running in three pilot cities in Europe: Rotterdam, Brussels and Birmingham. Data control mechanisms were hardly to examine.

Floodcheck (6)

Floodcheck rather serves as flood risk management tool for house-owners than a Citizen Science application. With the help of Floodcheck, house-owners are able to identify the flood risk for their properties as well as receive information about property securing. The mobile application is a project of the German regional cooperation *Emscher Genossenschaft und Lippe Verband* and thus solely available for house-owners within the regional environs.

4.2 Reflecting on Citizen Science mobile application criteria and first draft requirement catalogue

Within this step, the information gathered in section 4.1 as well as the mobile application's intention described in section 2 were taken up and applied to Lemmens et al.'s [Le21] criteria for Citizen Science mobile application development. In the following, a brainstorm of application features is presented as a first draft requirement catalogue, which the authors perceive as appropriate for a Citizen Science mobile application within the context of FloReST. Lemmens et al. [Le21] proposes in their work seven criteria: Look and Feel, Re-use, Co-creation, User Interface Design, Participant Motivation, Supporting

Infrastructure, Testing and Maintenance. At the current research stage, three of these criteria (Participant Motivation, Re-use, supporting Infrastructure) were taken up for the beginning, the remaining will be deeper investigated in the future research process. Moreover, two further criteria were added, namely Functionality and Data control [Ge21], as they appeared important to the researchers to elaborate on in the context of the research project. Tab. 2 presents an overview of the selected criteria and the corresponding brainstormed features. It needs to be pointed out that the following table is not a final requirement catalogue but rather a brainstorm of thoughts and ideas for the further application development process.

Functionality	**Participant Motivation**	**Re-use**	**Supporting Infrastructure**	**Data control**
What can be entered in the application?	*How to motivate citizens to use the application?*	*Is the app only applicable to FloReST?*	*Which technical infrastructure is needed?*	*How can data control be guaranteed?*
Points, lines and surfaces	User journey	Communi-cation tool	Map services	Data protection
Selection (text, photos, icons)	Appetizer	Data collection	Meteorological data	Manual control
Photo upload	Understandable	Data for internal purpose	Hybrid	Contact person
Free text entries	Transparency		Free software	
Problem report Fix location	Illustrative		Angula, React	

Tab. 2: Overview criteria and features

5. Limitations and outlook

The inspection of different mobile applications marked a first step in the process of Citizen Science mobile application development and in answering the question of which preliminary features should be included in the FloReST Citizen Science application. Accordingly, this information built the base of the follow-up brainstorm described in section 4.2. Being complemented with literature, a first draft of a requirement catalogue was produced.

The current research proposal is mainly focusing on the topic of Citizen Science, as this was the initial step in the early beginning process of Citizen Science application development. Since no mobile application for drain path identification was found, the issue of drain path awareness needs to be deeper investigated in the future in order to find interesting and inspiring ways to attract users to report on such. Regarding this aspect,

further research is also needed in the field of citizen attraction and how to make them using such an application.

In addition, the issue of data quality control is indispensable to be reflected in the further research process, from a theoretical but also from a practical perspective. As false data might distort the final drain path identification and risk mapping, a control is of great importance. In the case of the FloReST project, it can be assumed that the number of Citizen Scientists participating in the data collection process might be manageable. Thus, a manual control could be a low-cost and viable option.

Further, the discussion of the user interface design and testing process remains open. The application development process has not yet reached that point. This short paper is a report of the current stage of work and first thoughts considering the task of developing a mobile application in the context of the FloReST project. It cannot be understood as a full and completed work.

Nonetheless, this paper gives a first impression of the challenges within a Citizen Science mobile application development in general as well as first insights into the research of the Institute for Software Systems, especially into the FloReST project. In the future process of research, the questions of data control and validation, as well as data privacy, user attractiveness and prototyping will be addressed.

Bibliography

[ACC15] Alfonso, L.; Chacon, J.; Peña-Castellanos, G.: Allowing citizens to effortlessly becoming rainfaill sensors. E-proceedings of the 36th IAHR World Congress, The Hague, 2015.

[Ap22] Apogaeis, www.apogaeis.com/blog/mobile-application-development-top-10-factors-to-consider/, last access: 20.05.2022.

[Ar18] Arthur, R.; Boulton, C. A.; Shotton, H. et al.: Social sensing of floods in the UK. PLoS One, 13(1), 2018.

[AS17] Albahari, A.; Schultz, C. H.: A Qualitative Analysis of the Spontaneous Volunteer Response to the 2013 Sudan Floods: Changing the Paradigm. Prehospital and Disaster Medicine, 32(03), 240–248, 2017.

[BBJ09] Berndtsson, J.C.; Bengtsson, L.; Jinnob, K.: Runoff water quality from intensive and extensive vegetated roofs. Ecological Engineering, 35 (3), 369-380, 2009.

[Be10] Berndtsson, J.C.: Green roof performance towards management of runoff water quantity and quality: A review. Ecological Engineering, 36 (4), 351-360, 2010.

[Be19] Becker, J.S.; Potter, S.H.; McBride, S.K. et al.:When the earth doesn't stop shaking: How experiences over time influenced information needs, communication, and interpretation of aftershock information during the Canterbury Earthquake Sequence, New Zealand. International Journal of Disaster Risk Reduction, 34, 397-411, 2019.

[BNR09] Bliss, D. J.; Neufeld, R.D.; Ries, R.J.: Storm Water Runoff Mitigation Using a Green Roof. Environmental Engineering Science, 26 (2), 407-418, 2009.

[Bo09] Bonney, R.; Cooper, C. B.; Dickinson, J. et al.: Citizen science: a developing tool for expanding science knowledge and scientific literacy. Bioscience 59, 977–984, 2009.

[Bu11] Burkhardt, M.; Zuleeg, S.; Vonbank, R. et al.: Leaching of additives from construction materials to urban storm water runoff. Water Sci Technol, 63 (9), 1974–1982, 2011.

[CCO21] Charters, F. J.; Cochrane Aisling, T.A.; O'Sullivan, D.: The influence of urban surface type and characteristics on runoff water quality. Science of the Total Environment, 755 (1), 142470, 2021.

[CNS11] Crall, A. W.; Newman, G.J.; Stohlgren, T.J.: Assessing citizen science data quality: an invasive species case study. Conservation Letters, 4, 433–442, 2011.

[Co08] Cohn, J.P.: Citizen Science: Can Volunteers do real research?. BioScience, 58(3), 192-197, 2008.

[DBZ10] Dickinson, J.L.; Bonter, D. N.; Zuckerberg, B.: Citizen Science as an Ecological Research Tool: Challenges and Benefits. Annual review of Ecology and Systematics, 41, 149-172, 2010.

[Di12] Dickinson, J.L.; Shirk, J.; Bonter, D. et al.: The current state of citizen science as a tool for ecological research and public engagement. Front Ecol Environ, 10(6), 291-297, 2012.

[Fa17] Farnham, D. J.; Gibson, R.A.; Hsueh, D.Y. et al.: Citizen science-based water quality monitoring: Constructing a large database to characterize the impacts of combined sewer overflow in New York City. Science oft he Total Environment, 580, 168-177, 2017.

[Fa18] Fava, M.C.; Abe, N.; Restrepo-Estrada, C. et al.: Flood modelling using synthesised citizen science urban streamflow observations. Journal of Flood Risk Management, 12 (2), e12498, 2018.

[Fr22] Franklin, B.: Quora, www.quora.com/Where-and-when-did-Benjamin-Franklin-say-Tell-me-and-I-forget-teach-me-and-I-may-remember-involve-me-and-I-learn, 2022.

[FS18] Feng, Y.; Sester, M.: Extraction of pluvial flood relevant volunteered geographic information (VGI) by deep learning from user generated texts and photos. ISPRS Int. J. Geo Inf., 7 (2), 2018.

[GDC07] Göbel, P.; Dierkes, C.; Coldewey, W.G.: Storm water runoff concentration matrix for urban areas. Journal of Contaminant Hydrology, 91 (1-2), 26-42, 2007.

[Ge21] Geyer, H.: Citizen Science - praktische Erfahrungen am Beispiel des ArtenFinders Rheinland-Pfalz. Fischer-Stabel, P. (ed): Umweltinformationssysteme: Grundlagen einer angewandten Geoinformatik/Geo-IT (3rd Edition), Wichmann Verlag, 2021.

[Go21] Goded, T.; Tan, M.L.; Becker, J.S. et al.: Using citizen data to understand earthquake impacts: Aotearoa New Zealand's earthquake Felt Reports. Australasian Journal of Disaster and Trauma Studies, 25(3), 61-78, 2021.

[Ha13] Haklay, M.: Citizen Science and Volunteered Geographic Information: Overview and Typology of Participation. Crowdsourcing Geographic Knowledge. D. Sui et al. (eds.):

Crowdsourcing Geographic Knowledge: Volunteered Geographic 105 Information (VGI) in Theory and Practice, 105-122. Springer Science+Business Media, 2013.

[Ha15] Haklay, M.: Citizen science and policy: A European perspective. Woodrow Wilson International Center for Scholars, 2015.

[He21] Hegarty, S.; Hayes, A.; Regan, F.; Bishop, I.; Clinton, R.: Using citizen science to understand river water quality while filling data gaps to meet United Nations Sustainable Development Goal 6 objectives. Science of the Total Environment, 783, 146953, 2021.

[Ka20] Karvonen, A.: Urban Techno-Politics: Knowing, Governing, and Imagining the City. Science as Culture, 29(3), 417–424, 2020.

[Ke21] Kempf, H.: Experience from large-scale crowdsourcing via weather apps. Australasian Journal of Disaster and Trauma Studies, 25(3), 87-93, 2021.

[Ko16] Kosmala, M.; Wiggins, A.; Swanson, A. et al.: Assessing data quality in citizen science. Front Ecol Environ, 14(10), 551-560, 2016.

[Le17] Lévesque, D.; Cattaneo, A.; Deschamps, D. et al.: In the eye of the beholder: Assessing the water quality of shoreline parks around the Island of Montreal through citizen science. Science of the Total Environment, 579, 978-988, 2017.

[Le21] Lemmens, R.; Antoniou,V.; Hummer, P. et al.: Citizen Science in the Digital World of Apps. Vohland, K. et al. (eds.): The Science of Citizen Science, 461-474, 2021.

[Li18] Lin, W.-Y.; Wu, T.-H.; Tsai, M.-H. et al.: Filtering disaster responses using crowdsourcing. Automation in Construction, 91, 182–192, 2018.

[MCR13] Morgan, S.; Celik, S.; Retzlaff, W.: Green Roof Storm-Water Runoff Quantity and Quality. Journal of Environmental Engineering, 139 (4), x-x, 2013.

[NLR22] Newson, M.; Lewin, J.; Raven, P.: River science and flood risk management policy in England. Progress in Physical Geography, 46(1), 105-123, 2022.

[Pa18] Paul, J. D.; Buytaert, W.; Allen, S. et al.: Citizen Science for hydrological risk reduction and resilience building. WIRES Water, 5, 2018.

[PLC00] Pagotto, C.; Legret, M.; Le Cloirec, P.: Comparison of the hydraulic behaviour and the quality of highway runoff water according to the type of pavement. Water Research, 34 (18), 4446-4454, 2000.

[PN16] Pandey, N.; Natarajan, S.: How social media can contribute during disaster events? Case study of Chennai floods 2015. 2016 International Conference on Advances in Computing, Communications and Informatics, 2016.

[RP13] Riesch, H.; Potter, C.: Citizen science as seen by scientists: Methodological, epistemological and ethical dimensions. Public Understanding of Science, 23(1), 107–120, 2013.

[Sc21] Scaini, A.; Stritih, A.; Brouillet, C. et al.: Flood Risk and River Conservation: Mapping Citizen Perception to Support Sustainable River Management. Front. Earth Sci., 9, 2021.

[Se19] See, L.: A Review of Citizen Science and Crowdsourcing in Applications of Pluvial Flooding. Front. Earth. Sci, 7 (44), 2019.

[Sh20] Shinbrot, X.A.; Muñoz-Villers, L.; Mayer, A. et al.: Quiahua, the First Citizen Science Rainfall Monitoring Network in Mexico: Filling Critical Gaps in Rainfall Data for Evaluating a Payment for Hydrologic Services Program. Citizen Science: Theory and Practice, 5 (1), 19, 2020.

[St20] Stavi, I.; Siad, S. M.; Kyriazopoulos, A.P.; Halbac-Cotoara-Zamfir, R.: Water runoff harvesting systems for restoration of degraded rangelands: A review of challenges and opportunities. Journal of Environmental Management, 255, 109823, 2020.

[Th17] Thornhill, I.; Ho, J. G.; Zhang, Y. et al.: Prioritising local action for water quality improvement using citizen science; a study across three major metropolitan areas of China. Science of the Total Environemnt, 584-585, 1268-1281, 2017.

[Um22] Umwelt-Campus Birkenfeld, www.umwelt-campus.de/forschung/projekte/iot-werkstatt/pegel-juenkerath, last access: 20.05.2022.

[Vi21] Vinnell, L. J.; Becker, J. S.; Scolobig, A. et al.: Citizen science initiatives in high-impact weather and disaster risk reduction. Australasian Journal of Disaster and Trauma Studies, 25(3), 55-60, 2021.

[We16] Wein, A.; Potter, S.; Johal, S. et al.: Communicating with the public during an earthquake sequence: Improving communication of geoscience by coordinating roles. Seismological Research Letters, 87(1), 112-118, 2016.

[We19] Weeser, B.; Jacobs, S.; Kraft, P. et al.: Rainfall-Runoff Modeling Using Crowdsourced Water Level Data. Water Resources Research, 55 (12), 10856-10871, 2019.

[Wo21] Wolff, E.; French, M.; Ilhamsyah, N. et al.: Collaborating with communities: Citizen Science Flood Monitoring in Urban Informal Settlements. Urban Planning, 6 (4), x-x, 2021.

[WOW77] Wisner, B.; O'Keefe, P.; Westgate, K.: Global Systems and Local Disasters: The Untapped Power of People's Science. Disasters, 1(1), 47–57, 1977.

Evolution of Spatial Disaster Risk Assessments: a Bibliometric Analysis

Ingrid Wong [1], Klaus Greve [2] and Jörg Szarzynski[3,4,5,6]

Abstract: The advancements and availability of geospatial technology and data, combined with a growing understanding of the importance of disaster management and disaster risk reduction, lead to an increase in quantity and quality of publications in spatial disaster risk assessments. This paper presents an overview of an in-depth bibliometric study of the global evolution of such spatial disaster risk assessments from 2000 to 2021. The study revealed an upwards trend in scientific production, with hydrological hazards dominating the research field. However, climate change may further drive research for meteorological and climatological hazards. No direct influence was found between major disasters and the number of assessments. The study also shed light on the conceptual frameworks that guided spatial risk assessments, with hazard and vulnerability being the essential components.

Keywords: Disaster, Spatial Risk Assessment, Disaster Management, Disaster Risk Reduction, Bibliometric Analysis

Addresses Sustainable Development Goal 13: Climate action

1. Introduction

A crucial component in disaster management and disaster risk reduction is disaster risk assessment. This approach can be quantitative or qualitative, seeking to determine the risk of being impacted by disasters through identifying the hazards and their characteristics, analysing the exposure and vulnerability from the existing conditions of people, the environment they depend on, and everything in between, as well as evaluating existing preventative measures and coping capacities, and pinpoint contributing factors that leads to a disastrous event [Un17]. One of the most common approaches of disaster risk

[1] Meteorological Service of Canada, Environment and Climate Change Canada, 4905 Dufferin St, North York, ON, M3H 5T4, Canada, wong.ingrid.cc@gmail.com, https://orcid.org/ 0000-0003-3758-8664
[2] Center for Remote Sensing of Land Surface, University of Bonn, Meckenheimer Allee 166, 53115 Bonn, Germany, klaus.greve@uni-bonn.de, 0000-0001-6463-6161
[3] Institute for Environment and Human Security, United Nations University, Platz d. Vereinten Nationen 1, 53113 Bonn, Germany, szarzynski@ehs.unu.edu
[4] Center for Global Mountain Safeguard Research, Eurac Research, Bolzano, Italy
[5] Disaster Management Training and Education Centre for Africa (DiMTEC), University of the Free State (UFS), Bloemfontein, South Africa
[6] International Research Institute of Disaster Science (IRIDeS), Tohoku University, Sendai, Miyagi, Japan

assessment is through quantifying and visualising risks by combining and superimposing geospatial layers of information on hazard, exposure, vulnerability and capacity using geographic information systems (GIS) and remote sensing technologies. This approach is called spatial risk assessments [Fe12, Un19].

In this paper we provide an overview regarding first results of an in-depth bibliometric study of the global evolution of spatial disaster risk assessments from 2000 to 2021. Based on own experiences we know, that the enhanced availability of PC based GIS instruments, the increasing amount of free and open satellite products, as well as crowd based spatial, tabular and statistics data has led to an increase in number and quality of publications in spatial risk assessments. At the same time, a growing understanding can be observed of the importance of reducing the impact of natural hazards through systematic efforts to analyse and reduce the causal factors of disasters.

With our research we want to identify and illustrate, how the field of disaster spatial risk assessment has evolved in quantity and quality, and how disaster dynamics, especially those of major disasters and related conceptual and organisational frameworks, have influenced the development of this body of literature. This holds true, in particular, regarding the United Nations Disaster Risk Reduction (UNDRR) guidelines – the Hyogo Framework for Action 2005-2015 [Un05] and its successor, the Sendai Framework for Disaster Risk Reduction (SFDRR) 2015-2030 [Un15].

2. Methodology

A bibliometric and quantitative content analysis was conducted on global publications related to spatial risk assessments of natural hazards (geophysical, meteorological, hydrological and climatological), and limited to research papers published exclusively in English between 01 January 2000 and 30 September 2021. Bibliographic data was retrieved from the Web of Science (WoS) through a tailored search query on 29 October 2021. Four major parts in the search query were relevant: spatial risk assessment related terms, disaster related terms to focus on journal articles in the disaster risk reduction domain, hazard related terms based on the natural hazard definitions on the Emergency Events Database (EM-DAT), and exclusionary criteria that sort out papers that were out of scope for this study.

The full record of cited references of 1687 total results was exported into Plain Text File format and checked for duplicates and errors using Citavi. All 1687 entries remained after data brush-up and were imported to various software packages for further data analysis. The *Bibliometrix R-package* and its web-based application *Biblioshiny*, for instance, were used to conduct a descriptive analysis on the bibliographic data including summarising the publication year, journal and country. These applications were used to a) conduct co-word analyses on the keywords including retrieving the top 30 most frequently occurred keywords and plotting the thematic evolutions, and b) for co-citation analysis to explore the intellectual structure of the research field through identifying the most influential

publications over the 21-year period. Additionally, the number of articles per hazard type was derived through running the search query by each individual hazard type on WoS. Lastly, findings from the bibliometric analysis were further compared and connected with recorded disaster trends and major disaster events from EM-DAT, and with the timeline of the adoption of the UNDRR guidelines.

3. Results

3.1 Overview

From 01 January 2000 to 30 September 2021, 1687 papers were published in the research field of spatial disaster risk assessment, accounted, on average, for 80.3 papers per year. As seen in Fig. 1, 2001 had the least number of articles with just 0.1% of the total number of articles, and the number of articles published peaked in 2020 with 16.3% of the total number of articles. However, it is notable, that the number of articles in 2021 only counted till the end of September of that year, therefore, the number of articles published in 2021 could be higher than the recorded number in this study which was 246. The annual scientific production growth rate was 19.34%.

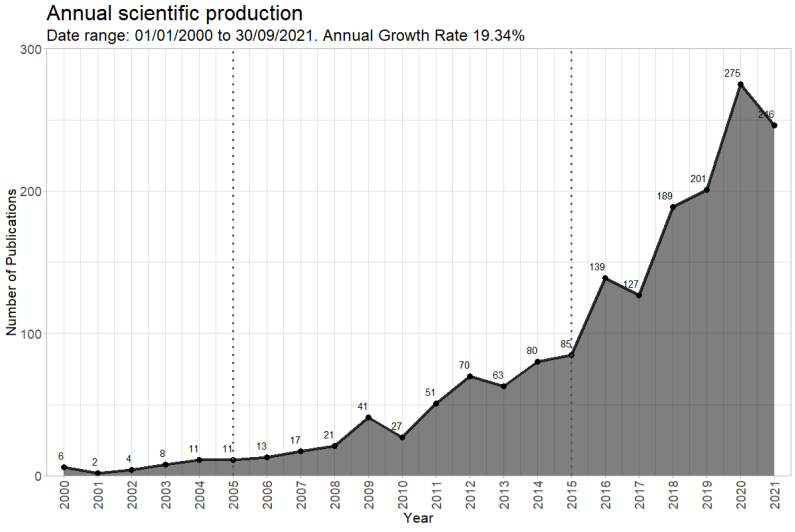


Fig.1: Annual scientific production from 01/01/2000 to 30/09/2021.
Vertical dotted lines are marked at 2005 and 2015, which were the years when the Hyogo Framework and Sendai Framework were adopted respectively.

There was a total of 5673 authors that contributed to the 1687 articles, with only 79 single-authored documents, putting the average collaboration index at 3.48, also implicating that

author collaborations are common in this research field.

A total number of 387 journal sources housed the 1687 articles, with the top five sources having over 50 articles each and 26% of all the articles combined. *Natural Hazards* was clearly at the lead, with 12.7% of all the articles published, and three times more compare to the journal with the second most articles (4.2%). Tab. 1 lists the top 15 journals.

Rank	Journal Name	Article count
1	NATURAL HAZARDS	214
2	NATURAL HAZARDS AND EARTH SYSTEM SCIENCES	71
3	WATER	54
4	INTERNATIONAL JOURNAL OF DISASTER RISK REDUCTION	52
5	SUSTAINABILITY	51
6	GEOMATICS NATURAL HAZARDS & RISK	43
7	LANDSLIDES	39
8	SCIENCE OF THE TOTAL ENVIRONMENT	39
9	JOURNAL OF FLOOD RISK MANAGEMENT	35
10	ENVIRONMENTAL EARTH SCIENCES	25
11	JOURNAL OF HYDROLOGY	25
12	REMOTE SENSING	24
13	ENGINEERING GEOLOGY	22
14	ISPRS INTERNATIONAL JOURNAL OF GEO-INFORMATION	22
15	STOCHASTIC ENVIRONMENTAL RESEARCH AND RISK ASSESSMENT	20

Tab. 1: Top 15 journal sources and their respective counts of articles.

The top five most cited papers within the 1687 publications are listed in Tab. 2, with the most cited paper being cited 750 times. All five papers were published before 2015 and four out of the five papers were published before 2010. These five papers were methodological focused that serve as guidelines for conducting risk assessments, hence were more widely cited. Three out of the five papers were landslide related, the other two were related to floods and wildfires, respectively.

Rank	Cited information and publication title	Cited count
1	Title: Landslide risk assessment and management: an overview DAI FC, 2002, ENG GEOL DOI: 10.1016/S0013-7952(01)00093-X	750
2	Title: Recommendations for the quantitative analysis of landslide risk COROMINAS J, 2014, B ENG GEOL ENVIRON	479

3	DOI: 10.1007/s10064-013-0538-8 Title: Flood risk assessment and associated uncertainty APEL H, 2004, NAT HAZARD EARTH SYS DOI: 10.5194/nhess-4-295-2004	300
4	Title: Human-caused wildfire risk rating for prevention planning in Spain MARTINEZ J, 2009, J ENVIRON MANAGE DOI: 10.1016/j.jenvman.2008.07.005	291
5	Title: Global landslide and avalanche hotspots NADIM F, 2006, LANDSLIDES DOI: 10.1007/s10346-006-0036-1	274

Tab. 2: Top five globally most cited publications within the 1687 search results.

Of the ten hazard types mentioned within all publications we can see the following statistics (Fig. 2): floods clearly dominated the numbers with its related terms mentioned in 41.6% of the papers, followed by landslides and mass movements mentioned in 20.6% and 17.2% of the papers, respectively, then storms in 13.3% of the papers, earthquakes in 11% of the papers, droughts in 7.3% of the papers, wildfires in 5.8% of the papers, volcanic activities and extreme temperatures were mentioned in 3.1% of the papers, and glacial lake outburst in just 0.5% of the papers. This trend is supported by the CRED & UNDRR reports with floods being the most common hazard, storms being the costliest hazard, and earthquakes being the deadliest hazard worldwide from 2000 to 2020 [CU20, CU21]. Landslides and mass movements were well-assessed even though the number of events, affected people and deaths, as well as the disaster cost were not among the worst, which could be due to the fact that landslides are closely related with floods. Droughts and extreme temperatures were under-reported as indicated in [CU20] and under researched based on the results of this study.

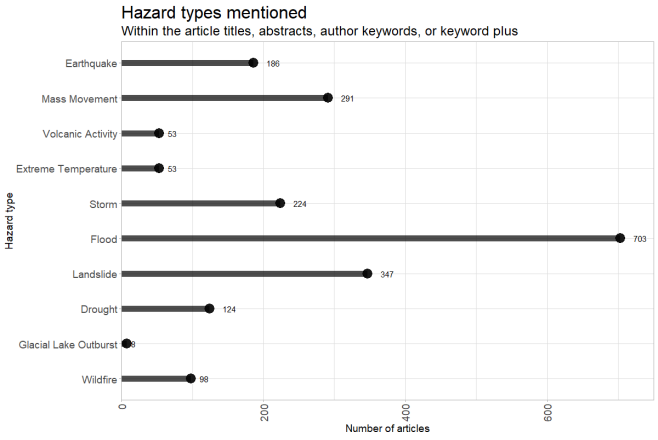


Fig.2: The number of articles with any of the ten hazard types mentioned within their bibliographic data (titles, abstracts, author keywords and keyword plus).

3.2 The impact of major disasters

The top five deadliest disasters from 2000 to 2019 were the 2004 Indian Ocean seaquake and tsunami, which directly affected Indonesia, Sri Lanka, India, Maldives, Thailand, Somalia, Myanmar, Seychelles and Malaysia, the 2010 Haiti earthquake, the tropical cyclone Nargis in Myanmar, the Sichuan earthquake in China in 2008, and the 2005 Kashmir earthquake in Pakistan [CU20, Kw14]. Fig. 3 and Fig. 4 plotted the annual assessments conducted for the affected countries.

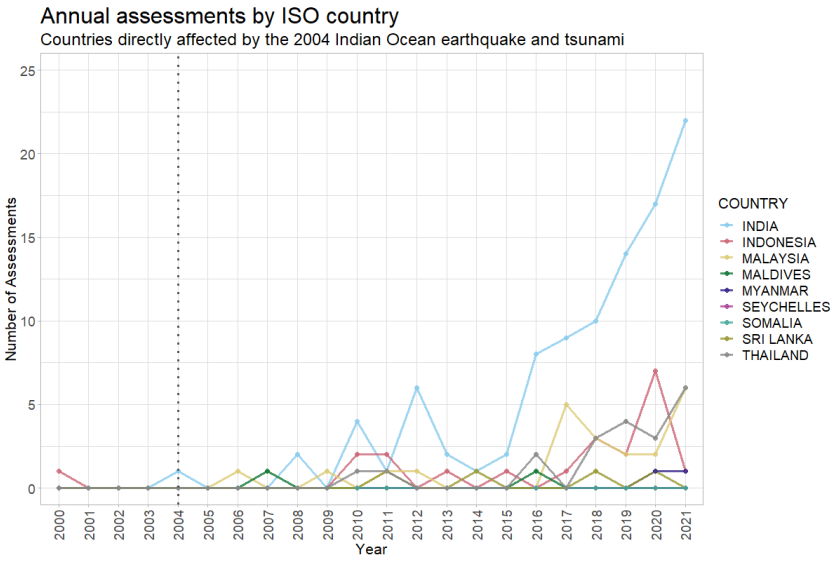


Fig. 3: Annual number of assessments conducted on the countries directly affected by the 2004 Indian Ocean seaquake and tsunami.

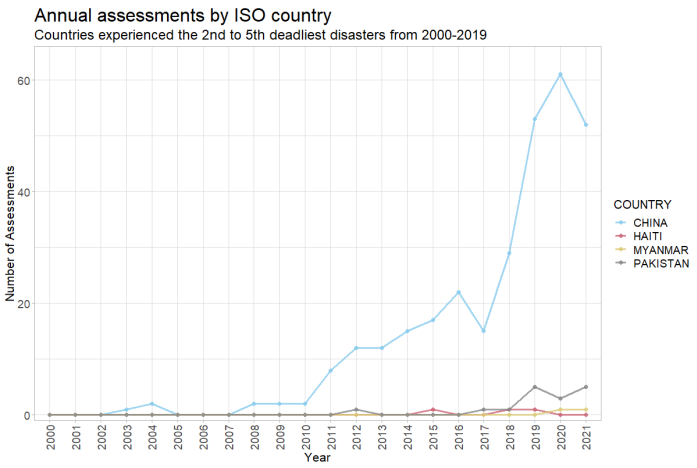


Fig. 4: Annual number of assessments conducted on the countries that experienced the second to fifth deadliest disasters from 2000 to 2019.

From the time series line graph for the countries devastated by the 2004 Indian Ocean seaquake and tsunami (Fig. 3), the number of assessments conducted on the selected countries did not seem to have increased post-2004. Number of assessments started to pick up rapidly for India and gradually but minimally for Indonesia, Malaysia and Thailand over the later years in the 2010s. From the time series line graph for the second to fifth deadliest disasters from 2000 to 2019 (Fig. 4), the number of assessments conducted for China started to increase in 2010 and continued to increase in a rapid rate in 2017. However, the number of assessments did not increase for Haiti after 2010, or Myanmar after 2008, and the assessments conducted for Pakistan started to pick up in 2019. Based on these findings, we see no direct reaction within the body of literature regarding major events.

3.3 Influence of conceptual frameworks

The graph of the annual production (Fig. 1, dotted lines) does not seem to show a shift in the growth pattern in 2005, when the Hyogo Framework was adopted, but the change of growth pattern coincided with the adoption of the SFDRR in 2015, where rapid increases of new publications were unprecedented. Hence, this data is indicating that the SFDRR may have contributed to the promotion of research of spatial risk assessment through its targets and priorities for action which emphasised the importance of this research field.

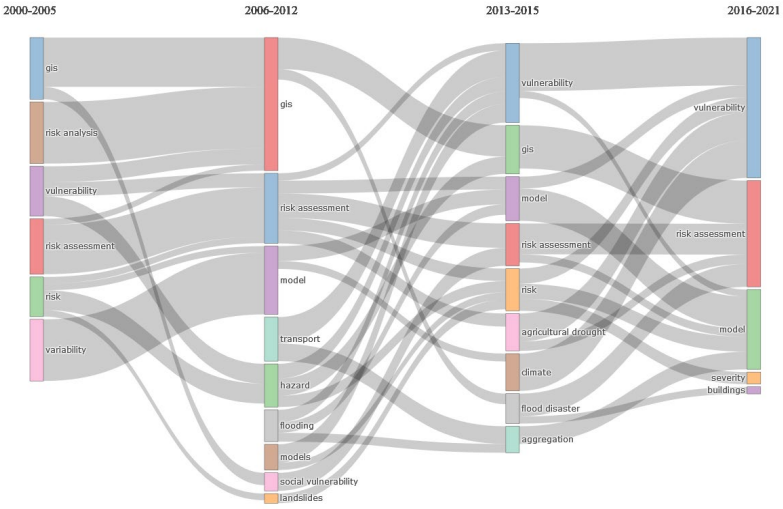


Fig. 5: Thematic evolution diagram of the top 10% most frequently occurred keywords grouped into various themed clusters, displaying how the clusters diverged, converged, and evolved. Three time slices were set with cut-off years at 2005, when the Hyogo Framework was adopted, 2012 when the IPCC SREX was published, and 2015 when the Sendai Framework was adopted.

A thematic analysis of the top 10% most frequently occurred keywords illustrated that in the past 21 years, the spatial risk assessment research field had its mainstream focus on "vulnerability", "climate change", "flood risk", and "impact(s)". Basic themes that form the foundation of the research field were "risk assessment", "GIS", "hazard", "analytic hierarchy process (ahp)", and "susceptibility". Additional basic themes, that were either still to be developed, or were past mainstream themes such as "risk", "model", "uncertainty", "system", and "inundation". Niche themes that were developed, but not widely researched on, were "management", "simulation", "probability", "buildings", "methodology", and seismic-related themes. Further breaking down the thematic analysis into thematic time slices showed the evolution of the research field over set time slices (Fig. 5). Prior to the adoption of the Hyogo Framework in 2005, research themes revolved around "GIS", "risk analysis", "vulnerability", "risk assessment", "risk", and "variability". From 2006 to the publishing of the IPCC SREX in 2012, the themes "GIS" and "risk assessment" retained and grew while the other themes diverged into the retained themes as well as themes like "model(s)" and "hazard". Smaller themes emerged such as "transport" and "flooding". Starting with 2013, until the adoption of the SFDRR in 2015, smaller themes converged and formed back into the major theme "vulnerability", while "GIS", "model" and "risk assessment" became more foundational themes and broke into smaller and more specific themes while retaining a portion of its importance in the field. New themes such as "climate", "agricultural drought", and "aggregation" emerged. From 2016 onwards until September 2021, "vulnerability", "risk assessment", and "model" continued to congregate smaller themes from the previous time slices and form into larger

mainstream and foundational themes. Smaller niche themes such as “severity” and “buildings” also developed from diverged themes.

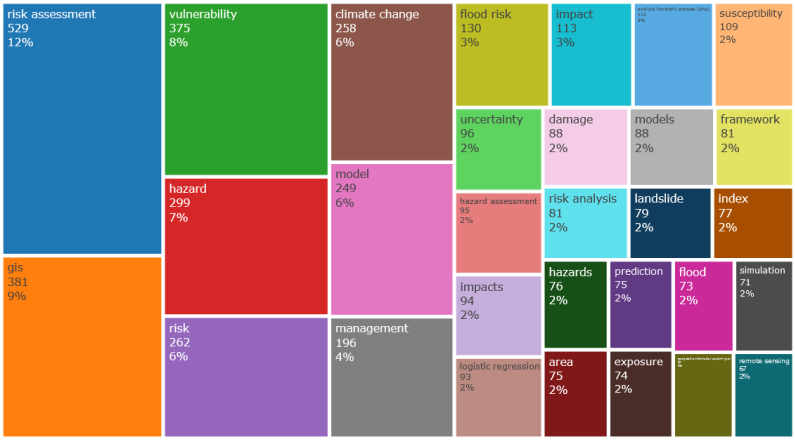


Fig. 6: TreeMap of the top 30 most frequently occurred keywords.
In total there were 4461 occurrences of the top 30 keywords. Numbers under the terms indicate the frequencies and their percentages within the top 30 occurrences.

As for “exposure”, although it was present in the top 30 keywords (Fig. 6), the frequency of occurrence was a lot less compared to its counterparts in the function of risk. The reason for this, could be that exposure was considered to be a sub-component of vulnerability prior to the adoption of the Hyogo Framework in 2005, when the main school of thought of risk was a function of just hazard and vulnerability [Wi04]. Throughout the first decade, of the 21st century into early 2010s, the concept of exposure has been a sub-component of vulnerability in various main schools of thoughts [Bi13, Bo01, Tu03]. It was not until the IPCC SREX, published in 2012, that risk was introduced as a function of hazard, vulnerability, and exposure, separated from vulnerability [In12]. This could imply that the usage of “exposure” as a keyword in the papers was not common until after 2012. The final component of the risk equation is “capacity”, but the term did not make it into the top 30 keywords. This term was brought up as a risk assessment component side by side with vulnerability in early 2000s [Wi04], but it was mainly considered a sub-component of vulnerability in the 2010s [Bi13], and reached back to a similar level of importance with the other risk components in the disaster risk reduction research community in 2017 [Un17]. The keyword “susceptibility” was mentioned more often than “exposure”. This term was not part of the search query, but like “exposure” and “capacity”, it is also one of the sub-components of vulnerability defined in the [Un17] definition chosen for this study, as well as other schools of thought such as [Bi13]. These findings illustrate, that there was a large variation on how authors in this research field defined risk, and what kind of risk frameworks they have chosen within their assessments, as further observed from “framework” being a top 30 keyword in this study.

4. Conclusions

Over the past 21 years, the research field of spatial disaster risk assessment was in an upwards trend regarding annual scientific production. A potential trigger of increased scientific production was the adoption of the SFDRR in 2015, which stressed the importance of understanding risk and risk assessment methodologies. However, major disasters did not trigger an increase of assessments in the affected countries.

The most representative journals for the research field were *Natural Hazards* and *Natural Hazards and Earth System Sciences (NHESS)*, which are journals that exhibit stronger focus on hazards compared to disasters or disaster risk. Therefore, it is also worth mentioning the journal that published the fourth-highest number of papers, which was the *International Journal of Disaster Risk Reduction (IJDRR)*. It would be interesting to see, in the next years and decades to come, whether researchers in this field would shift towards journals that are more risk-focused.

Hazard and vulnerability were the essential components of spatial risk assessments, and the other components of risk were dependent on the frameworks that researchers chose to utilize. Hydrological hazards (i.e. floods and landslides) dominated the research field, geological hazards (i.e. earthquakes and volcanic activities) were niche, but well-developed topics in the field, while meteorological (i.e. storms and extreme temperatures) and climatological hazards (i.e. droughts and wildfires) received less attention. However, this may change because it is believed that climate change will further drive research in this field as the effects of natural hazards are expected to be exacerbated as climate change continues.

The authors hope, that this study provided a supportive reference for future research and collaboration in the disaster management and disaster risk reduction domain, as well as a guide for conducting similar bibliometric analysis or content analysis in the future.

Bibliography

[Bi13] Birkmann, J.; Cardona, O. D.; Carreño, M. L.; Barbat, A. H.; Pelling, M.; Schneiderbauer, S.; Kienberger, S.; Keiler, M.; Alexander, D.; Zeil, P.; Welle, T.: Framing vulnerability, risk and societal responses: the MOVE framework. Natural Hazards 67, pp. 193-211, 2013.

[Bo01] Bohle, H.-G: Vulnerability and Criticality: Perspectives from Social Geography. IHDP Update, Bonn, pp. 2, 2001.

[CU20] Centre for Research on the Epidemiology of Disasters (CRED); United Nations Office for Disaster Risk Reduction (UNDRR): The Human Cost of Disasters: An overview of the last 20 years 2000-2019. 2020.

[CU21] Centre for Research on the Epidemiology of Disasters (CRED); United Nations Office for Disaster Risk Reduction (UNDRR): 2020: The Non-COVID Year in Disasters. 2021.

[Fe12] Fekete, A.: Spatial disaster vulnerability and risk assessments: challenges in their quality and acceptance. Natural Hazards 61, pp. 1161–1178, 2012.

[In12] Intergovernmental Panel on Climate Change (IPCC): Managing the risks of extreme events and disasters to advance climate change adaptation: A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change. Cambridge University Press, 2012.

[Kw14] Kweifio-Okai, C.: Where did the Indian Ocean tsunami aid money go? The Guardian, International Edition, 2014.

[Tu03] Turner, B. L.; Kasperson, R. E.; Matson, P. A.; McCarthy, J. J.; Corell, R. W.; Christensen, L.; Eckley, N.; Kasperson, J. X.; Luers, A.; Martello, M. L.; Polsky, C.; Pulsipher, A.; Schiller, A.: A framework for vulnerability analysis in sustainability science. Proceedings of the National Academy of Sciences of the United States of America 100(14), pp. 8074–8079, 2003.

[Un05] United Nations International Strategy for Disaster Reduction (UNISDR): Hyogo Framework for Action 2005-2015: Building the Resilience of Nations and Communities to Disasters. Geneva, Switzerland, 2005.

[Un15] United Nations International Strategy for Disaster Reduction (UNISDR): Sendai Framework for Disaster Risk Reduction 2015-2030. Geneva, Switzerland, 2015.

[Un17] United Nations Office for Disaster Risk Reduction (UNDRR), Terminology on Disaster Risk Reduction 2017, https://www.undrr.org/terminology, accessed: 04/10/2021.

[Un19] United Nations Office for Disaster Risk Reduction (UNDRR): Global Assessment Report on Disaster Risk Reduction 2019. Geneva, Switzerland, 2019.

[Wi04] Wisner, B.; Blaikie, P.; Cannon, T.; Davis, I.: At Risk: natural hazards, people's vulnerability and disasters, Second Edition, Routledge, New York, 2004.

Development of a framework for decision support in the context of climate adaptation

Simeon Wetzel [1] and Stephan Mäs [2]

Abstract: Implementing effective adaptation data that enable this knowledge transfer. For this purpose, this article presents a framework that shall simplify the development of these tools and the measures to climate change impacts concerns policymakers worldwide. At the local scale, there is a gap between scientific findings and a translation of these into concrete measures. It requires a network of tools and access to climate data. The underlying data management concept is intended to provide an infrastructure that requires only a few interventions in the operating process for both data suppliers and system administrators. Most of the infrastructure components have already been realised in the context of the ongoing research project KlimaKonform. In addition, there is an outlook on future implementations.

Keywords: Climate Change Adaptation, Climate Adaptation, SDI, Spatial Data Infrastructure, Data Management, Climate Impacts, FAIR

Addresses Sustainable Development Goal 13: Climate action

1. Project background and motivation

The impacts of climate change are a major global challenge that affect almost every region and all parts of society. Since not all climate impacts or avoidable, it is necessary to prepare and adapt to the potential risks like singular extreme weather events, potential hazards and foresee the consequences for agriculture and human health [Ip14]. It is therefore of particular importance that policymakers implement effective adaptation measures that mitigate the vulnerability of human and natural systems and, thus, pave the way to a climate-resilient future [Ip19]. Adaptation to and mitigation of climate change impacts is also addressed in one of the 13 Sustainable Development Goals (SDG; topic: Climate Change) of the United Nations (UN) [Un22]. Even though climate change impacts are in effect on a global scale, there is a need for locally tailored adaptation measures [DEA18].

In the context of the research project KlimaKonform[3], we investigate methods and

[1] Technische Universität Dresden, Chair of Geoinformatics, Helmholtzstraße 10, Dresden, 01069, simeon.wetzel@tu-dresden.de, https://orcid.org/0000-0001-7144-3376
[2] Technische Universität Dresden, Chair of Geoinformatics, Helmholtzstraße 10, Dresden, 01069, stephan.maes@tu-dresden.de, https://orcid.org/0000-0002-9016-1996
[3] https://klimakonform.uw.tu-dresden.de/

decision supporting tools that aim to guide local policymakers in the process of developing adaptation measures. The study area of the project is a rural low mountain region in central Germany[4]. Due to the topography and demography, as well as economic particularities, the region has a special exposure to climate impacts. In the course of the project, the physical, socio-economic and ecological climate impacts in the region will be modelled based on representative examples. In addition, the adaptation capacities and costs will be investigated. Ultimately, this requires a close cooperation between scientists and local stakeholders in order to identify their needs for the development of measures. The aim is to provide tailored climate information that can be directly translated into decision-relevant knowledge. Since the landscape type of the study area is similar to wide parts of Germany, a transferability of the project insights on other regions is expected.

Due to the contribution of many scientific partners from universities and scientific institutes, such as Technische Universität Dresden (TUD[5]), Leibnitz Institut für Ökologische Raumordnung (IÖR[6]), Helmholtz-Zentrum für Umweltforschung (UFZ[7]), the project relies on a large basis of climate analysis and data. Additionally, the existing climate information platform (ReKIS[8]) shall be extended. This platform provides a variety of heterogeneous climate data that cover the study area. This data will be used for the development of climate data products. The offered data range from future climate simulations (e.g. CMIP5[9], WEREX-V[10]), reference datasets (historic data) to raw data (from weather stations).

As a starting point for the development of the climate data products, it is required to bundle these data from various sources, describe them properly (using metadata) and enable possibilities for straightforward data exchange. In recent decades, the geospatial domain has been developing concepts for data management and exchange to support the integration of multiple data sources into a harmonized and coherent data set and the data provision in Spatial Data Infrastructures (SDI) [Be14, Ya19]. In this type of distributed infrastructure, standardized web interfaces are used (mostly specified by the Open Geospatial Consortium, OGC[11]) with the objective of providing interoperable (platform-independent) data exchange. A core component of a SDI are metadata. Metadata are often defined as data about data [Ya08]. They contain important information such as contact points, access information or technical properties (e.g. spatial coverage, spatial/temporal resolution) [Ha18]. In addition, metadata increase the discoverability of datasets, ensure improved transparency and at the best enable the re-use of data [Be14, SM18]. In SDIs metadata are usually managed in a structured form using standard metadata schemas that

[4] The project focuses on selected municipalities in the districts of Burgenlandkreis (Saxony-Anhalt), Greiz (Thuringia) and Vogtlandkreis (Saxony), For more detail, see the map under: https://klimakonform.uw.tu-dresden.de/index.php/modellregion
[5] https://tu-dresden.de/
[6] https://www.ioer.de/
[7] https://www.ufz.de/
[8] https://rekis.hydro.tu-dresden.de/
[9] https://www.wcrp-climate.org/wgcm-cmip/wgcm-cmip5
[10] For further information about this model, see: [BPS16]
[11] https://www.ogc.org/

enable machine readability [Wi16].

The data providers of the project and the information platform are either climate scientists or responsible employees from the federal state offices. So far, they have little or no previous knowledge to store metadata on specially designed platforms such as WebGIS servers. Experience from other projects has shown similar circumstances. Nevertheless, in order to collect as much metadata as possible, a concept is needed where data providers can store meta-information with tools they already know. In the interest of a sustainable system, this stored metadata is to be automatically applied to the various components of the SDI. The intention is to design a system that requires only minimum intervention from data providers and operators. This article describes the architecture specifics of this data management concept. With the proposed framework, we intend to create a shared and transparent data pool, which potentially closes the gap between scientific results and recommended actions for local policymakers.

2. Architecture concept

As a data management concept, we propose to implement approved elements of a spatial data infrastructure. The key component of our architecture is the existing climate information platform ReKIS. Further components are to be developed and then coupled with this platform (Fig. 1).

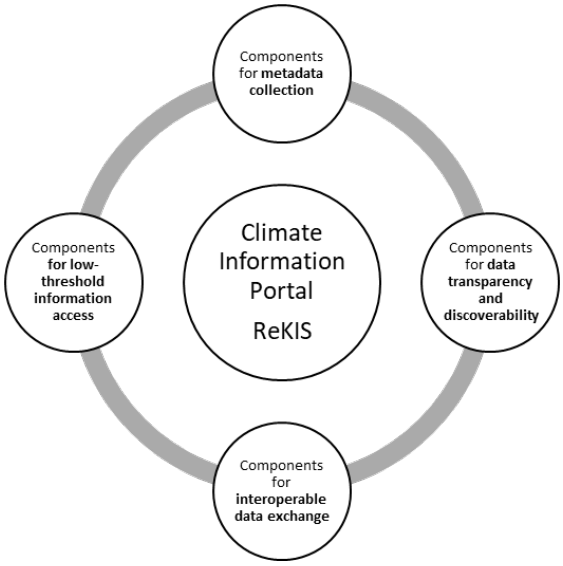


Fig. 1: New architecture components to be developed

1. Components for metadata collection:

   During the process of collection and publication of metadata in the original climate platform (ReKIS), we experienced some obstacles. Firstly, most of the data providers of our project and the ReKIS platform have no prior knowledge of standard metadata schemas or online tools for metadata collection. Therefore, they need practicable tools for metadata collection that do not add technical hurdles preventing them from depositing data and metadata.

   Hence, we decided to implement a semi-automatic approach similar to what has been presented in other works like [ATG22, No20]. This includes a simple Excel template as an easy-to-use method to collect the metadata (manually). In addition, we implemented tools that automatize the remaining process of integrating and distributing the metadata on the other (web-) components. The collected metadata is then transformed into the standardized schema GeoDCAT-AP[12] [RVR19]. GeoDCAT-AP extends the de-facto standard schema for European Open Data Portals DCAT-AP[13] with spatial properties of the standards ISO 19115[14]/19119[15] (Geographic information – Metadata / Services).

2. Components that manage and use the metadata to increase **data transparency** and **discoverability**:

   For metadata management and presentation, a catalogue service that includes all metadata records has been set up. Our catalogue is based on an open source solution called CKAN[16] (Comprehensive Knowledge Archive Network). It also supports the upload of data-supplementing resources like documentations or links to scientific publications. CKAN is an open data platform that is widely used by governments (e.g. Canada, Australia, European Data Portal) or city administrations (e.g. Berlin, Helsinki). It is extensible with own plugins and already implements features for geo-related data[17]. CKAN also provides an API (Application Programming Interface). This interface is capable of converting all records into a JSON representation and is therefore easy to integrate into other applications such as websites. We used this API to automate the process of metadata upload as well as connecting the catalogue with the main climate information platform ReKIS (see chap. ***3 Implemented functionality*** for a description of this connection).

3. Components that support inter-organizational **data exchange**:

   As mentioned earlier, the platform shall support a multitude of participating data providers and the resulting climate data products should be usable for stakeholders

[12] https://inspire.ec.europa.eu/good-practice/geodcat-ap
[13] https://joinup.ec.europa.eu/collection/semantic-interoperability-community-semic/solution/dcat-application-profile-data-portals-europe/release/11
[14] https://www.iso.org/standard/26020.html
[15] https://www.iso.org/standard/39890.html
[16] https://ckan.org/
[17] http://docs.ckan.org/projects/ckanext-spatial/en/latest/

from many different organizations. In order to guarantee seamless data exchange over the web and so build up an effective distributed information infrastructure, it is recommended to host data as a service via standardized web interfaces [Pe10]. These services can be directly connected to GIS applications (both web and local). Subsequently, the data users only stream the data content, whereas metadata and data sources are maintained by the infrastructure hosts [Eu18].

Hence, we set up an instance of GeoServer[18] that implements the open standards specified by OGC and published our geospatial data as web services. Precisely, our server hosts the following services:

a) **Web Map Service[19] (WMS)**: e.g. geographic base data such as administrative units, land use scenarios, water body layers, etc.

b) **Web Feature Service[20] (WFS)**: e.g. flood risk maps, hydro-geological overview maps, digital terrain models, etc.

c) **Web Coverage Service[21] (WCS)**: e.g. grid interpolated climate data (daily/yearly) for the federal states. These include variables like bias-corrected precipitation, mean/max/min temperatures, relative humidity, global radiation, etc.

The metadata of these services are stored in our catalogue service. The OGC community is currently working on new standards[22] that will supersede the old ones. As soon as available, it is planned to integrate these new standards via the WebGIS server as well.

4. Components that enable **low-threshold access** to data and scientific research results:

To derive decision-making tools for climate adaptation, climate information must reflect specific local climate change impacts that are tangible for the local policymakers [Ns21, Si18]. These can be examples of flood events, consequences of droughts or generally the change of the local water balance [Si18]. Especially when stakeholders are not climate experts themselves (like in our case), it is important to make the offered information easily accessible and understandable [Re18].

In order to create a low-threshold presentation of the researched climate information, we decided to add so-called story maps as a further component. In recent years, story maps have been widely used as a tool for knowledge transfer; mainly in the educational context [BHW18, Co18, Th19] but also as a suitable tool for science communication [Ou21, Vo21]. They are an illustrative method to present information with a spatial reference. They can contain interactive maps, text blocks, data visualizations, and other media content. In order not to set up another web server that

[18] https://geoserver.org/
[19] https://www.ogc.org/standards/wms
[20] https://www.ogc.org/standards/wfs
[21] https://www.ogc.org/standards/wcs
[22] https://ogcapi.ogc.org/

has to be maintained, we decided to use the solution from ESRI[23].

The geospatial web services hosted on our WebGIS server (see 3 *Components that support inter-organizational data exchange*) are used as a layer for web maps. In turn, these web maps can be directly embedded into the story maps. When we update the data sources of the maps, no edits in the story map are necessary, because the data content is dynamically pulled via the services. This is a further detail towards building a sustainable infrastructure.

3. Implemented functionality

The developed spatial data infrastructure includes the above-described components. The current infrastructure implements a straightforward workflow from data provision to its use. For both data providers and users, there are only two points of intervention in this process (Fig. 2). Data providers need to store their data on a FTP-based server intended for this

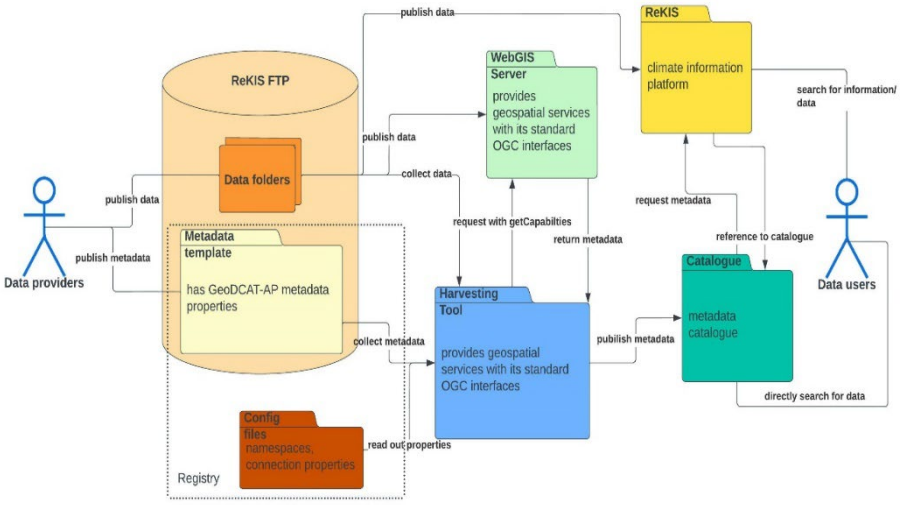


Fig. 2: Workflow from data provision to data use

purpose (see Fig. 2: ***ReKIS FTP***) and fill in a template for respective metadata (see Fig. 2: ***Metadata template***). In a next step, the data get published and available for download via a web service on the main website (see Fig. 2: ***ReKIS***). Currently the data upload data to the WebGIS Server (see Fig. 2: ***WebGIS Server***) is done manually. We are aiming to automate this step as well (c.f. chap. ***4. Conclusion and outlook***). A python-based tool

[23] https://storymaps.arcgis.com/stories

(see Fig. 2: ***Harvesting Tool***) collects metadata directly from the metadata template and from the services hosted by the WebGIS server (WMS, WFS, WCS). Comparable to ETL (Extract Transform Load) tools, the Harvesting Tool maps all the fields of the sources to the catalogue metadata schema and publishes respective metadata records. YAML based configuration files contain all mapping and connection parameters (see Fig. 2: ***Config Files***). With these configuration files, we intend to keep our tools flexible and extensible to new potential data and metadata sources. The objective is to build a data infrastructure that can be flexibly extended with new components as requirements change. Previous experiences with the climate information platform ReKIS show that such climate information platform has constantly to be adapted to the ever-growing amount of information, the technological progress as well as the evolving needs of the end users [Kr21].

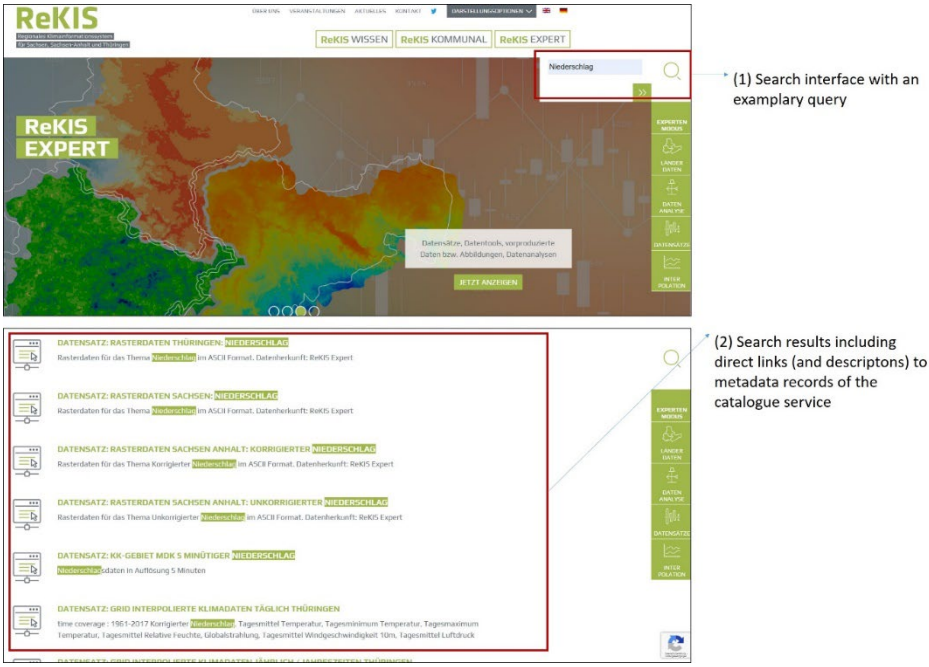


Fig. 3: Integration of the catalogue component into the search interface of the main climate information system

During the implementations, we received feedback from the ReKIS operators that platform users may not know what a metadata catalogue is. Therefore, the provision of a simple Web link to our catalogue service is not sufficient for users. We solved this by a discreet integration of the catalogue service into the search interface of ReKIS (Fig. 3). If users search for a keyword on ReKIS, a request is made to the API of the catalogue in the background. If available, the catalogue API returns metadata records and appends them to the search results page. These search results include a link and a brief description. By

clicking on the search results, our catalogue opens in a new browser tab and users can find a data description and access and/or download information. As a result, the two SDI components are integrated and inexperienced users might not even recognize that the provided functionality is actually distributed between the two.

4. Conclusion and outlook

The presented infrastructure is in operational use[24] and regular feedback from the project participants is evaluated to improve the offered functionality. The extensions to the architecture primarily bring benefits for the project participants, who use it for publishing decision-supporting climate data products.

However, there are still features that have to be realized (Fig. 4) in future. As mentioned earlier, the publication of geospatial web services on our WebGIS server is still not working automatically. Further, there is currently no interface to upload the collected metadata to our WebGIS server. The server uses a GeoServer extension[25] that allows customisable configuration of metadata fields. Therefore, it was no issue to adapt the metadata fields to the schema of the metadata catalogue. Furthermore, GeoServer provides a RESTful API[26] that enables automatic upload of data and metadata. The next step is to develop a further Python module that uses this API to

- upload datasets that are stored by the data providers (on ***ReKIS FTP***, c.f. Fig. 2)
- upload/update metadata that the data providers deposit in the metadata template

To implement this, an existing python library[27] that provides methods to connect with the GeoServer-API can be used.

In addition, some of the story maps are still in a conceptual state. In order to summarize and communicate the scientific results of the climate research, the partners have yet to complete their research and environmental modelling. As mentioned earlier, we intend to use the hosted geo services in maps that are embedded in the story maps (Fig. 4). Once the story maps are finalised with content, we will integrate them into ReKIS (similar to the catalogue service).

After all new architectural components are completed and integrated, we aim to evaluate their contribution to creating an open, transparent, and easy-to-access infrastructure. It is envisaged that dedicated user surveys will be conducted for this purpose. Since the WebGIS server, the harvesting tool and the catalogue service use open interfaces, further components can be added if necessary.

[24] URL of the catalogue service: https://klimakonform-dmp.geo.tu-dresden.de/
[25] https://docs.geoserver.org/stable/en/user/community/metadata/index.html
[26] https://docs.geoserver.org/stable/en/user/rest/
[27] https://github.com/gicait/geoserver-rest

Until the project end of KlimaKonform in April 2023, it is intended to complete the story maps and to publish them on the ReKIS plattform. Further, the functionality to automatically upload (meta-)data on the WebGIS server will be implemented. As the development of the new OGC API standards will not be completed until the end of 2023[28], these will be implemented in the WebGIS server after the end of the project.

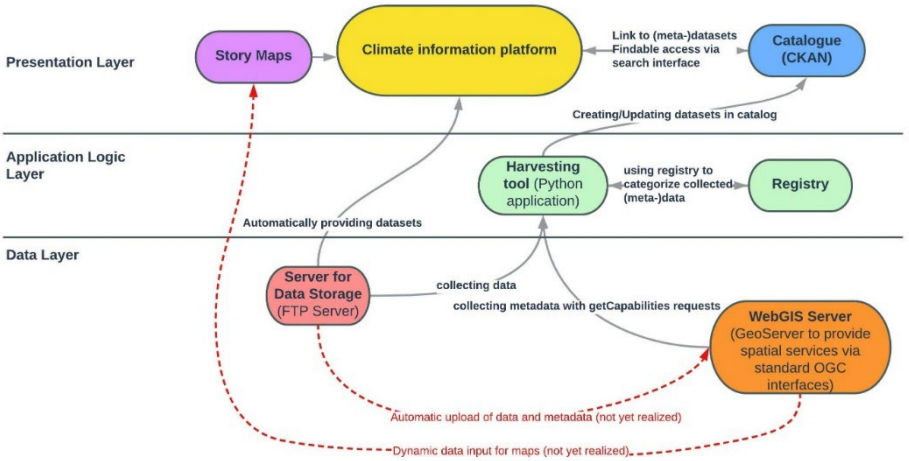


Fig. 4 Future implementation plans to be realised

Bibliography

[ATG22] Agrawal, S.; Tripathi, A. K.; Gupta, R. D.: Development and implementation of automatic metadata generation framework for SDI using OSS: a case study of Indian NSDI. Arabian Journal of Geosciences, 15/5, 2022.

[BHW18] Berendsen, M. E.; Hamerlinck, J. D.; Webster, G. R.: Digital story mapping to advance educational atlas design and enable student engagement. ISPRS International Journal of Geo-Information, 7/3, 2018.

[Be14] Bernard, L.; Mäs, S.; Müller, M.; Henzen, C.; Brauner, J.: Scientific geodata infrastructures: Challenges, approaches and directions. International Journal of Digital Earth, 7/7, pp. 613–633, 2014.

[BPS16] Bernhofer, C.; Pluntke, T.; Schaller, A.: Analyse und Bewertung des WEREX-V-Ensembles hinsichtlich der Abbildung von Extremen in Sachsen. Schriftenreihe Des LfULG, Heft 9, 2016.

[Co18] Cope, M. P.; Mikhailova, E. A.; Post, C. J.; Schlautman, M. A.; Carbajales-Dale, P.:

[28] See the development roadmap of the OGC API at https://ogcapi.ogc.org/apiroadmap.html

Developing and Evaluating an ESRI Story Map as an Educational Tool. Natural Sciences Education, 47/1, pp. 1–9, 2018.

[DEA18] Donnelly, C.; Ernst, K.; Arheimer, B.: A comparison of hydrological climate services at different scales by users and scientists. Climate Services, 11(June), pp.24-35, 2018.

[Eu18] European Environment Agency: EEA public map services, https://www.eea.europa.eu/code/gis, last modified 06/2018, last access: 2022/06/20.

[Ha18] Habermann, T: Metadata life cycles, use cases and hierarchies. Geosciences (Switzerland), 8/5, pp. 1-15, 2018.

[Ip14] IPCC, International Panel on Climate Change: Climate Change 2014 Part A: Global and Sectoral Aspects in Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, 2014, https://www.ipcc.ch/site/assets/uploads/2018/02/WGIIAR5-PartA_FINAL.pdf, last access 2022/06/20.

[Ip19]: IPCC, International Panel on Climate Change: Global warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, 2019, https://www.ipcc.ch/site/assets/uploads/sites/2/2019/06/SR15_Full_Report_High_Res.pdf, last access: 2022/06/20-

[Kr21] Kronenberg, R.; Franke, J.; Neumann, T.; Struve, S., Bernhofer, C.; Sommer, W.: Das Regionale Klimainformationssystem Re-KIS – eine gemeinsame Plattform für Sachsen, Sachsen-Anhalt und Thüringen. In Umweltinformationssysteme. Grundlagen einer angewandten Geoinformatik/Geo-IT (3. vollst). Peter Fischer-Stabel (Ed.) - Herbert Wichmann-Verlag, 2021.

[No20] Nogueras-Iso, J.; Ochoa-Ortiz, H.; Ángel Jañez, M.; R. R. Viqueira, J.; Po, L.; Trillo-Lado, R.: Automatic Publication of Open Data from OGC Services: the Use Case of TRAFAIR Project GEOProcessing 2020. The Twelfth International Conference on Advanced Geographic Information Systems, Applications, and Services, pp. 75-80, 2020.

[Ns21] Nsengiyumva, G.; Dinku, T.; Cousin, R.; Khomyakov, I.; Vadillo, A.; Faniriantsoa, R.; Grossi, A.: Transforming access to and use of climate information products derived from remote sensing and in situ observations. Remote Sensing, 13/22, 2021.

[Ou21] Oubennaceur, K.; Chokmani, K.; Alem, A. El; Gauthier, Y.: Flood risk communication using arcgis storymaps. Hydrology, 8/4, 2021.

[Pe10] Percival, G: The application of open standards to enhance the interoperability of geoscience information. International Journal of Digital Earh, 3/1, pp. 14-30, 2010.

[RVR19] Raes, L.; Vandenbroucke, D.; Reznik, T.: GeoDCAT-AP. OGC Doc. No. 18-001r1, pp. 1-132, 2019, http://www.opengis.net/doc/dp/GeoDCAT-AP, last access: 2022/06/20.

[Re18] Reiter, D.; Meyer, W.; Parrott, L.; Baker, D.; Grace, P.: Increasing the effectiveness of environmental decision support systems: lessons from climate change adaptation projects in Canada and Australia. Regional Environmental Change, 18/4, pp. 1173-1184,

2018.

[SM18] Schauppenlehner, T.; Muhar, A.: Theoretical availability versus practical accessibility: The critical role of metadata management in open data portals. Sustainability (Switzerland), 10/2, 2018.

[Si18] Singh, C.; Daron, J.; Bazaz, A.; Ziervogel, G.; Spear, D.; Krishnaswamy, J.; Zaroug, M.; Kituyi, E.: The utility of weather and climate information for adaptation decision-making: current uses and future prospects in Africa and India. Climate and Development, 10/5, pp. 389-405, 2018.

[Th19] Thürkow, D.; Lindau, A. K.; Schmidt, G.; Illiger, P.; Krause, C.; Gerstmann, H.; Schürmann, A.: Using Interactive Story Maps Enriched by Direct Knowledge Queries for the Development of E-Learning Modules on Climate Change KN. Journal of Cartography and Geographic Information, 69/3, pp. 195-204, 2019.

[Un22] UN, United Nations: Climate change | Department of Economic and Social Affairs, https://sdgs.un.org/topics/climate-change, last access 2022/20/06.

[Vo21] Vollstedt, B.; Koerth, J.; Tsakiris, M.; Nieskens, N.; Vafeidis, A. T.: Co-production of climate services: A story map for future coastal flooding for the city of Flensburg. Climate Services, 22(January), 10025, 2021.

[Wi16] Wilkinson, M. D.; Dumontier, M.; Aalbersberg, Ij. J.; Appleton, G.; Axton, M.; Baak, A.; Blomberg, N.; Boiten, J. W.; da Silva Santos, L. B.; Bourne, P. E.; Bouwman, J.; Brookes, A. J.; Clark, T.; Crosas, M.; Dillo, I.; Dumon, O., Edmunds, S.; Evelo, C. T.; Finkers, R.; Gonzalez-Beltran, A.; Gray, A. J.G.; Groth, P.; Goble, C.; Grethe, J. S.; Heringa, J.; Hoen. P. A.C. 't: Hooft, R.; Kuhn, T.; Kok, R.; Kok, J.; Lusher, S. J.; Martone, M. E.; Mons, A.; Packer, A. L.: Persson, B.; Rocca-Serra P.; Roos, M.; van Schaik, R.; Sansone, S.-A.; Schultes, E.; Sengstag, T.; Slater, T.; Strawn, G.; Swertz, M. A.; Thompson M.; van der Lei, J.; van Mulligen, E.; Velterop, J.; Waagmeester A.; Wittenburg, P.; Wolstencroft, K.; Zhao, J.; Mons, B.: Comment: The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data, 3, pp. 1-9, 2016.

[Ya19] Yamashkin, S. A.; Yamashkin, A. A.; Fedosin, S. A.: Project-Oriented Spatial Data Infrastructures. International Journal of Civil Engineering and Technology (IJCIET), 10/2, pp. 1181-1190, 2019.

[Ya08] Yan, Y.; Jia, W.; He, Z.: Research on querying for metadata service based on geo-ontology. International Geoscience and Remote Sensing Symposium (IGARSS), 2008.

Living lab research project "5G Smart Country" - Use of 5G technology in precision agriculture exemplified by site-specific fertilization

Moid Riaz Bhatti [1], Ali Akyol[1], Henrik Rosigkeit[2], Linda Matzke[2], Isabel Grabenhorst and Jorge Marx Gómez[1]

Abstract: The research project "5G Smart Country" aims at developing ideas for the development and testing of 5G applications for agriculture and forestry under real conditions. Agricultural and forestry data are collected from a wide variety of sources, such as satellites, drones, and robots with special sensors. Artificial intelligence (AI) and data analytics algorithms help make the required decisions, particularly for automatic differentiation between crops and weeds for mechanical weed control, demand-driven fertilization (variable rate application, VRA)—also by means of small-scale application (pointed fertilizing)—automated tracking of wildlife populations, real-time assessment of harvest (smart harvesting), forest inventory maintenance, and targeted logging. Here we present a system architecture and software model for digital crop management and show how multispectral analysis is used to develop vegetation indices to conduct VRA.

Keywords: Living lab, smart farming, BMDV, 5G, site-specific fertilization, VRA, satellite/sentinel, vegetation indices, NDVI, GNDVI, digital plant model, AI.

Addresses Sustainable Development Goal 13: Climate action

1. Introduction

Today's agriculture faces a variety of challenges. The projected increase in population from currently 7.9 to 9.7 billion people in 2050 and the resulting industrialization of cities will continue to reduce the amount of land per capita that can be used for agriculture [Sc21, EH15, Mu21]. The supply of food for the increasing world population is also being constrained by climate change, which is causing a deviation in the optimal temperature and water availability of crops. As a result, yield uncertainties are increasing [EH15]. The spatial heterogeneity of soil properties on agricultural land leads to varying nutrient removal and yields due to the variability of plant development. With the amendment of the Fertilizer Ordinance (Düngeverordnung), the requirements for fertilization have been

[1] Oldenburg University, Chair Business Informatics (Very Large Business Applications), Ammerländer Heerstr. 114-118, 26129 Oldenburg, Germany, moid.riaz.bhatti@uni-oldenburg.de https://orcid.org/0000-0002-3426-2960, {ali.akyol, jorge.marx.gomez}@uni-oldenburg.de
[2] Landwirtschaftskammer Niedersachsen, Fachbereich Versuchswesen Pflanze, An der Domäne 1, 38154 Königslutter, Germany, {henrik.rosigkeit, linda.matzke, isabel.grabenhorst}@lwk-niedersachsen.de

further tightened [Dü20]. How precision farming with site-specific application/ management (VRA) can be made more efficient with 5G technology is one of the questions addressed by the research project "5G Smart Country" funded by the German Federal Ministry of Digital Affairs and Transport (BMDV). In the following work-in-progress paper, an overview of the collaborative project is given based on site-specific Nitrogen (N) fertilization (N-VRA) in winter wheat (*Triticum aestivum* L.). A system architecture for the collection of field data, networking of various machines, and data visualization as well as a software model for digital crop cultivation are presented and the possible scenarios for 5G use in precision farming are explained.

2. Project presentation and work packages

As part of the 5G innovation competition (5G-Innovationswettbewerb), the BMDV is funding cities, regions, and research organizations to develop and test ideas for 5G applications [Bm21]. With prospects of digital transformation of rural areas, the districts of Helmstedt and Wolfenbüttel in Lower Saxony successfully applied to the BMDV for the project "5G Smart Country[3]" [Bm21]. The focus of the project is on testing 5G applications under real agricultural and forestry conditions. The sub-project and the subject of this paper “Smart Farming” has several project partners[4] from science and industry, and six work packages (WP[5]) [Wo21].

3. Fundamentals of mineral fertilization of winter wheat

Breeding progress, intensification of cultivation strategies, and the use of mineral fertilizers have historically led to an increase in wheat yields. In Germany, the cultivation of winter wheat is particularly widespread [Kr12]. Besides yield increase, N fertilization can also cause leaching of the unused amounts of N with water [Bu15]. With an oversupply of N in the soil due to fertilization and mineralization of the residues of soil organic matter resulting in plant-available forms of ammonium (NH_4^+) and nitrate (NO_3^-), increased losses occur leading to nitrate leaching and ammonia immobilization [Ha14].

To guarantee a fertilization that meets the needs as much as possible, the N available to plants in the soil, N_{min}, is determined as a minimum factor by sampling the sites at the beginning of the vegetation period, and the fertilization strategy is planned in relation to

[3] The estimated project duration is 3 years (Nov. 2021 – Nov. 2024); funding amount: 3.9 mill. euros.

[4] Landwirtschaftskammer Niedersachsen, Universität Oldenburg, TU Braunschweig, Julius-Kühn-Institut, Deutsches Zentrum für Luft- und Raumfahrt, Strube D&S GmbH, Rauch Landmaschinenfabrik GmbH, Domäne Schickelsheim, Vodafone GmbH.

[5] WP1: Digitized crop production systems; WP2: Trait recording on individual plants; WP3: Detecting, locating, and regulating weeds; WP4: Recording wildlife populations in agricultural areas; WP5: Pointed fertilizing; WP6: Smart grain harvesting and information gathering.

the crop development, e.g., by means of BBCH[6] scale [Fi07]. The first application in spring at the time of plant development at BBCH stages 21-25 has a direct influence on stand density at the time of tillering. At the onset of heading (BBCH 30-32), yield plants are differentiated by reducing weak secondary shoots and fixing ear density [Kr12]. Furthermore, N fertilization results in the formation of additional roots on the plant, which promote anchorage in the soil and nutrient uptake on the main shoot [St94]. Optionally, a third application before ear emergence in BBCH-49 is also considered which can be decisive for yield and quality due to its influence on the storage cells in the grain.

3.1 Site-specific fertilization of winter wheat

Increasing N supply to a crop generally leads to a greater crop biomass and hence grain yield, but it can also make the plants susceptible to lodging, which limits yields and quality. Though harvesting of crops with insufficient fertilization increases N use efficiency (NUE)—yield per unit of N available to the crop—it also leads to depletion of soil mineral reserves and deterioration of soil quality [Ha14]. N-VRA considers specific soil conditions of the smaller sections and plant nutritional status resulting in a higher NUE and a lower risk of N leakage into the environment [Me22, Mi20]. An efficient N-VRA in the cultivation of winter wheat brings economic (yield-related) and ecological (regarding substance leaching and emission) advantages [Ra02]. In this context, the heterogeneity of the field is of particular importance. To increase yields, sites with poorer plant development are promoted, especially in the first two fertilizer applications, and those with above-average plant development are fertilized less. This leads to a more homogeneous development of the crop. To maintain the quality of the varieties, high-yielding zones are specifically supplied with sufficient N in the third application, while at the same time sites with a lower crop development are only supplied with a basic amount.

4. Technical concept

The project foresees primarily 5G technology for fast and reliable communication between machines (process control), sensors (data collection), servers (AI/data analytics), and users. The expected data heterogeneity will be served by a data lake.

4.1 System architecture

The project "5G Smart Country" envisages collection of a wide variety of data on winter wheat and sugar beet (*Beta vulgaris* L.). The data will be processed using AI/data analytics and will be displayed as live maps, for instance, for biomass, nutrient quantities, wildlife, stubble, and weed. A possible system architecture is shown as frontend-backend layers [Mo22, FG16] in Fig. 1. The frontend serves as the display layer allowing for

[6] Biologische Bundesanstalt für Land- und Forstwirtschaft, Bundessortenamt und Chemische Industrie (BBCH)

userinteraction with the system, and the backend is responsible for data storage and processing.

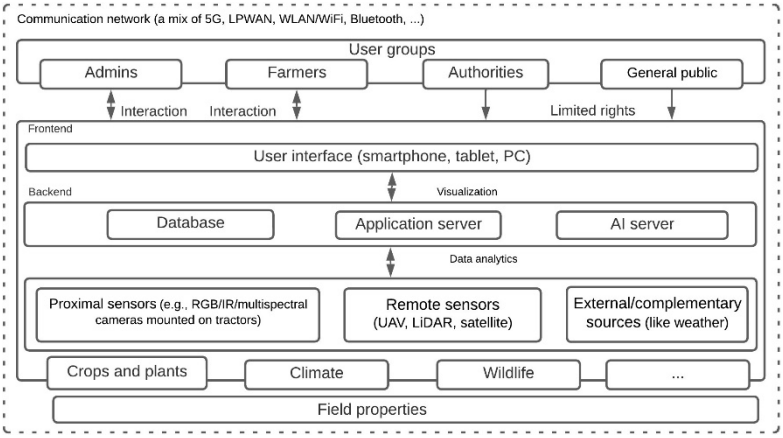


Fig. 1: Possible system architecture [own illustration].

The portal is being designed to be able to manage different individual user groups (farmers, scientists, authorities, public) with a multi-device capable frontend making it operable via different end-devices like smartphone, tablet, and PC. The backend layer collects data from different data sources (from simple proximal sensors to satellites), integrates the measurement data with complementary data (e.g., weather), processes the data set using data analytics/AI algorithms, and provides results such as N availability, biomass, stubble, and wildlife tracking maps.

4.2 Data communication

With higher data transmission speed and connection density, and especially lower latency, 5G technology can create new opportunities for faster and more precise decisions and process control, particularly based on machine-to-machine (M2M) communication [Ha22]. However, data volumes and rates for Internet-of-Things (IoT) sensors in smart farming will not always and everywhere need the highest 5G capabilities, and for such use cases other 'lightweight' technologies would be more optimal. IoT is increasingly using the Low-Power Wide-Area Networks (LPWAN) technology, where we have lower data rates (in the kbit range) than Wireless Local Area Network (WLAN/WiFi), Bluetooth, and cellular technology, but much longer range (10–40 km in rural and 1–5 km in urban areas) and much longer battery life (10+ years) due to low power consumption. An established LPWAN protocol is the Long-Range Wide-Area Network (LoRaWAN), which is built on the Long-Range Modulation (LoRa) technique—a chirped spread spectrum-based communication technique—making the signal transmission robust to channel interference and is, therefore, suitable for long-distance energy-efficient transmissions [CZ20]. The

whole communication architecture will be a mix of diverse data transmission rates and corresponding techniques. Some use cases where higher 5G performance would be needed are real-time communication between the agricultural machines (VRA), UAVs (crop analytics), robots (weed detection/elimination, pointed fertilizing, and detection of biotic (such as pest infestation) and abiotic (such as heat) stresses [HH09]) and the servers. Some of the data exchange standards/formats deemed suitable for the communication of geographic data are GeoJSON (encoding and exchange of geospatial data), Shapefiles (distribution of vector/map data), GeoTIFF (raster image file format for satellite and aerial imagery data), LAS (3D LiDAR point cloud data), WMS (display of geospatial data on the Internet), and REST (communication between machines).

4.3 Collection and visualization of field and plant data

The system has different data sources for precision farming with their own requirements for data transmission and algorithms: 1. proximal/near-ground sensors (e.g., RGB/IR cameras mounted on harvesters/tractors); 2. remote sensors (e.g., UAV, LiDAR, satellite) [Ta22]; 3. external sources (e.g., weather). An established technique for soil/land reconnaissance and crop analysis is multispectral analysis, where the reflected light from the object of interest is decomposed into the individual wavelength ranges and analyzed [Cu80]. Multispectral analysis grants conclusions about soil properties and plant health as well as other properties of the crop from the examined areas [Se20]. Based on multispectral analysis, a field can be subdivided into subareas (sections) suitable for VRA.

To carry out VRA, a preliminary analysis of several factors is conducted, e.g., plant condition and soil nutrient content. Such an analysis can hardly be done manually on even moderately large farms. For this purpose, multispectral sensors mounted on satellites, UAV, or ground-based equipment are generally used. From multispectral data, various vegetation indices (VI) can be calculated by using ratios, differences, ratios of differences and sums, or by forming linear combinations of spectral band data. VIs enhance the vegetation signal, while minimizing the effect of solar irradiance and soil background [JH91], and can be integrated into GIS maps for site-specific analysis, processing, and yield forecasting [Di16, PG21]. One prominent example of VIs is the Normalized Difference Vegetation Index (NDVI) which uses the fact that a healthy plant absorbs most of the incident visible wavelengths (particularly the red (R) ones) while transmitting or reflecting most of the falling near-infrared (NIR) wavelengths. NDVI is calculated as the normalized difference of the spectral reflectance in the R and NIR regions (NDVI=(NIR–R)/(NIR+R)) [Ro73, JH91]. NDVI values for vegetative areas vary between 0 and 1. They increase towards 1 with an increase in the soil cover by the growing plants, their Leaf Area Index (LAI), chlorophyll content, and plant N-nutritional status [VSB19].

An example of NDVI maps is shown in Fig. 2, where Sentinel-2 satellite images [SDH] of a winter wheat field in March and April were used to calculate NDVI on a 10 m grid (QGIS, version 3.16.14). The field can thus be divided into small sections that can be viewed and managed individually. With sufficiently high-resolution multispectral data, it

is even possible to view the individual plants in row crops. The storage and processing of the site-specific data and the required fertilization process control decision mechanisms are part of the digital plant cultivation model which is described in the following section.

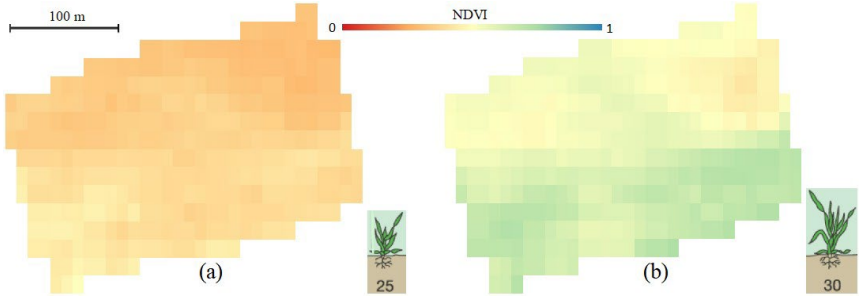


Fig. 2: NDVI maps of a winter wheat field in different plant growth stages (Sentinel-2 L2A data with 0% cloud cover; NDVI scale: 0: hardly any; 1: good vegetation). (a) BBCH-25: tillering, (b) BBCH-30: just before the second fertilizer application with the beginning of the stem and the first internode elongation. Cartoons: the plants in their corresponding phenological stages [Br15].

4.4 Software model for digital plant cultivation

The digital plant cultivation system is based on the object-oriented software design methodology (OO) [BJR07] in the form of classes. Fig. 3 shows a Unified Modeling Language (UML) class diagram [Fo04]. The main classes are Environment, Soil, Seed, Fertilizer (field/site-specific, i.e., uniform/VRA), Plant, and Wheat. The classes contain agricultural parameters as fields and methods for data retrieval, processing, and process control, respectively. The basic idea behind the software model is explained below.

The Environment class hosts environmentally relevant information such as temperature and relative humidity. Current as well as historical environmental data serve as a basis for agricultural planning and process control. The Soil class models the physical and chemical properties of the soil where crops are planted. The soil type is derived from the composition of sand, clay, and silt [US99]. Other important parameters include pH, soil salinity, and nutrients in the soil. Seed quality significantly affects plant stand, crop quantity and quality, and is modeled by the Seed class, which collects data on seed quality (seed purity) and seeding parameters (seeding depth, rate…) that enable plant growth and yield prediction. The fact that different phenological stages in plant growth have their own needs for nutrients is considered by the Plant class which models plant phenology with data on initial parameters (plant type/position, soil salinity), vegetation parameters (vegetation stage (VS), transpiration rate), morphological parameters (LAI), and vegetation indices (VI), e.g., NDVI or the GNDVI [Gi96] which is computed similarly to the NDVI, but the green (G) band is used instead of the R band (GNDVI=(NIR–G)/(NIR+G)). GNDVI is more sensitive to chlorophyll concentration in a wide range of chlorophyll variations than the 'red' NDVI allowing more precise estimation of pigment

concentration and is linearly correlated with LAI and biomass [Gi96, Hu08]. Using this set of data, the Plant class provides an objective basis for yield prediction. The characteristics of the produce depend on the plant type and the Wheat class models this with data on physical properties (hectoliter weight, foreign matter, falling number...) [St17]. Fertilizer application is modeled by the Fertilization class. VRA is controlled according to the VSs and VIs which are determined from the sensor data and allow yield forecasting which is integrated in the Wheat class [HH09, JC19, Le86].

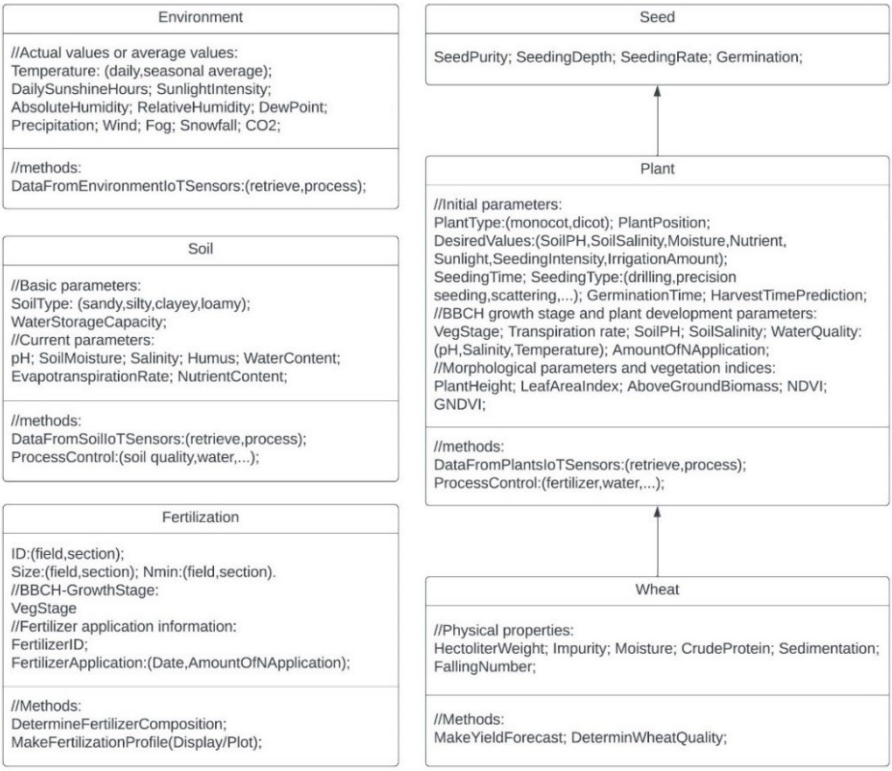


Fig. 3: UML class diagram for the digital plant cultivation system. The arrows show OO inheritance relationship between the classes [Fo04], e.g., wheat plant is a special plant and therefore the class wheat inherits data and methods from the class plant and extends this with its own specific features.

Based on this georeferenced data, AI and statistics are used to gain insights and control processes such as trajectory generation for UAV flights for VRA and waypoint map for robot control for mechanical weed elimination or microdosing. Phenotypic measurement data are documented and processed in a standardized manner according to the Crop Ontology Model (COM) [Pi22]. Here, a variable is defined as a combination of three quantities: Trait (what is being studied: e.g., plant height), Method (how is the

measurement made: e.g., measurement, calculation), and Scale (how is the measurement represented: qualitative or quantitative (unit)). In summary, 1 variable = {1 trait, 1 method, 1 scale}. A variable name is constructed like this: Trait_Method_Unit. Tab. 1 illustrates this with an example data set from the COM:

Variable name	Trait name	Trait abb.	Method name	Method explanation	Unit
PH_M_cm	Plant height	PH	PH measurement	The distance from the ground to the top of the plant measured with a ruler	cm
GY_Calc_gm2	Dehulled grain yield	GY	GY computation	Measured on harvested and dehulled grain. Divide weight by plot area.	g/m²

Tab.1: A data set from the COM.

5. Conclusion

In this paper we have presented the BMDV project "5G Smart Country". The objective is to develop and implement new concepts for 5G-based precision farming, particularly plant analytics and intelligent process control aimed at increasing agricultural efficiency by increasing yields while optimizing resource use and complying with fertilizer, pesticide, and climate-oriented environmental regulations. The planned key milestones include 5G-based fertilization (VRA, pointed), machine/software-based detection of crop stresses, automated initiation of actions, e.g., robotic weed detection and elimination, and wildlife tracking. In the end, a web-based software system/app will be available to farmers, scientists, authorities, and the public, providing information vital for smart farming (weather, field heterogeneity, plant growth, N demand/uptake, weeds, biomass, yield forecast, stubble...) in the form of maps, tables, and files.

6. Outlook

The software model for digital crop cultivation will be further developed to include additional components in the farming processes as well as communication with sensors and machines and implemented in a programming language. Furthermore, a data visualization app with weather, wildlife, stubble and weed maps as well as a satellite data processing pipeline for the study of crop development and determination of VIs for VRA and yield forecasting will be developed.

Acknowledgement

The authors thank the BMDV for funding this research project (grant number: 45FGU117)

Bibliography

[Bm21] 5G-Innovationsprogram.www.bmvi.de/DE/Themen/Digitales/Mobilfunk/5G-Innovations programm/5g-innovationsprogramm.html, last access: 20.05.2022.

[BJR07] Booch, G.; Jacobson, I.; Rumbaugh, J.: Object-Oriented Analysis and Design with Applications. 3rd ed., Pearson Education, MA, 2007.

[Br15] Brockerhoff, H.: Entwicklungsstadien im Getreide, LWK-NRW, 2015.

[Bu15] Bull, I. et al.: Qualitätsdüngung zu Winterweizen: in Mitteilungen der Landes-forschungsanstalt für Landw. und Fischerei: Heft 55, 2015.

[Cu80] Curran, P.: Multispectral remote sensing of vegetation amount. Progress in Physical Geography 4(3), pp. 315-341, 1980.

[CZ20] Chaudhari, B.S.; Zennaro, M.: LPWAN Technologies for IoT and M2M Applications. Academic Press, CA, 2020.

[Di16] Diepenbrock, W.; Ellmer, F.; Leon, J.: Ackerbau, Pflanzenbau und Pflanzenzüchtung. Verlag Eugen Ulmer, 4th ed., 2016.

[Dü20] Bundesanstalt für Landwirtschaft und Ernährung. Düngeverordnung (DüV), 2020.

[EH15] Ehrlich, P.; Harte, J.: Opinion: To feed the world in 2050 will require a global revolution. Proc. Natl. Acad. Sci. USA, 112(48), pp. 14743–14744, 2015.

[Fi07] Finck, A.: Pflanzenernährung und Düngung in Stichworten, 6th ed., Borntraeger, 2007.

[FG16] Foster, E.C.; Godbole, S.: Database Systems: A Pragmatic Approach. 2nd ed., Apress, Berkeley, CA, 2016.

[Fo04] Fowler, M.: UML Distilled: A Brief Guide to the Standard Object Modeling Language. 3rd ed., Pearson Education, MA, 2004.

[Gi96] Gitelson, A.A. et al.: Use of a green channel in remote sensing of global vegetation from EOS-MODIS. Remote Sensing of Environment, 58(3), pp. 289-298, 1996.

[Ha14] Hawkesford, M.: Reducing the reliance on nitrogen fertilizer for wheat production, Journal of Cereal Science, 59(3), pp. 276–283, 2014.

[Ha22] Hassan, M.: Wireless and Mobile Networking. CRC Press, FL, 2022.

[HH09] Hopkins, W.G.; Hüner, N.P.A.: Introduction to Plant Physiology. 4th ed., John Wiley & Sons, Inc, NJ, 2009.

[Hu08] Hunt, E.R. et al.: Acquisition of NIR-green-blue digital photographs from unmanned aircraft for crop monitoring. Remote Sensing, 2(1), pp. 290-305, 2010.

[JC19] Jose, S.; Clennet, C.: Trees, Leaves, Flowers & Seeds. A Visual Encyclopedia of the

Plant Kingdom. DK Publishing, NY, 2019.

[JH91] Jackson, R.D.; Heute, A.R.: Interpreting vegetation indices. Preventive Veterinary Medicine, 11(3-4), pp. 185-200, 1991.

[Kr12] Kropf, U.: Getreide-Guide, Top agrar Ratgeber, 2012.

[Le86] Leonard, D.: Soils, Crops, and Fertilizer Use: A Field Manual for Development Workers. 4th ed., Peace Corps, 1986.

[Me22] Mezera, J. et al.: Comparison of Proximal and Remote Sensing for the Diagnosis of Crop Status in Site-Specific Crop Management. Sensors, 22(1), p. 19, 2022.

[Mi20] Mittermayer, M. et al.: Erfassung der räumlichen Variabilität von Boden- & Pflanzenparametern: Grdl. für die teilflächenspezif. N-Bilanzierung. LNI, GI, Bonn, 181, 2020.

[Mo22] MongoDB: What is the MEAN Stack? Introduction & Examples, www.mongodb.com/mean-stack, last access: 05.05.2022.

[Mu21] Mugiyo, H. et al.: Evaluation of Land Suitability Methods with Reference to Neglected and Underutilized Crop Species: A Scoping Review. Land, 10(2), p. 125, 2021.

[PG21] Panek, E.; Gozdowski, D.: Relationship between MODIS Derived NDVI and Yield of Cereals for Selected European Countries. Agronomy, 11(2), p. 340, 2021.

[Pi22] Pietragalla J. et al.: Guidelines for creating crop-specific ontologies to annotate phenotypic data, Alliance Bioversity International-CIAT, version 2.1, 2022.

[Ra02] Raun, W. et al.: Improving N Use Efficiency in Cereal Grain Production with Optical Sensing and VRA. Agronomy Journal, 94(4), 815, 2002.

[Ro73] Rouse Jr, J. W. et al.: Monitoring vegetation systems in the Great Plains with ERTS. 3rd Earth Resources Techn. Satellite-1 Symp., Washington, D.C., 351, pp. 309-317, 1973.

[Sc21] Scholik, N.: Limit. Ein Essay zur Lage der Weltbevölkerung. Teil 1: Grenzen und Herausforderungen. Z. Außen Sicherheitspolit., 14, pp. 189–212, 2021.

[SDH] Scientific Data Hub, ESA, https://scihub.copernicus.eu/dhus/#/home.

[Se20] Segarra, J. et al.: Remote Sensing for Precision Agriculture: Sentinel-2 Improved Features & Applications. Agronomy 10, p. 641, 2020.

[St17] Steber, C.M.: Avoiding problems in wheat with low Falling Numbers. Crops & Soils 50.2, pp. 22-25, 2017.

[Ta22] Tao, H. et al.: Proximal and remote sensing in plant phenomics: 20 years of progress, challenges and perspectives. Plant Communications, 3, p. 100344, 2022.

[US99] U.S. Department of Agriculture, Natural Resources Conservation Service.: Soil Taxonomy: A Basic System of Soil Classification for Making and Interpreting Soil Surveys. Handbook 43, 2nd ed., 1999.

[VSB19] Vizzari, M.; Santaga, F.; Benincasa, P.: Sentinel 2-based N VRT fertilization in wheat: Comparison btw traditional & simple precision practices. Agronomy, 9(6), p. 278, 2019.

[Wo21] Lk Wolfenbüttel: Forschungsantrag "5G Smart Country", 2021.

Volker Wohlgemuth et al. (Hrsg.): EnviroInfo 2022,
Lecture Notes in Informatics (LNI), Gesellschaft für Informatik, Bonn 2022 111

An Artificial Intelligence of Things based Method for Early Detection of Bark Beetle Infested Trees

Peter Knebel[1], Christian Appold[1], Achim Guldner[1], Marius Horbach[1], Yasmin Juncker[1], Rüdiger Machhamer[1], Simon Müller[1] and Alfons Matheis[1]

Abstract: Bark beetles, like the European Spruce Bark Beetle (Ips typographus), are inherent parts of a forest ecosystem. However, with favorable conditions, they can multiply quickly and infest vast amounts of trees and cause their extinction. Therefore, it is important for forest officials and rangers of e. g. a national park, to monitor the population of the beetles and the infested trees. There are several ways to approach this, but they are often costly and time-consuming. Therefore, we design and test a bark beetle early warning system with AI-based data analysis: Audio data, data on pheromones and information for a drought stress assessment of the affected trees are to be collected and used as a basis for the analysis. The aim is to devise a micro-controller-based sensor system that detects the infestation of a tree as early as possible and warns the forest officials, e. g. via a message on their cell phone.

Keywords: Soundscape Ecology; Bark beetle detection; IoT sensors; AIoT-based evaluation

Addresses Sustainable Development Goal 15: Life on land

1. Introduction

Soundscape ecology [Pi11] is an interdisciplinary, systemic science approach that connects, among others, biology, landscape ecology, conservation research, eco-acoustics, bio- acoustics, and computer science. The origins of this research concept date back to the 1970s. [Sc93] Bernie Krause [Kr13] categorized sounds in terms of their origin into geophones (sounds of the natural environment, e. g. wind and water noise), biophones (sounds produced by non-human creatures) and anthrophones (sounds produced by humans, e. g. speech, machine or traffic noise). The composition of such noises form a characteristic soundscape of a locality or region. A soundscape can be systematically recorded and analyzed similar to a fingerprint. The analysis enables conclusions to be drawn about the local ecological situation, e. g. with regards to the biodiversity rate. [FT14] Soundscapes are not static, but are in a continuous process of change. The dynamics are shaped by human behaviors of the respective land use. Thus, it is possible to gain insights to regionally specific changes in terms of land use, but also with regards to climate change and its consequences on the basis of long-term observations of the acoustic

[1] Trier University of Applied Sciences, Umwelt-Campus Birkenfeld, PO-Box 1380, 55761 Birkenfeld, Germany
{p.knebel | chpp3226 | a.guldner | s12a0e | y.juncker | r.machhamer | s16c79 | a.matheis}@umwelt-campus.de

environment. [Fa14]

The starting point of our research approach is that bark beetles, in particular the European Spruce Bark Beetle (*Ips typographus*), can be detected quickly as an invasive species with the help of soundscape ecology tools. Throughout their life cycle, these beetles destroy and feed on the phloem, a layer inside the bark of trees, especially regular spruces (*Picea abies*). [Sc17] This extends from adult male individuals, which burrow through the outer layers of the bark into the phloem and bore tunnels for reproduction, over female individuals, which create further tunnels in a fanned-out structure for oviposition, to larvae and pupae, which feed on the phloem when they grow. The principle of the audio-based detection is to record these activities in-situ and attribute them to the beetle. [VH15]

However, we do not only rely on audio data as indicator, but also transfer the idea of soundscapes to pheromones. Bark beetles use pheromones to communicate, making them essential for their reproduction. Species-specific pheromones can be divided into releaser and primer pheromones and are composed chemical components. Volitile aggregation pheromones that attract the beetles to a tree spread rapidly, even without strong air movement. The beetle's sensory hairs contain up to 80 receptors for the pheromones, allowing them to be perceived even in very low concentrations. Male beetles begin searching for suitable breeding sites after winter dormancy or as young beetles after maturation feeding as so-called pioneer beetles. The onset of swarming depends on weather conditions, for example, Ips typographus swarms in dry weather and temperatures above 16.5 °C. [Ga19] After the pioneer beetles found a tree, a mass infestation is triggered by the aggregation pheromones. The beetles produce these pheromones in their digestive tract, from Alpha-Pinen and Myrcene. The main pheromone is 2-methyl-3-buten-2-ol, supplemented by the long-range attractants (S)-cis-verbenol and ipsdienol, among others. [Zu94]
Thus, we cover two approaches: 1. With the help of audio monitoring, we detect feeding sounds of the larvae or the sound of beetles boring into the phloem. 2. With the help of an artificial nose (volatile organic compounds (VOC) sensor), we detect pheromones and resin odors of the trees, which are typical for a bark beetle infestation.

2. Problem statement

When bark beetles infect a tree, it is quickly no longer economically viable to use it for lumber production. In most cases, the only option is to dispose of it, resulting in reduced revenues up to total loss. [De18] If the infestation is discovered too late, the rest of the population in the surrounding area is also vulnerable. Therefore, especially on the part of the forestry, the early detection of infestations or of the danger of an infestation is highly relevant. The changes in temperature due to climate change allow bark beetles to be active earlier in the year and thus they can even develop more generations than before. [WS98, TP08]
Currently, there are various ways to detect a bark beetle infestation. A ranger can suspect

an infestation based on borings on the ground or strong resin odors in the air, and confirm it after removing tree bark and looking for burrows or nuptial chambers. [Hö16] However, this requires an experienced ranger. In addition, rain or wind may obliterate the borings or suppress the odor. Areas that are difficult to reach may not be checked regularly enough, if at all. Many forest areas do not have a forester or ranger in charge. There is also already some work done on the (semi-)automated detection of bark beetles, such as the assessment of airplane- or drone images [He15, DD00], but they are costly and often quite inaccurate. At present, probably the most accurate method are trained dogs, which can unerringly "sniff out" a bark beetle infestation. [JBS19] However, the dogs need frequent breaks after 20 to 30 minutes and the number of trained dogs also does not suffice to ensure a comprehensive examination. Researchers at the Rottenburg University of Applied Sciences and the Universities of Göttingen and Freiburg are working on bark beetle detection from the air with the help of drones. [He20] With the help of a long nozzle below the drone, they also want to detect the resin odors caused by a bark beetle infection. Here, the flying skills of the drone pilot are a decisive factor. They must approach the forest canopy closely from above in order to be able to measure sufficient data. This method is well suited to assess forest sites that are difficult to reach. However, it is also quite personnel-intensive. In addition, strong winds and rain can make the flights difficult or impossible.

3. Proposed Method

To approach the automated detection of a bark beetle infestation in a cost- and energy efficient way, we focus our work on the design and evaluation of an Artificial intelligence of things (AIoT) based micro-controller sensor system. We use existing, open source IoT devices such as Espressif Systems on a Chip (SoC)[2] and Raspberry Pis as well as cheap sensor modules, like BME680[3] and piezoelectric contact pickups.

Currently, the data is analyzed with a TensorFlow model on a PC, but the goal is to transfer the trained model to tinyML, so that the sensor data can be classified directly on the IoT device and only the classification results from several sensor nodes are transferred to a Raspberry Pi-based central processing unit via a WiFi or LoRaWAN network. It would then relay the warnings to the forest owner. That way, we hope to reduce the energy consumption of the system, so that it can be deployed in remote areas without electrical connection.

Data for training the classification algorithms are taken from various sources. On the one hand, we obtained data recorded with experimental setups on deadwood in the lab and in the forest, on the other hand, data was provided by associated researchers, who investigate the bark beetle with the help of audio data [Du07], e. g. to observe its life cycle.

[2] https://www.espressif.com/en/products/socs [2022-04-25]
[3] https://www.bosch-sensortec.com/products/environmental-sensors/gas-sensors/bme680/ [2022-04-25]

3.1 Audio-based detection

To detect the presence of European Spruce Bark Beetles in the trees bark, we designed a prototype to pick up structure borne sound from inside the phloem on the basis of a vibration transducer as described in [Du07]. For the main component, we use a piezoelectric element in form of a round disc, normally intended for DIX-electronics as an acoustic guitar-pickup[4], which we encase in a 3D-printed housing to shield it from environmental influences. We then screw a flat-head, woodworking screw into the bark of the tree that is to be monitored.

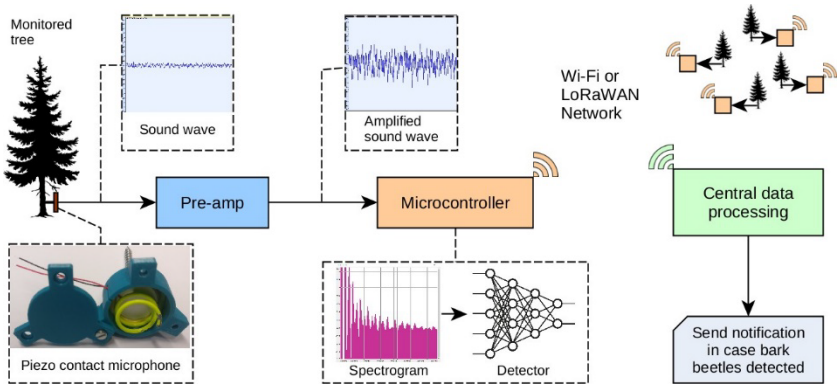


Fig. 1: Audio-based detection

In the housing, we incorporate a slot for the screw and an opposing spring, so that the piezoelectric pickup is firmly pressed against the screw. The sounds from inside the tree's bark are transmitted over the spring to the pickup which converts them into an electric signal, functioning as transducer. We fitted a standard 6.35mm stereo jack to the two wires that are soldered to the functional materials of the piezoelectric element. The signal is then passed through an amplifier and, at present, recorded with a professional handheld recorder as waveform audio format. In the envisioned final setup the amplification and audio jack will be connected directly to the microprocessor as shown in Fig. 1.

3.2 Pheromone-based detection

To detect the pheromones, we use an array of BME680 VOC sensors connected to a SoC to generate the data. The sensor generates relative resistance values as a function of ambient air VOC, which we use to explore whether the presence of bark beetles can be detected. We evaluate the data using algorithms from [Dz19], where we used a similar setup with one VOC-sensor to classify liquids using a support vector machine. The results show that this method is feasible, so we plan to use it for the classification of bark beetles

[4] https://robu.in/interfacing-of-piezoelectric-sensor-with-arduino/[2022-05-02]

as well. The sensors are attached to infested and not infested trees in similar environments. In the learning phase, we first perform extensive measurements of the ambient air. A small, heatable plate inside the sensor generates a temperature-dependent graph (similar to Fig. 5), whose characteristics allow detecting the concentration of many VOCs that evaporate at different temperatures. This data is analyzed, for example, with principal component analysis (PCA) and classification algorithms to estimate the causes of this air composition. Since the heating plate of the sensor is very small, the temperatures are reached within milliseconds and the required energy per measurement is very low. More concrete energy consumption measurements are planned as soon as the functionality is confirmed. Fig. 2 depicts the setup, again with several sensor nodes reporting to one central unit.

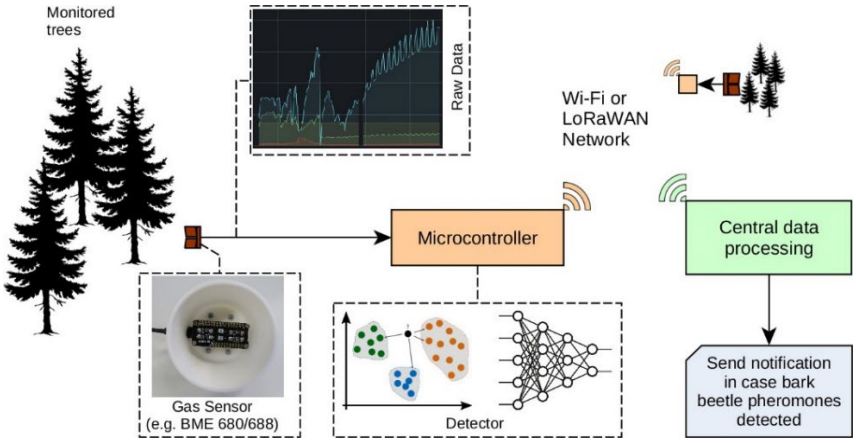


Fig. 2: Pheromone-based detection

4. Preliminary Results

As a proof-of-concept, we designed evaluation approaches for the sound-based and pheromone-based detection, which we describe here.

4.1 Sound-based classification

As a first approach for training the sound-based model, we used data collected through recordings in the Hunsrück-Hochwald National Park. We separated the sounds in two classes ”bark beetle“ and ”no bark beetle“. The second class consists of other noises, like animal sounds. In total, 47 sound files of different length ($< 20s$) were used. Since the training set is small and as such bears the risk of overfitting, we only split the data into a training and a test set. We plan on acquiring more beetle data during the summer of 2022 to extend the training process.

We trained two Keras[5]-models for classification, also making use of the python library librosa[6] for pre-processing. One model was trained with spectral features of the data (numerical data), the other with spectrograms (image data).

As training base for model 1, we used Mel Frequency Cepstral Coefficients (MFCC) which find usage in machine learning primarily for speech recognition (e. g. [Ko15]), but have also recently been used in animal sound classification [TH21]. We trained model 1 on the values of the first 13 coefficients. For the calculation, we split each sound file into snippets of 0.33 second length, from which we calculated the MFCCs. To further enlarge the small data set, we augmented the data from each snippet: Based on the positive results in [Ko15] for speech recognition, we changed the speed of the files arbitrarily between 90 and 110 %. We trained model 2 on mel-spectrograms, which are a common tool in audio classification and have also been used in animal sound recognition before (e. g. [Si20, NMP19]). Since our data already contained a lot of noise (e. g. wind), we did not use any other type of augmentation, like adding random noise, yet.

Analogous to model 1, we split the sounds into 1 second snippets and for each snippet derived the mel-spectrogram with librosa. We plotted the spectrogram without axes and saved them as image files. Spectrograms don't allow for a lot of possibilities when it comes to augmentation, since e. g. rotating or flipping them would alter the data too much. As mentioned in [NMP19], the best options are masks in the frequency and time domain. To increase the data, the spectrograms were augmented on these domains and saved respectively.

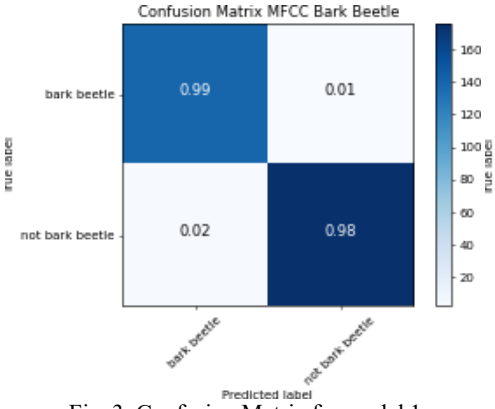


Fig. 3: Confusion Matrix for model 1

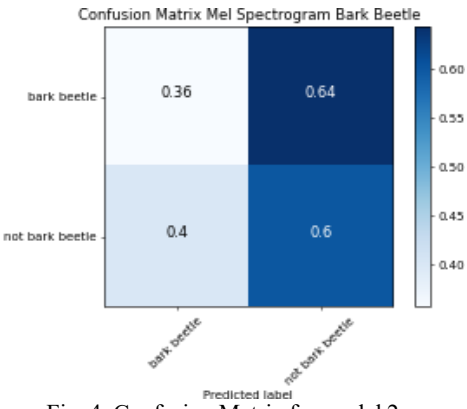


Fig. 4: Confusion Matrix for model 2

We trained both models for 110 epochs, using the sigmoid activation function and the adam optimizer. Not only are the models to be compared in their accuracy, but also in their energy efficiency, since we potentially want to deploy the detector on battery- or solar-powered devices. Considering the confusion matrices (Fig. 3 and 4), model 1 outperforms model 2,

[5] https://keras.io/[2022-04-01]
[6] https://librosa.org/[2022-05-02]

which presents a considerably higher false negative/positive rate (40 − 64% compared to < 0.5 %).

The required storage space for the pre-processed training data needed is smaller for model 1 (around 3 mb) than model 2 (around 6 mb), which will be more significant with bigger data sets. Furthermore, model 1 itself needs considerably less storage space (356 kb) than model 2 (150 mb). We also noticed differences in power consumption. Model 2 consumes over three times the energy of model 1 (1.82 vs. 0.59 watthours) in pre-processing and training. model 2 also needed more time than model 1.

4.2 Pheromone-based classification

So far, we base our assumption that the pheromones can be detected upon measurement results (see Fig. 5) obtained in a lab-setup with the pheromones mentioned in section 1. To produce reliable results for the pheromone-based detection, we are currently in the field to record more data in this years' swarming season using a setup with the VOC-Sensor as described above. To conduct the measurements, we heat the sensor to 30 °C and increase the temperature by one degree every 0.1 seconds until it has reached 400 °C. After each of these 371 steps, the sensor measures the relative resistance of its gas sensitive layer. Since such a measurement takes about two minutes, we use PCA, which provides us with the three most valuable features out of these 371, to be able to perform a much more energy-efficient and fast (about three seconds) measurement after the learning phase. Once this methodology proves successful for this use case, we plan to use the algorithms from [Ma20] to simplify the learning process in the presence of frequently changing forest air conditions or species specificity through an adaptive application.

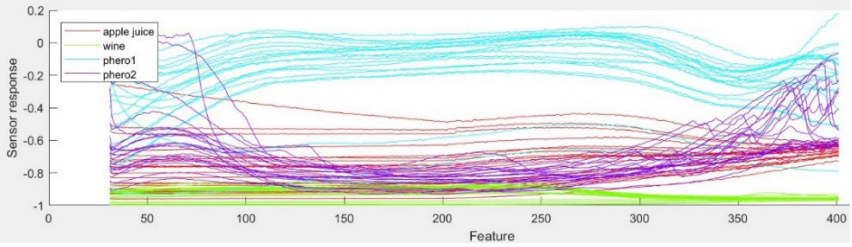


Fig. 5: Data visualization from the VOC sensor detecting bark beetle pheromones

4.3 Threats to the approach

In this section, we discuss the threats that could render the approach unsuitable for the intended usage, based on preliminary assessments of its viability. Although bark beetles as pest for the timber industry are not a new phenomenon, there are still many unanswered questions regarding their biology, like the exact time intervals and climate conditions favorable to their life cycle or determining which life stages (larvae, pupae, imago)

represent which sounds and which part of the tree (trunk or at or directly under the crown) the pioneer beetles fly to. These questions obviously have implications for our sensors and the data they collect. We need to identify the right place to attach the sensors to or discern the beetle's sounds over its life stages, since we want to detect the beetles as early as possible. Furthermore, we need the data from the field to ascertain that the sounds can be unequivocally linked to the beetles and that similar sound – like those from other insects in the tree – can be distinguished.

Biology aside, there are also technical challenges that need to be overcome. As we apply tried and tested algorithms and hardware to the area of bark beetle detection, we focus on the major threats to our AI-setup: classification accuracy, the potential capability of the hardware as AIoT device, the energy consumption of the devices for its application in the field, and the sensors and data. Regarding the models for the audio-based detection, we need a bigger data set to build a more robust model. Furthermore, the data needs to be more suited to the given problem, i. e. data regarding healthy trees and trees with bark beetles for anomaly detection instead of comparing the existing bark beetle recordings with other random sounds. Because of the small data set, the current models were also not tested against unknown data, so the actual performance is to be determined in the future. An economically viable system must not only be inexpensive, but also have a long runtime. Therefore, we want to use SoCs. However, modern power banks switch off automatically when the load is too low, so we need to use a more tailored battery setup.

5. Conclusion and outlook

The approach to gather sound- and pheromone data with cheap and efficient IoT hardware and assess them with off-the-shelf algorithms seems promising. As a proof of concept, we showed that we can detect bark beetles by analyzing audio data, and a visual analysis of VOC data from the lab also looks promising. We currently test the sensor prototypes in the field. For this purpose, the Hunsrück-Hochwald National Park Authority has provided us with various forest areas with indisputable infestation status. We will then evaluate the data we obtain here further and expand the training of the models.

In the future, we plan on supplementing the existing sensors by further environmental sensors, e. g. for recording temperature or soil and air humidity. We argue that trees suffering from drought stress – i. e. the supply of water is too low – are more susceptible to the bark beetle. Not only are they able to produce less resin for defense [Ba21, p. 24], but we also assume that the bark beetle is able to explicitly target such trees. For this purpose, it will be necessary to define drought stress and gather resilient measurements for it. How the transmission of information to the responsible person regarding the infestation can be technically implemented still needs to be clarified. To avoid the problem of lacking network coverage, we plan on testing solutions like LoRaWan and adapt our detection setup if necessary. For the classification, we plan on considering anomaly detection e. g. through kNN and SVM which is also capable of running on SoCs.

Acknowledgement

This project was partly funded by the German Federal Ministry of Food and Agriculture project KI-Pilot under Grant 2820KI001 and partly by the Ministry for Science and Health of Rhineland-Palatinate (Ministerium für Wissenschaft und Gesundheit, MWG) as part of the research training group (Forschungskolleg) "AI-based Cyber-Physical Process Systems".

Bibliography

[Ba21] Bachlehner, Albert: Physiologische und hydrologische Parameter im Zusammenhang mit der Attraktivität trocken-gestresster Fichten (Picea Abies) für den Buchdrucker (Ips Typographus). Master's thesis, 2021. https://abstracts.boku.ac.at/download.php?dataset_id=19993&property_id=107 [2022-05-11].

[DD00] Dubayah, Ralph O.; Drake, Jason B.: Lidar Remote Sensing for Forestry. Journal of Forestry, 98(6):44–46, 06 2000.

[De18] Deter, Alfons: Bis zu zwei Milliarden Euro Schaden am deutschen Wald, 2018. https://www.topagrar.com/management-und-politik/news/bis-zu-zwei-milliarden-euro-schaden-am-deutschen-wald-9840658.html [2022-05-03].

[Du07] Dunn, David: Microphones, Hydrophones, Vibration Transducers: Rolling Your Own, 2007. https://www.zachpoff.com/site/wp-content/uploads/David-Dunn-Microphones_Hydrophones_Vibration-Transducers Rolling_Your_Own Dunn2007.pdf [2022-04-25].

[Dz19] Dziubany, Matthias; Garling, Marcel; Schmeink, Anke; Burger, Guido; Dartmann, Guido; Naumann, Stefan; Gollmer, Klaus-Uwe: Machine learning-based artificial nose on a low-cost iot-hardware. In: Big Data Analytics for Cyber-Physical Systems, pp. 239–257. Elsevier, 2019.

[Fa14] Farina, Almo: Soundscape Ecology. Springer Netherlands, 2014.

[FT14] Frommolt, Karl-Heinz; Tauchert, Klaus-Henry: Applying bioacoustic methods for long-term monitoring of a nocturnal wetland bird. Ecological Informatics, 21, 2014.

[Ga19] Gabriel, Oliver: Kleines Einmaleins des Borkenkäfers, March 2019. https://www.forstpraxis.de/kleines-einmaleins-des-borkenkaefers/ [2022-05-11].

[He15] Heurich, Marco; Krzystek, Peter; Polakowski, Fabian; Latifi, Hooman; Krauss, Andreas; Müller, Jörg: Erste Waldinventur auf Basis von Lidardaten unddigitalen Luftbildern im Nationalpark Bayerischer Wald. Forstliche Forschungsberichte München, 214:101–113, 2015.

[He20] Henning, Hete: Borkenkäfer-Früherkennung aus der Luft, 2020. https://www.uni-goettingen.de/de/document/download/efe458ff383b9ef8e472a48024f580f6.pdf/Rottenburger%20Post%202020.pdf [2022-05-03].

[Hö16] Höbarth, M.; Leitner, T.; Perny, B.; Reh, M.; Wöhrle, M.; Zobl, A.: Borkenkäfer -

Vorbeugung und Bekämpfung. Technical report, LFI Österreich, Waldverband Österreich, Vienna, Austria, 2016. https://de.readkong.com/page/borkenk-fer-vorbeugung-und- bek-mpfung-5680852 [2022-05-11].

[JBS19] Johansson, Annette; Birgersson, Göran; Schlyter, Fredrik: Using synthetic semiochemicals to train canines to detect bark beetle–infested trees. Annals of Forest Science, 76(2), 2019.

[Ko15] Ko, Tom; Peddinti, Vijayaditya; Povey, Daniel; Khudanpur, Sanjeev: Audio augmentation for speech recognition. In: INTERSPEECH. 2015.

[Kr13] Krause, Bernie: Das große Orchester der Tiere. Kunstmann Antje GmbH, September 2013.

[Ma20] Machhamer, Rüdiger; Garling, Marcel; Czenkusch, Levin; Ueding, Kristof; Guldner, Achim; Schneider, Jens; Latif, Azhar; Schmeink, Anke; Naumann, Stefan; Gollmer, Klaus- Uwe; Dartmann, Guido: Development of Edge Runtime Learning Systems for an Artificial Nose Classifying Drinks. In: 9. Medit. Conf. on Embedded Computing (MECO). 2020.

[NMP19] Nanni, Loris; Maguolo, Gianluca; Paci, Michelangelo: Data augmentation approaches for improving animal audio classification, 2019. https://arxiv.org/abs/1912.07756 [2022-04-01].

[Pi11] Pijanowski, Bryan C.; Farina, Almo; Gage, Stuart H.; Dumyahn, Sarah L.; Krause, Bernie L.: What is soundscape ecology? An introduction and overview of an emerging new science. Landscape Ecology, 26(9):1213–1232, 2011.

[Sc93] Schafer, R. Murray: The Soundscape: Our Sonic Environment and the Tuning of the World. INNER TRADITIONS, 1993.

[Sc17] Schebeck, Martin; Hansen, E. Matthew; Schopf, Axel; Ragland, Gregory J.; Stauffer, Christian; Bentz, Barbara J.: Diapause and overwintering of two spruce bark beetle species. Physiological Entomology, 42(3):200–210, 2017.

[Si20] Singh, Neha: Classification of Animal Sound Using Convolution Neural Network. Master's thesis, Technological University Dublin, 2020.

[TH21] Trouvain, Nathan; Hinaut, Xavier: Canary Song Decoder: Transduction and Implicit Segmentation with ESNs and LTSMs. In: Artificial Neural Networks and Machine Learning – ICANN 2021. Springer International Publishing, Cham, pp. 71–82, 2021.

[TP08] Tomiczek, C.; Pfister, A.: Was bedeutet der Klimawandel für die Borkenkäfer?,2008.https://www.waldwissen.net/de/waldwirtschaft/schadensmanagement/insekten/klimawandel-und-borkenkaefer [2022-05-18].

[VH15] Vega, Fernando E.; Hofstetter, Richard W.: Bark Beetles: Biology and Ecology of Native and Invasive Species. ACADEMIC PR INC, January 2015.

[WS98] Wermelinger, B.; Seifert, M.: Analysis of the temperature dependent development of the spruce bark beetle Ips typographus (L) (Col., Scolytidae). Journal of Applied Entomology, 122(1-5):185–191, 1998.

[Zu94] Zuber, Markus: Ökologie der Borkenkäfer. Biologie in unserer Zeit, 24(3):144–152, 1994.

ARTIFICIAL INTELLIGENCE APPLICATIONS

Volker Wohlgemuth et al. (Hrsg.): EnviroInfo 2022,
Lecture Notes in Informatics (LNI), Gesellschaft für Informatik, Bonn 2022 123

Detection of snow-coverage on PV-modules with images based on CNN-techniques

Dennis Hepp[1], Sebastian Hempelmann[2], Grit Behrens[3] and Werner Friedrich[4]

Abstract: The transition from fossil fuels to renewable energy is considered as very meaningful to mitigate climate change. To integrate weather-dependent energies firmly into the power grid, a forecast of the energy yield is very important. This paper is about renewable energy generation by photovoltaic (PV) systems. The yield of PV-systems depends not only on weather conditions, but in wintertime also on the additional factor "snow cover". The aim of this work is to detect snow cover on photovoltaic plants to support the energy yield forecast. For this purpose, images of a PV-plant with and without snow cover are used for feature extraction and then analyzed by using a convolutional neural network (CNN).

Keywords: convolutional neural network; machine learning; python; image recognition; snow detection; photovoltaic

Addresses Sustainable Development Goal 7: Affordable and clean energy

1. Introduction

The climate change affects us all. If it is not sufficiently curbed, our planet could become uninhabitable in the future. [IP22] To prevent this, for example, CO_2 emissions must be drastically reduced. [IP21] With energy consumption on the rise, it only makes sense to reduce CO_2 emissions in the energy production sector as well. This can be achieved by switching from fossil fuels to renewable energies. [IE22] One of the weather-dependent renewable energy sources is solar power, which is generated by PV-systems. These plants can be monitored and based on this data it is also possible to make forecasts for the future. [He20] For the forecast, some factors must be considered. PV-systems depending on weather conditions and location. Plenty of sunshine is an indicator of high electricity yields. However, these can be reduced by objects that cast shadows such as trees or clouds, soiling on the plant, or age-related side effects. [Qu11] To make an accurate energy yield forecast, these factors, as well as the prevailing weather, must be included in the calculation. If the PV-system does not supply as much electricity as predicted, an error

[1] Bielefeld University of Applied Sciences, Faculty of Engineering and Mathematics, Langer Weg 9a, 33332 Gütersloh, Germany, dennis.hepp@fh-bielefeld.de
[2] Bielefeld University of Applied Sciences, Faculty of Minden Campus, Artilleriestrasse 9, 32427 Minden, Germany, sebastian.hempelmann@fh-bielefeld.de
[3] Bielefeld University of Applied Sciences, Faculty of Minden Campus, Artilleriestrasse 9, 32427 Minden, Germany, grit.behrens@fh-bielefeld.de
[4] DSL & Photovoltaikberatung, Frankfurter Str. 9b, 65551 Limburg, Germany, wf@etechnikfreak.de

message is generated in the monitoring system. Which either will be used to coordinate maintenance work or improving the forecast. This is the aim of the research project “PV digital 4.0”, which is funded under grant number 13FH020PX6. This work deals with the snow-coverage on PV plants and is structured as follows, in chapter 2 the aim of the work will be described. Further, this paper involves the detection of snow coating on PV-systems based on images and a convolutional neural network (CNN), which is discussed in detail in Chapter 3. The results are discussed in Chapter 4 and Chapter 5 concludes and gives an outlook for the future.

2. Aim of the work

The goal of this work is to detect snow coverage on PV systems using pictures. To put it into a larger context, this detection is used to label the monitoring data. The labelled data allows us to train a standalone neural network, which can detect snow on plants independent of their location and orientation towards the sun, just using the monitoring data together with weather data. [HBB22]

This is one of the in chapter 1 mentioned factors, which need to be considered in the project “PV Digital 4.0” to increase the accuracy of the yield prediction.

3. Materials and methods

In this work a software based on *Python* has been developed using modules like *datetime* and *glob*. For downloading image data *selenium 4.1.5* was used, to pre-process the data the Python library *OpenCV 4.5.5.64* was utilized and to train a network which should finally detect snow on the PV-systems a CNN was built with *keras 2.9.0* (based on *TensorFlow 2*).

3.1 Overview

A convolutional neural network (CNN) is a special type of neural networks (NN). The neural network is the model used in deep learning. It consists of several layers in which a selected number of neurons are located. These neurons have trainable weights and are connected to each other. How these connections look like is determined by the type of neural network. The architecture, i.e. the number of layers and neurons, is developed by the programmer. Typical neural networks learn a global pattern. CNNs, on the other hand only ever learn a local pattern, so it can recognize edges, textures, etc. in images and is perfectly designed for image recognition. [Ch17] To make this possible, CNN comes up with some special methods, which are explained in more detail in following subsections.

3.2 Convolutional neural network

Convolutional neural networks are typically used for image recognition. The CNN takes an image as their input, and so-called kernels (or filters) are applied to extract features. The input image data for the CNN is handed over with the shape (image_height, image_width, image_channels). A grayscale picture has just one channel, which means every pixel in the picture can be described with one value. These values are the input for the CNN. The colored pictures used in this work come with three channels (RGB: red, green, blue). The kernel is applied on every channel and then summed up. Kernels are matrices and the parameters are trained like weights in a neural network. [Ch17] This 3-D grid is typically processed in the "convolutional block", which is a combination of the mathematical operation "convolution" (convolutional layer) and the "pooling" method. This will be repeated several times. [GBC16] Throughout the convolutional process, the first convolutional layer can learn small patterns, like edges. The next convolutional layer will combine these small patterns to learn bigger, combined patterns. [Ch17] Once a pattern was learned, the CNN will be able to recognize this pattern in any part of the image (equivariant representations). After passing through the convolutional blocks, the output data will be modified with "flattening", so the output can be classified by an MLP. [Fr19] The mentioned methods are described further in the following paragraphs.

Convolution

The convolution is a mathematical operation, which takes the input and overlays a so-called kernel. This kernel is a matrix, which can also be seen as weights from a typical NN and will be used for the whole CNN (parameter sharing). It is smaller than our input, which leads to "sparse interaction" and will be used for each channel. So, the kernel is put on the input. The single values are multiplied with each other and then summed up. Once completely shifted over the whole input a new, smaller matrix, the "feature-map" results. If the down scaling is not wanted, (zero-) "padding" can be used, which won't be discussed further in this work. [Bi08]

Pooling

A part of the convolutional block is pooling. The aim of a CNN is to learn bigger patterns, just using the convolution operation would take a long time to recognize such big patterns. Therefore, the data is compressed and generalized in the pooling layer to speed up the process. The most used pooling approach is the "max-pooling". Max-pooling is modifying the output further by looking at a certain window in the feature-map and only takes the biggest value in that to create the output of the convolutional layer, our final feature-map. [Fr19]

Flattening

Flattening is used to reorder the entries of the 3-D grid resulting from the convolutional layers. The grid will be stringed together to a vector, which then can be used to for the classification task with the MLP. [Fr19]

3.3 Data base

Fig. 1: Image taken by the camera (due privacy, this picture was cropped).

For the described experiment, a camera was set up in the center of Germany 173m above sea level in Hesse (coordinates: 50.37194660775961, 8.122048403586822). The camera is pointing to a private roof with a PV-plant on it (shown in Fig. 1) and is accessible via Wi-Fi. A script was written, which can connect to the interface of the camera and allow us to receive images. The images were sent as HTTP-request to a LINUX server running the script. Via cronjob, this process was executed every five minutes during the day from the end of November 2021 to April 2022. Due to transmission errors, the camera was not always available, so there were gaps in the data. In total 19.730 images could be collected in the experiment. Among these, 111 of the images show snow cover on the PV-system. Afterwards, the visible parts of the plant were cut out in rectangles (see Fig. 2) with OpenCV and were saved. This resulted in a total of 256 usable images of a snow-covered PV-system and a lot more non-snow-covered images, that could be used for training.

Fig.2: The PV-plant can be divided into three smaller parts.

3.4 Data selection

For good results, the training data set should be well-balanced with the same number of pictures with snow cover and without snow cover. For this purpose, care must be taken in selecting days with similar weather conditions, since the position of the sun affects the light conditions in the image, as well as the reflection on the solar panels. In addition, depending on the time of day, the roof, and dormer casts a shadow (e.g. Fig. 3) on some PV-modules.

(a) with snow

(b) without snow

Fig. 3: In between these two sunny-day pictures are three days. This grants us similar conditions, like the angle of the shadow.

3.5 Pre-processing

Unfortunately, in the period in which the pictures were taken, it didn't snow often, so the amount of data is small. Good results come naturally from good data, so the mass of qualitative images is an important factor. To increase the number of images indirectly, the *ImageDataGenerator* of *keras* was used, which randomly rotates, shifts, shears, and zooms into the training images during the training process of the model and changes the

size of the images to 512 x 512. This is also called “data augmentation”. The data augmentation is applied for each epoch. So, in each epoch, our CNN gets a randomly generated dataset based on the rules set in the *ImageDataGenerator* and our base dataset. [Fr19]

3.6 Data analysis

For image recognition we used a CNN (shown in Fig. 4) as described in 3.2. The CNN takes an input with the shape (512, 512, 3) because we rescaled all images in 3.5. We used RGB pictures, therefore we needed to set the image_channel to three. In total, we used four convolutional blocks, a 2-D convolution with the activation function “ReLU” followed by a 2-D MaxPooling. The Kernel for the convolution was given the size of 3 x 3 and the MaxPooling window is set to 2 x 2. This helps us to focus on the important features. The first convolution is set to give us an output dimension of 32, so our first output shape will be (510, 510, 32). After the MaxPooling we have a shape of (255, 255, 32). The second convolution is set to 64 and the third and fourth to 128. At the end we got a shape of (30, 30, 128) this will be flattened to a 1-D vector with the shape of (115200), which is basically the multiplication of the shape of our last convolutional layer (30*30*128 = 115200). Our dense layer of the MLP is set to an input of 512 neurons, also with the “ReLU” activation function. Followed by a dense layer with one neuron and the “sigmoid”-function, which will take the input to give an output between 0 and 1. A threshold is then used for the classification, every output above the threshold will be classified as 1 for “snow-coverage” and every output below the threshold will be classified as 0 for “no snow-coverage”. For the loss function, we used the “binary cross-entropy”, which is common for CNNs. [Fr19] The common optimizer “Adam” was chosen for the optimization task of this network. The dataset was split into three data sets: 70% train data, 20% validation data and 10% test data.

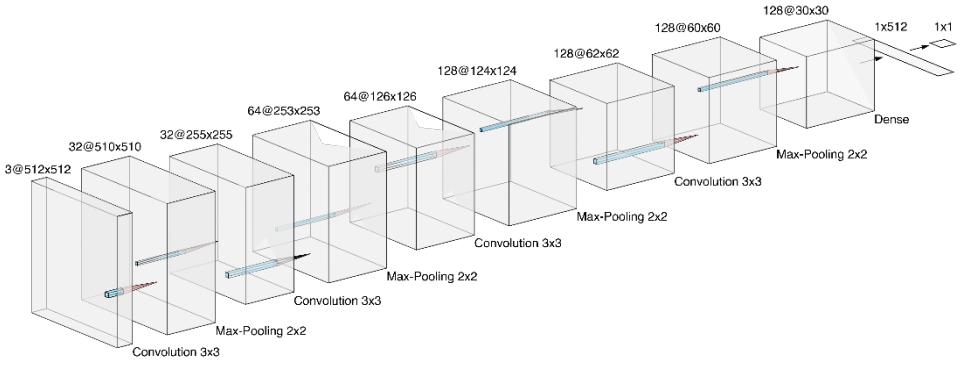


Fig. 4: This is a visualization of the used CNN. The first cube represents our input layer with the shape (512, 512, 3). The convolution and the MaxPooling process is visualized by the blue cutouts pointing to the next cube. Every convolution step is increasing the third

dimension, which were the RGB-values of the input image in the first step. The MaxPooling is decreasing the first and second dimension, which is the height and width of the image. After a few steps e.g., the sixth cube got a shape of (124, 124, 128), which means the image got compressed to a height and width 124 x 124 with 128 features. After the eighth cube, the data will be flattened and commit it to our first Dense layer with 512 neurons.

4. Results

The results of the training are visualized in Fig. 5. The accuracy while training achieved more than 96%. At the end, the CNN was able to predict 100% of the validation data set. The loss value of the training was around 0.16 and for the validation set 0.07. Using the test data to test the CNN leads us to 92% accuracy. The gap between the accuracy of the test and train data set is caused by overfitting. Also, overfitting led our valid prediction to 100%.

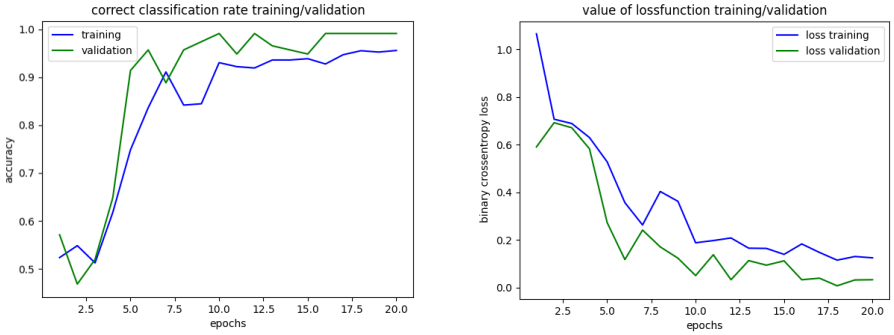


Fig. 5: Accuracy and loss visualized. The accuracy value is in between 0 and 1, which can be transformed into %. This value shows how much percent of the data set was classified correctly. The loss value is given by our loss function (binary cross-entropy), it is the feedback-signal for the CNN, which is used in the training process. The aim of the CNN is to minimize this value while training.

Due to randomly generated test, train, and valid data sets one of the rare cases, having sunny weather with and without snow (shown in Fig. 3), wasn't well distributed. So, the case "with snow" on the PV-plant and "sunshine" made it into the train but not into the valid data set, while "sunshine without snow" on the other hand made it into both sets. Which means the net couldn't validate the special case shown in Fig. 3a. As we used images with the same weather conditions (described in 3.4), the CNN could have learned to recognize sunshine/shadow pictures as no snow-coverage, even when there was snow (see Fig. 6).

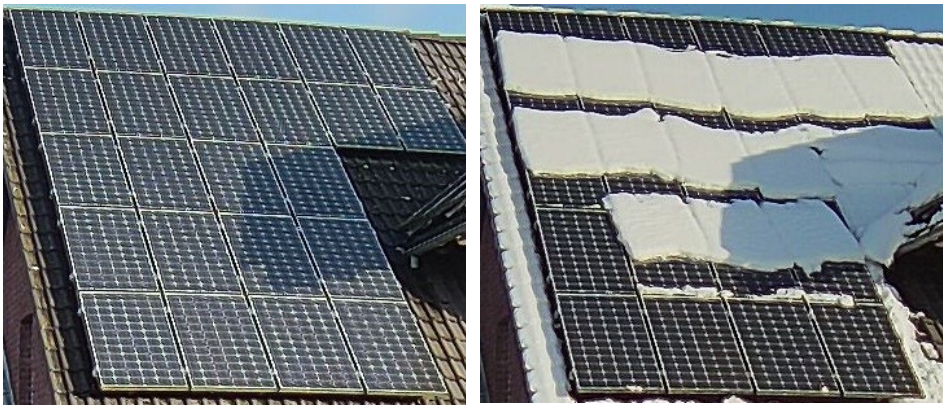


Fig. 63: Both pictures are from the test data set. The left one is showing the class "no snow-coverage", the right one shows "snow-coverage". Both pictures were classified as "no snow-coverage" by our trained model.

5. Conclusion

Even with a small data set we were able to get good results. Overfitting is a problem for our CNN because snow on PV-plants won't ever have the same features as shapes and density. For this, some techniques like the data augmentation mentioned in 3.5 or adding a “Dropout”-layer can be used to prevent this problem. This CNN will be improved with new data from the next winter (in 2022/2023). With the new database, we will increase the accuracy of the CNN. The plan is to take pictures with more cameras on an industrial PV-plant in Germany, which is higher above the sea level and tends to have more snow during wintertime. The next step for the automation of the snow data labeling is either to improve the CNN to automatically detect the snow coverage in percent or to develop an algorithm, which can calculate the snow amount on the plant. With this automated labeling process and with more data, we can also improve the work in [HBB22] and train a better model to recognize snow coverage on PV-plants just using the monitoring and weather data.

Bibliography

[Bi08] Bishop, C.M.: Pattern Recognition and Machine Learning. Springer, 2006.

[Ch17] Chollet, F.: Deep Learning with Python. Manning, 2017.

[Fr19] Frochte, J.: Maschinelles Lernen. Hanser, 2019.

[GBC16] Goodfellow, I.; Bengio, Y.; Courville, A.: Deep Learning. MIT Press, 2016.

[IE22] IEA, International Energy Agency, Global electricity demand is growing faster than renewables, driving strong increase in generation from fossil fuels, https://www.iea.org/news/global-electricity-demand-is-growing-fasterthan-renewables-driving-strong-increase-in-generation-from-fossil-fuels, accessed: 12/07/2022.

[He20] Hempelmann, S. et al.: Evaluation of unsupervised anomaly detection approaches on photovoltaic monitoring data. In: 47th IEEE Photovoltaic Specialists Conference (PVSC). Calgary, AB, Canada, 15.06.2020 - 21.08.2020: IEEE, pp. 2671–2674.

[HHB22] Hepp, D.; Hempelmann, S.; Behrens, G.: Detektion von Schneebedeckung mittels stringbasierter Monitoringdaten und Methoden des Maschinellen Lernens. In: 37. PV-Symposium. Conexio-PSE GmbH, Pforzheim, pp. 777-788, 2022.

[IP21] IPCC, 2021: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, 2021.

[IP22] IPCC, 2022: Climate Change 2022: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, 2022.

[Qu11] Quaschning, V.: Regenerative Energiesysteme. Technologie - Berechnung – Simulation. Hanser, 2011.

Volker Wohlgemuth et al. (Hrsg.): EnviroInfo 2022,
Lecture Notes in Informatics (LNI), Gesellschaft für Informatik, Bonn 2022 133

Information Disclosure by Decentralized Coordination in Virtual Power Plants and District Energy Systems

Jörg Bremer[1] and Sebastian Lehnhoff[2]

Abstract: Grouping small, hardly predictable, and volatile energy resources to jointly operating virtual power plants with sufficient flexibility for coordination is widely seen as a key aspect of integrating renewable energy into the grid. For several reasons, self-organizing, agent-based systems are probably the best technology for coordination. A major drawback of many currently existing solutions is the necessity to communicate plain information for negotiation and optimization. Such information contains e.g. possible energy generation schemes or aggregated costs. Previous works have already shown that identification of anonymously sent information is possible. In this paper, we demonstrate the possibility of disaggregating cost structure information as an example of possible leakage of business information in the case of participation in virtual power plants or district energy systems. From this perspective, we derive measures to ensure privacy preservation in decentralized coordination algorithms.

Keywords: Virtual Power plant; Distributed Optimization; Self-Organization; Data Privacy

Addresses Sustainable Development Goal 7: Affordable and clean energy

1. Introduction

Integrating as much as possible renewable energy sources into the energy grid is one of the most crucial task today – not only for fighting global warming, but also for ensuring energy safety. One goal here is achieving energy efficiency by a higher penetration of renewable feed-in [Be13; Ka11]. Higher penetration of renewable energy needs modern concepts for integration into the power grid due to their volatile nature and low flexibility. Low voltage coupling points with the grid lead to partially inversed power flows and demand local optimization. One way to cope with volatility and small size is bundling of different energy resources and orchestrating them via communication and joint control. This concept is also known as virtual power plant (VPP) [NM12].

For many use cases, it is advantageous to bundle energy resources within a local region. Here, energetic neighborhoods come into play, especially regarding multi-modal energy systems in which for example a complex interplay of electricity and district heating grid are scrutinized for synergies [Th17]. The basic coordination problem and thus also the

[1] University of Oldenburg, Energy Informatics, Uhlhornsweg 84, 26129 Oldenburg, Germany, jbr@offis.de
[2] University of Oldenburg, Energy Informatics, Uhlhornsweg 84, 26129 Oldenburg, Germany, lehnhoff@uol.de

resulting problems are mostly the same as in VPP.

Multi-agent-based systems are widely considered to be a valid approach for coordinating a large number of distributed entities or sensing or operation equipment, and to solve distributed optimization and control problems in cyber-physical systems, especially for horizontal-distributed control tasks. In addition, hierarchical topologies supporting vertical- distributed control are already available. For optimizing industrial production and logistics processes, multi-agent systems have been on the research agenda and been used for many years. Use cases comprise for example supply chain management [CM06], or production planning [Tö02], and scheduling [OK07].

Solving a problem with the help of autonomously acting distributed entities – such as agents – naturally raises an issue with safety concerns; especially if a decentralized consensus solution has to be implemented in critical infrastructures or processes [Fa18]. Solving such problems within agent coalitions needs the exchange of information for negotiation. Exchanging information via messages is necessary to build up agents' beliefs for problem solving and inevitably allows insight into other agents' options.

As an example, in the predictive scheduling use case in VPPs frequently operational schedules are exchanged as proposal for the own choice of action [HS16]. Each schedule contains data of the possible portion of energy that may be generated (or consumed) during the same given time period. Today, the resolution is usually 15 minutes per time interval (often for a day). In future, finer grained resolutions can be expected. With each round during negotiation, a new possible operable schedule is sent to several other agents together with transient information on other agents' schedules. As the underlying problem is of a multi-objective nature, several performance indicators for evaluating a schedule according to different criteria often accompany the mere electro-technical information.

This is also known as gossiping principle that comes into play into many decentralized coordination algorithms [KV07]. Such information can be collected and aggregated by malicious agents. In the case of energetic neighborhoods, actors from a close vicinity are drawn closely together for long-term collaboration. This means that data from more than just a single coordination process could be collected to extract some meaningful information; [Da18] already showed that collected information can easily be assigned to specific energy devices, and even to corresponding businesses. Different machine learning methods have been tested to achieve this. Moreover, collected and aggregated schedules allow for deriving detailed information on internal processes - heating profiles, and thus working hours, machinery load factors in case of internal consumption optimization with batteries, or current capacity utilization [BL19; Da18].

In this contribution, we extend the concern and demonstrate that is also possible to derive cost and pricing information from individual tariffs even if contained only in aggregated form. The rest of the paper is organized as follows. We recap some previous work on data privacy in distributed algorithms and present a use case study that reveals information disclosure even for aggregated data in distributes optimization.

2. Related Work

Advances in information technologies have further increased long existing concerns of privacy. When it comes to autonomous agents acting on behalf of a business, several privacy and information leakage concerns can be raised. A good overview can be found in [SEG14].

Surprisingly low effort has so far been put into the question of data privacy when it comes to (decentralized) algorithm design. A method based on the alternating direction method of multipliers (ADMM) for solving decentralized optimization in an agent system with preserved privacy can be found in [ZAW19]. ADMM solves convex optimization problems by breaking them into smaller pieces that can for example be solved individually by agents [Bo11]. To incorporate privacy preservation, partially homomorphic cryptography has been integrated. Unfortunately, applications are limited to convex functions and the method cannot be applied to non-convex black-box optimization, what is often the case in decentralized agent coordination scenarios. An extension of the use of ADMM to distributed machine learning can be found in [Wa20].

Two frequently occuring tasks in multi-agent systems are distributed constraint satisfaction (DisCSP) and distributed constraint optimization (DCOP). For these two problem classes, algorithms have been developed that aim at preserving anonymity in multi-agent problem solving. The major concern in DisCSP and DCOP algorithms is that they usually leak information that can be exploited to infer private information of other agents [GPT06]. By integrating anonymity into specialized protocols, shared information cannot be linked to the corresponding agent. Examples can be found in [BM03; SGG07; SM04; YSH05]. Some of them still leak at least some information. All these approaches still communicate plain information and try just to disguise the sender of the information. On the other hand, [Da18] demonstrated for some coordination tasks in decentralized energy management that enough information is exchanged to still identify the origins of shared information.

A way more frequent use case in the energy sector is coordination of energy generation and consumption. In [BL19] a prototypical application was scrutinized that uses order preserving encryption [Ag04] to solve the predictive scheduling problem in virtual power plants. The possibility of direct integration as well as the performance which barely degraded by encoding offers some advantages over other encryption schemes. Simply the objective function that is minimized by the agents (locally as well as globally) has to be restated. Because no mathematical functions are supported, objective functions can only rely on the order of input values from different agents. An implementation for the sum of input values has been shown in [BL19]. On the other hand, this is already the biggest disadvantage of this method. Only a few number of special cases can be implemented with these methods.

Collected information can be analyzed with appropriate machine learning methods. In [Da18], an algorithm for decentralized, agent-based scheduling in virtual power plants has

been scrutinized. It was found that schedules can be properly assigned to specific devices (and thus businesses) with machine learning. In district energy systems, the concrete business behind an agent may even already be known due to public information on the other members in the energetic neighborhood. In the researched examples, the collected schedules allowed for deriving detailed information on internal processes and thus on heating profiles, working hours, machinery load factors in case of internal consumption optimization with batteries, current capacity utilization, etc. The same holds true for consumption patterns.

3. Data Disclosure in Self-Organization

As a case study for a possible leakage of information during self-organized coordination in VPP we consider the disclosure of individual cost or prices (depending on whether we spy out a generator or a consumer).

3.1 Problem Description

Predictive scheduling is the problem of finding an operation schedule (determining the individual course of generated or consumed power) for each energy resource within a VPP for a given future time horizon. Today, often planning is made for 96 time intervals of 15 minutes each for the next day.

This constitutes a distributed combinatorial nature of the optimization problem [18] for which several solution have been proposed [20], [35], [14]. Decentralized algorithms are seen as the most promising approach due to the distributed architecture and problem size. Additionally, in district energy systems of energy cooperatives of individual and self-dependent actors, centralized authorities that dictate the generated amount of power may spoil acceptance.

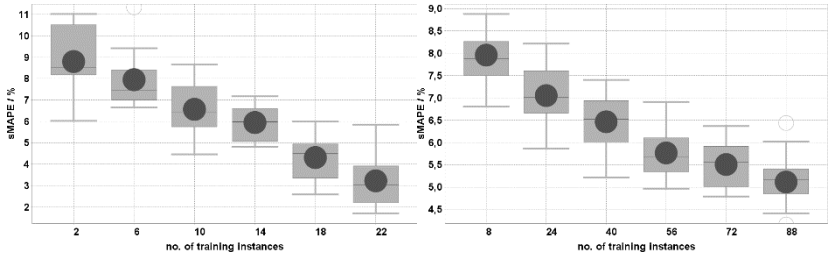
Solving distributed problems with agent-based, decentralized approaches leads to information exchange to build up the agents' beliefs for problem solving. This information could be collected and aggregated. In case of energetic neighborhoods, actors from a close vicinity are drawn closely together for long-term collaboration. In this case, schedules from more than just a single optimization process could be gathered. Collected schedules would allow for deriving detailed information on internal processes – heating profiles, and thus working hours, machinery load factors in case of internal consumption optimization with batteries, current capacity utilization, etc. The same holds true for consumption patterns. Thus, schedules and thereof derived phase spaces of device operations should actually not be publicly known.

3.2 Case Study

For our case study, we consider a small business that plans internal production schemes after some individually with the energy provider negotiated time of use tariff. In this way, for each time interval of the planning horizon, a different energy price has to be paid [So15]. W.l.o.g., we generated random prices $c \in [40, 80]$ cent for our experiments. We assume the following scenario. A group of distributed energy resources (as members of a VPP or a district energy system) is conducting a distributed (day-ahead) planning of energy consumption and generation with the goal of balancing as much as possible and to minimize overall energy cost. To incorporate individual cost in the decentralized balancing algorithm, different schedules are sent by the agents as proposal during negotiation. Each schedule s_i must be annotated with total individual energy cost

$$c_i = \sum_{j=1}^{d} s_i[j] \cdot c[j] \tag{1}$$

as these cannot be calculated by the other agents. Tariffs are not public. During coordination, a set of n different schedules could be collected by a fraudulent agent. If there is a number of schedules available equal to or greater than the schedule dimension, the system of equations is fully determined and the exact energy cost for each time interval can be derived. But, also if the system is undetermined, we can try to build an approximate model of the cost. We conducted some experiments with particle swarm optimization to fit the model.



(a) Scenario 1: 24 intervals, 1 day (b) Scenario 2: 96 15-minute intervals, 1 day

Fig. 1: Model fitting results of tariff estimation by with particle swarm optimization for two one day scenarios.

For the experiments, one agent was chosen to be the fraudulent agent during optimization. It was the task of this agent to collect all schedules of the other agents together with the total energy cost according to different individual time of use tariffs. From the collected unique schedules of a specific agent, tariff information can be calculated if the system is fully determined. This is the case as soon as the number of collected schedules is equal to or larger than the number of time intervals in a schedule. But even with less schedules an approximation can be calculated. We use particle swarm optimization to fit a model for the

under determined system. For each model fitting a slightly different best guess for the tariff approximation will be found. Thus, we repeated model fitting 20 times and took the mean tariff as the approximation.

Fig. 1 show the result for two different scenarios: 24 and 96 time intervals with different numbers of collected schedules that could be used for model fitting. As optimization protocol

dim.	no. of agents		
d	10	25	50
96	32.51 ± 15.65	36.23 ± 19.35	38.11 ± 14.23
24	10.53 ± 4.04	11.68 ± 1.74	15.39 ± 7.24
8	6.64 ± 3.05	7.25 ± 1.79	7.46 ± 2.92

Tab. 1: Number of mean unique schedules (per agent) of other agents that a fraudulent agent sees during a single distributed optimization process.

we used the one proposed in [HS17]. For measuring the quality of the model, we used the symmetric mean absolute percentage error (sMAPE):

$$sMAPE(f,a) = \frac{100}{n}\sum_{i=1}^{n}\frac{|f_i - a_i|}{(|a_i| + |f_i|) - 2} \tag{2}$$

The results show that a relative good fitting can be achieved with a rather small number of schedule information. Tab. 1 shows the result of another experiment. Each agent counted the number of unique schedules from other agents that were seen during the optimization process. This result shows that already during a single optimization process enough schedules are communicated to leak information also for indirect data that is contained only in aggregated form (like time interval individual pricing information).

From these results as well as from the findings of [Da18], we clearly see that is necessary to raise awareness for privacy interests and appropriate measures in distributed problem solving in (future) cyber-physical energy systems with a high share of autonomous functions. Research is still at the beginning, when it comes to suitable encryption schemes that could be used to secure the shared information in such systems. Some first examples for the centralized server can be found in [KLG19], but proper best practices and design schemes for systematically integrating these approaches into distributed problem solving are missing so far and are worth to be given more attention in the future.

4. Conclusion

With this contribution, we wanted to raise a general awareness of the information disclosure problem in distributed and self-organized systems. When local information on possible behavior is spread to other actors in a multi-agent system in order to achieve some

consensus on coalition behavior, the chance of revealing private information is often unintentionally given. As today, obviously there is no technology, which could be used out-of-the-box to tackle data masking in decentralized algorithms.

Thus, more research in the field of encrypted (distributed) optimization and self-organization is highly recommended.

Bibliography

[Ag04] Agrawal, R.; Kiernan, J.; Srikant, R.; Xu, Y.: Order preserving encryption for numeric data. In: Proceedings of the 2004 ACM SIGMOD international conference on Management of data. S. 563–574, 2004.

[Be13] Becker, T.; Boschert, S.; Hempel, L.; Höffken, S.; Obst, B.: Complex urban simulations and sustainable urban planning with spatial and social implicati- ons. In: Int Ann Photogrammetry, Remote Sensing Spat Inf Sci, ISPRS 8th 3DGeoInfo conference and WG II/2 workshop. Bd. 2, W1, 2013.

[BL19] Bremer, J.; Lehnhoff, S.: Encrypted Decentralized Optimization for Data Masking in Energy Scheduling. In: Proceedings of the 2019 3rd International Conference on Big Data Research. ICBDR 2019, Association for Computing Machinery, Cergy-Pontoise, France, S. 103–109, 2019, isbn: 9781450372015, url: https://doi.org/10.1145/3372454.3372487.

[BM03] Brito, I.; Meseguer, P.: Distributed forward checking. In: International Confe- rence on Principles and Practice of Constraint Programming. Springer, S. 801– 806, 2003.

[Bo11] Boyd, S.; Parikh, N.; Chu, E.; Peleato, B.; Eckstein, J. et al.: Distributed optimization and statistical learning via the alternating direction method of multipliers. Foundations and Trends in Machine learning 3/1, S. 1–122, 2011.

[CM06] Chaib-Draa, B.; Müller, J.: Multiagent based supply chain management. Springer Science & Business Media, 2006.

[Da18] Dabrock, K.: Privacy in der automatisierten prdiktiven Einsatzplanung von Energieanlagen im Smart Grid, Magisterarb., University of Oldenburg, Dept. of Energy Informatics, Germany, 2018.

[Fa18] Farris, I.; Taleb, T.; Khettab, Y.; Song, J.: A survey on emerging SDN and NFV security mechanisms for IoT systems. IEEE Communications Surveys & Tutorials 21/1, S. 812–837, 2018.

[GPT06] Greenstadt, R.; Pearce, J. P.; Tambe, M.: Analysis of privacy loss in distributed constraint optimization. In: AAAI. Bd. 6, S. 647–653, 2006.

[HS16] Hinrichs, C.; Sonnenschein, M.: Design, analysis and evaluation of control algorithms for applications in smart grids. In: Advances and New Trends in Environmental and Energy Informatics. Springer, S. 135–155, 2016.

[HS17] Hinrichs, C.; Sonnenschein, M.: A distributed combinatorial optimisation heuristic for the scheduling of energy resources represented by self-interested agents. International

Journal of Bio-Inspired Computation 10/2, S. 69–78, 2017.

[Ka11] Karnouskos, S.: Demand side management via prosumer interactions in a smart city energy marketplace. In: 2011 2nd IEEE PES international conference and exhibition on innovative smart grid technologies. IEEE, S. 1–7, 2011.

[KLG19] Kaiser, P.; Langer, A.; Gaedke, M.: MPCC: Generic Secure Multi-Party Com- putation in Centralized Cloud-based Environments. In: Proceedings of the 2019 3rd International Conference on Big Data Research. S. 60–66, 2019.

[KV07] Kermarrec, A.-M.; Van Steen, M.: Gossiping in distributed systems. ACM SIGOPS operating systems review 41/5, S. 2–7, 2007.

[NM12] Nikonowicz, L. B.; Milewski, J.: Virtual power plants-general review: structure, application and optimization. Journal of power technologies 92/3, S. 135, 2012.

[OK07] Opadiji, J. F.; Kaihara, T.: Distributed production scheduling using federated agent architecture. In: International Conference on Industrial Applications of Holonic and Multi-Agent Systems. Springer, S. 195–204, 2007.

[SEG14] Such, J. M.; Espinosa, A.; García-Fornes, A.: A survey of privacy in multi-agent systems. The Knowledge Engineering Review 29/3, S. 314–344, 2014.

[SGG07] Smith, M.; Grosz, B.; Greenstadt, R.: SSDPOP: Improving the Privacy of DCOP with Secret Sharing. Proceedings of Autonomous Agents and Multi- Agent Systems, AAMAS-2007/, 2007.

[SM04] Silaghi, M.-C.; Mitra, D.: Distributed constraint satisfaction and optimization with privacy enforcement. In: Proceedings. IEEE/WIC/ACM International Conference on Intelligent Agent Technology, 2004.(IAT 2004). IEEE, S. 531– 535, 2004.

[So15] Sonnenschein, M.; Lünsdorf, O.; Bremer, J.; Tröschel, M.: Decentralized control of units in smart grids for the support of renewable energy supply. Environmental Impact Assessment Review 52/, S. 40–52, 2015, issn: 0195-9255.

[Th17] Thiem, S.; Danov, V.; Metzger, M.; Schäfer, J.; Hamacher, T.: Project-level multi-modal energy system design-Novel approach for considering detailed component models and example case study for airports. Energy 133/, S. 691– 709, 2017.

[Tö02] Tönshoff, H.; Herzog, O.; Timm, I.; Woelk, P.: Integrated process planning and production control based on the application of intelligent agents. Proc. 3rd CIRP ICME/, S. 135–140, 2002.

[Wa20] Wang, X.; Ishii, H.; Du, L.; Cheng, P.; Chen, J.: Privacy-preserving distributed machine learning via local randomization and ADMM perturbation. IEEE Transactions on Signal Processing 68/, S. 4226–4241, 2020.

[YSH05] Yokoo, M.; Suzuki, K.; Hirayama, K.: Secure distributed constraint satisfac- tion: Reaching agreement without revealing private information. Artificial Intelligence 161/1-2, S. 229–245, 2005.

[ZAW19] Zhang, C.; Ahmad, M.; Wang, Y.: ADMM Based Privacy-Preserving Decentra- lized Optimization. IEEE Transactions on Information Forensics and Security 14/3, S. 565–580, 2019.

Volker Wohlgemuth et al. (Hrsg.): EnviroInfo 2022,
Lecture Notes in Informatics (LNI), Gesellschaft für Informatik, Bonn 2022 141

Using Deep Learning for automated birth detection during farrowing

Jan-Hendrik Witte[1], Johann Gerberding[2], Clara Lensches[3], Imke Traulsen[4]

Abstract: Pig livestock farming has been undergoing major structural change for years. The number of animals per farm is constantly increasing, while competition is becoming more intense due to volatile slaughter prices. Sustainable, welfare-oriented livestock farming becomes increasingly difficult under these conditions. Studies have shown that animal-specific birth monitoring of sows can significantly reduce piglet losses. However, continuous monitoring by human staff is inconceivable, which is why systems need to be created that assist farmers in these tasks. For this reason, this paper aims to introduce the first step towards an automated birth monitoring system. The goal is to use deep learning methods from the field of computer vision to enable the detection of individual piglet births based on image data. This information can be used to develop systems that detect the beginning of a birth process, measure the duration of piglet births, and determine the time intervals between piglet births.

Keywords: precision livestock farming, birth monitoring, deep learning, computer vision

Addresses Sustainable Development Goal 9: Industry, innovation and infrastructure

1. Introduction

The structures of modern pig livestock farming, and piglet production have changed significantly in recent years. The situation report of the German Farmers' Association shows the opposite trend of a steadily decreasing number of farms with a simultaneous increasing number of sows held per farm [De20]. A total of 70% of all sows housed in Germany are kept on the largest 2,000 individual farms, each with 250 breeding sows per farm. Meanwhile, the slaughter price has been highly volatile in recent years, which further intensifies competition and poses major challenges for the farmer now and in the future. At the same time, politics and society alike are calling for more sustainable and more animal-friendly husbandry [Be14], which creates additional pressure and makes economically profitable livestock farming increasingly difficult. These challenges cannot be met with conventional methods, which is why new and innovative solutions are needed. As a result, research in the domain of precision livestock farming (PLF) has increased in recent years. PLF describes systems that utilize modern camera and sensor technologies

[1] University of Oldenburg, jan-hendrik.witte@uni-oldenburg.de
[2] University of Oldenburg, johann.gerberding@uni-oldenburg.de
[3] University of Göttingen, clara.lensches@uni-goettingen.de
[4] University of Göttingen, imke.traulsen@uni-goettingen.de

to enable automatic real-time monitoring in livestock production to supervise animal health, welfare and behaviour [Be14] [D'18]. This involves the automated acquisition, processing, analysis and evaluation of sensor-based data like temperature, humidity, NH3 or CO2 concentration [Co18] as well as image and video data [Ch20]. These distinct types of information and data sources hold the potential to enable data-driven assistance systems that support farmers in their daily work and would help them adapt to the constantly changing conditions in sow livestock farming.

To build such systems, methods are first needed that allow the automated processing of these types of data streams in the form of image, video, and sensor data. Video and image data alone can be used for a variety of PLF related use cases, many of which can already be found in the literature. For example, methods from the field of computer vision (CV) can be utilized to detect changes in activity of sow behaviour during final gestation [Kü20], which contains valuable information for interpreting the sow's behaviour and could be used for various subsequent processes. Similarly, Lao et al. investigated the automatic behaviour detection of lactating sows based on image data [La16]. Especially the image-based detection of objects in the context of PLF poses a particular challenge with various problems such as the grouping, overlapping and occlusion of animals, their different postures, orientations and positions, as well as constantly changing environmental factors such as different lighting conditions, soiling of animals or occlusion caused by objects in the pen. Due to their ability to generalize, the use of deep learning (DL) methods from the field of CV has been proven effective in addressing these challenges.

One topic area that has not yet been considered in the literature using these techniques is the birth monitoring of sows. Various studies in the field of birth monitoring have already proven that constant and targeted observation can reduce piglet loss during the birth process [Ho95]. White et al. [Wh96] were able to reduce piglet losses from 18.2% to 10.1%, through targeted birth monitoring based on a custom protocol. However, intensive, permanent observation of the farrowing process of individual births is not feasible in practice. There is a need to create systems that allow automated monitoring of birth processes that informs the farmer as soon as individual problems like stillbirths or prolonged farrowing is detected. This paper aims to lay the foundation for the development of such systems by developing a model for automated birth detection based on video streams.

The paper is structured as follows: First, the current state of the art in birth monitoring in sows will be presented. The primary focus lies on papers that apply DL models and architectures as well as their respective performance. This is followed by the introduction of the approaches for automated frame-based birth detection considered in this paper. Subsequently, a description of the workflow for data collection, preparation and analysis as well as model selection for each respective approach will be presented. In addition, the data sets created for each approach are described, as well as the test environment in which the different models were instantiated, trained, and evaluated. To conclude, the current status of this research is described. Here, the problems and challenges are addressed,

potential solutions are presented, and the future proceedings are described in more detail.

2. Related Work

There are currently no papers available which have addressed the topic of automated birth detection based on DL methods from the field of CV. So far, the literature has considered use cases that address the automatic detection of different body conditions of the sow as well as use cases that are located before and after the actual farrowing event. In terms of body condition, Cang et al. [Ca19] use a custom Faster-R-CNN with an additional regressive branch for initial sow detection and subsequent weight estimation with an average absolute error of 0.644 kg and a relative error of 0.374%, while Huang et al. [Hu19] apply convolutional neural networks (CNN) to determine body condition scores of individual sows. Behaviours such as nest building, which can be observed prior to the actual farrowing, have been addressed in the literature by using accelerometer data from sensors to classify nest building behaviour with a generalized linear model, achieving an accuracy of 85% on the applied test set [Oc15]. Kasani et al. use different DL architectures to detect and classify sow posture into laying left, laying right, sitting and standing. The authors evaluated variations of DenseNet, VGG and Inception architectures as well as MobileNet based on a custom data set, in which the DenseNet121 achieved an accuracy of 99,83% in the classification of sow posture [Ka21]. However, most papers in the literature address the automated behaviour and posture detection of lactating or nursing sows [La16], [Wa21], [Ya18], [Zh18], [Zh20a]. Zheng et al. [Zh20b]for example use Faster R-CNN for sow posture classification into standing, sitting, ventral lying and lateral lying and achieve a mean average precision (mAP) of 0.927, while Zhang et al. [Zh19] apply a combination of MobileNet and SSD network for sow behaviour detection in drinking, urinating and mounting behaviour with an accuracy of 0.965, 0.914 and 0.923 respectively and an overall mAP of 0.934.

We found one work in progress paper which introduces an embedded system to monitor farrowing, in which the actual birth detection of piglets is considered. Silapachate et al. [Si18] applied histogram equalization, background subtraction and edge detection for image pre-processing and plan to apply histogram of oriented gradients, different machine and DL models like support vector machines or CNNs to train a binary classification model that "distinguish video frames with a newly farrowed piglet and those without". The method presented in this paper differs in the following aspects:

- The piglet birth itself should be classified. Unlike Silapachote et al. [Si18], this should not be based on a newly detected piglet in the pen, but on the distinctive visual features in the area of the vulva during the birth event.
- In addition to the binary classification approach, the use case will also be addressed based on an object detection approach by using bounding boxes to recognize, localize and classify the farrowing event.

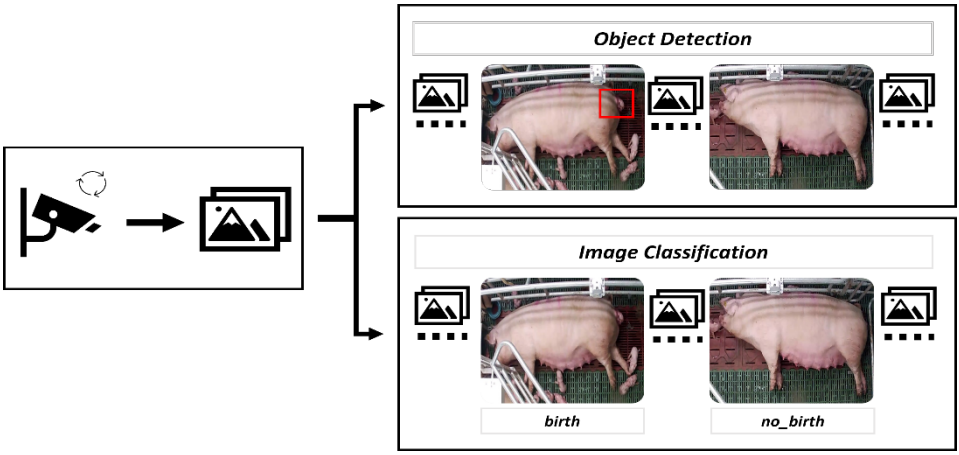


Fig. 1: Overview of approaches

3. Approaches for farrowing detection

Automated detection of a single birth event in the overall farrowing process is defined as both an object detection as well as an image classification use case. Object detection describes the detection and localization of objects of a defined class by enclosing bounding boxes around the respective object in the image, while image classification describes the task of assigning a given image to a defined label or class [Wa19]. From a system point of view, the basic idea is to split a given video stream into single frames and classify on each frame whether a birth event is taking place or not. Fig. 1 describes this process. Depending on the investigated method, the image data is processed differently. In the case of the object detection approach, the birth of a piglet is detected and localized using bounding boxes. Frames in which the object detection model predicts a bounding box with high confidence therefore in theory contains a farrowing event, frames in which no bounding box was placed correspondingly do not contain a birth event. In the case of the image classification approach, the images are processed as a whole and classified into *birth* and *no_birth* as a binary classification task. Both approaches have immediate advantages and disadvantages in terms of implementation and data preparation. While data preparation for an object detection task requires manual placement of bounding box annotations on each image, preparation of the image classification dataset only requires the categorization in one of the two defined classes, which can be done much more efficiently. However, the bounding box annotation provides a direct bounding of the context to be considered within the image, which is beneficial for the actual detection and localization of the birth event. Since the image classification model processes the image as a whole, the corresponding approach does not have this property, which could make the classification of the frames more difficult. Data preparation, dataset creation, model

selection and evaluation are performed individually for each approach.

4. Materials and methods

4.1 Data Collection

To address the considered use case of frame-based automatic birth detection, an interdisciplinary data collection workflow has been defined to obtain the necessary data basis for dataset creation, model training and evaluation. Within the DigiSchwein project [Ga21], camera recordings of individual birth processes of sows are recorded and stored on a data platform. So far, experiments were conducted between May and October 2021 at the agricultural research farm for pig husbandry of the Chamber of Agriculture Lower Saxony in Wehnen. Of the planned 96 farrowing processes, 26 have already been conducted and recorded. So far, six of these farrowings were analyzed. To expand the database, four more farrowing processes have been added that were recorded within the InnoPig project at the agriculture research farm Futterkamp (Chamber of Agriculture in Schleswig-Holstein, Germany). Each farrowing pen was equipped with a commercial camera system (AXIS M3024-LVE Network Camera) which was installed in top view above the rear part of the sow. Each sow was recorded during the entire farrowing and lactation period. The videos from the DigiSchwein project were recorded at 10 fps, while the from the Futterkamp research farm have 5 fps. Both video recordings have a resolution of 1280x800. These recordings were analyzed by animal scientists to provide time stamps indicating the points in the video at which a birth event occurred. The evaluation of the video files was performed by using the open source Behavioral Observation Research Interactive Software (BORIS) [Fr16]. The starting point of the continuous observation was an hour before the beginning of each farrowing process which was defined as the birth of the first piglet of a litter. Video observation stopped at the end of the post-partum phase. The annotated time stamps in the BORIS software contain the start and end points of individual birth events, which are then used to extract the corresponding frames from the video recordings in which the respective birth event was identified. Each frame was then manually reviewed to determine if a birth event could be detected so that farrowing events could be accurately described on a frame-by-frame basis. A frame was annotated as soon as the content matched the following criteria:

- Visibility of parts of the newborn piglet in the area of the sow's vulva.
- Visibility of the expansion or extension of the sow's vulva.
- A combination of both criteria.

Fig. 2 shows an example of each of these criteria. In total, a number of 176 single birth events were extracted from the collected recordings, which were subsequently used for data analyses, exploration, and preparation.

Fig. 2: Example of birth definitions

4.2 Data Exploration

Exploratory data analysis was conducted to extract specific indicators and key metrics about the respective farrowings. The duration of individual piglet births as well as the time intervals between the birth of two successive piglets were extracted and examined as relevant indicators and are presented in Fig. 3. The results show that a substantial proportion of piglet birth durations are within the range of one to five seconds. Frame-by-frame analysis of birth events has also revealed that birth events are less than one second long. At 5 or 10 fps, this would mean that the shortest observed birth event of 0.8 seconds is 4 or 8 consecutive frames. At the same time, there are also significant outliers when considering the duration of individual births. Cases were identified in which a single birth event birth event was up to 24 seconds long. Conversely, this means that these partial birth sequences are up to 120 and 240 frames long at 5 and 10 fps, respectively. The same applies to the time intervals between individual birth events in the overall farrowing process. It can be observed that most of the intervals are between 10 and 20 minutes long, while there are also exceptions in which the intervals are up to two hundred minutes long. This information can be used to define specific thresholds at which the farmer could be informed about, for example, delayed subsequent births or similar complications. Within the DigiSchwein project, further farrowing processes will be analyzed and examined so that this database can be steadily expanded over time.

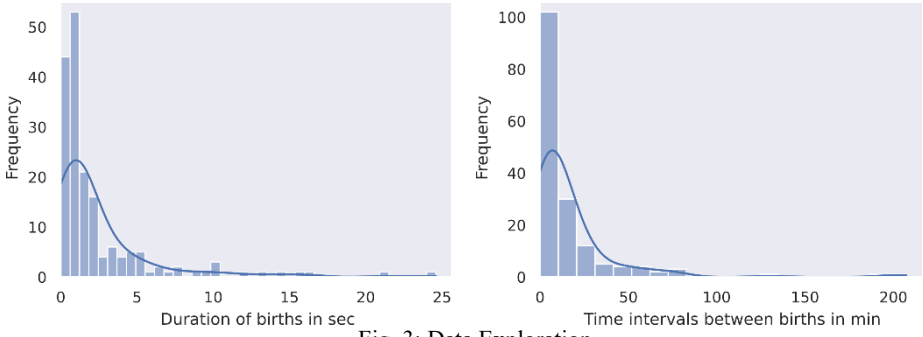


Fig. 3: Data Exploration

4.3 Model Selection

The model selection for the respective approach was conducted by defined selection criteria. These criteria are based both on models and architectures that were already used in literature as well as on the requirements for PLF systems that have been mentioned in the PLF literature. The following criteria were defined:

- **Prediction accuracy**: The prediction of the respective models should be as accurate as possible [No19].
- **Prediction speed**: Model inference should be in real-time [Le19].
- **Cost efficiency**: The respective models should be as resource efficient as possible to allow a potential deployment to low cost hardware [Ba12].

The website paperswithcode[5] provides an overview of all published real-time object detection architectures and their benchmark results on the COCO test-dev, a popular dataset on which model performance is evaluated and benchmarked. This overview served as a basis for selecting the object detection model as well as the image classification model. The following models and architectures were selected for the image classification and object detection approach:

EfficientNet: EfficientNets are among the top performers in image classification on benchmark datasets such as ImageNet, while being smaller and faster than other architectures such as ResNet or Inception [Ta19]. At the core, EfficientNets are based on a traditional CNN architecture. By applying the introduced compound scaling method to uniformly scale network depth, width and resolution as well as a neural architecture search, different EfficientNet variants were created depending on the selected compound coefficient [Ta19]. In the context of this paper, EfficientNet-B0 was used since is the smallest of the EfficientNet variants and therefore fits the specified criteria.

YOLOv5: Since the YOLO architecture has already been used in the PLF literature for various use cases [Al20a], [Sh21], [Le19], has high performance, an active developer community and also meets the defined criteria as it has a fitting balance between speed, performance and hardware requirements, it was selected as the object detection model for the initial prototyping process. YOLOv5 is the latest instalment of the YOLO architecture, but there is currently no official paper for this version. The latest paper release is YOLOv4 by Bochkovskiy et al. [Bo20], which applies specific methods and concepts summarized under the terms *bag of freebies* and *bag of specials* to improve accuracy and execution speed compared to YOLOv3 and other architectures such as EfficientDet. The comparison of the two official implementations of YOLOv4 [Al20b] and YOLOv5 [Jo21] resulted in the selection of the YOLOv5 implementation, as it was more suitable for the context of this paper.

In addition to EfficientNet for the image classification approach, following architectures

[5] https://paperswithcode.com/sota, last visited: 01.04.2022

have been evaluated as well: ResNet [He15] and SwinTransformer [Li21]. ResNet represents the baseline approach in various PLF related publications, while SwinTransformer, based on the Transformer architecture [Va17] is currently the baseline for various state of the art models in the natural language processing (NLP) as well as the CV domain.

4.4 Dataset and test environment

Overall, a total of 3.216 images were extracted from the 176 individual birth events of the examined farrowing processes, in which a birth event could be detected based on the defined visual criteria in Sec. 4.1. These were used as the foundation to create the training and test dataset for the object detection as well as the image classification approach.

Object detection dataset: In total, all 3.216 images were annotated with bounding boxes. The open source tool Labelme was used to annotate the images for model training and evaluation [Wa16].

Image Classification dataset: For the classification approach, different sampling strategies were conducted for frame selection. The best model performance was achieved by using data sets that were generated using of a hard sampling strategy based on inspired by Shrivastava et al. [Sh16]. Based on the assumption that frames immediately after and before a birth event are more difficult for the model to classify, a total of 3.216 negative examples were included to the image classification dataset in addition to the 3.216 positive examples, with one-third of the 3.216 negative examples representing frames found immediately before and after a birth event. This results in a dataset containing 6.432 images.

Both data sets were split into train and test set using an 80/20 ratio. To ensure that the test set contains only or limited data that the respective model has not yet seen, images of two birth processes were used exclusively for testing purposes and are therefore not included in the training data set. Both data sets were split into train and test set using an 80/20 ratio. To ensure that the test set only contains only data that the respective model has not yet seen, images of two complete birth sequences were used exclusively for testing purposes and are therefore not included in the training set.

Model training was performed on a desktop workstation with two Nvidia RTX 3090 with 24 GB VRAM each, a Threadripper 3960X and 64 GB RAM. For the object detection task of, the YOLOv5 implementation of Jocher et al. [Jo21] was applied. Standard parameters were used for training. The model was trained for 20 epochs with a batch size of 16 and the images were scaled to 640×640. Based on the selection criteria, the smallest checkpoint, YOLOv5s, was used for initial training and to enable transfer learning.

For the image classification task, the PyTorch Image Models [Wi19] framework was applied for model training and testing. Images were resized to 224×224 pixels and model training was set to 20 epochs with a batch size of 64 and a learning rate of 0.0001. Cross

entropy loss was used as the loss function, Adam [Ki14] for the optimizer. Image augmentation was also applied by randomly rotate the image within a given degree, horizontal flipping, RGB-shifts as well as changes in brightness and contrast.

4.5 Challenges and limitations

At the current state of this research, there are several challenges and limitations that may limit the generalizability and transferability of the results of this paper, which will be addressed in this section. The data recording was conducted in several pens, but since the pens are all located at the Wehnen site in the Lower Saxony Chamber of Agriculture and are therefore very similar in structure and visual layout, both data diversity and transferability or generalizability could be limited. This cannot be resolved by adding new training data to the already annotated dataset presented in this paper, unless video recordings from other pens would be added to the dataset. Furthermore, the annotation effort to create the training dataset is very high. First, birth recordings, which are usually several hours long, must be analysed by skilled personnel and birth starts or other important events must be tagged. Then, the individual images extracted from these tagged timestamps must also be annotated manually, which can, depending on how the data should be annotated, take several seconds per image. In this case, either with bounding boxes to create an object detection dataset, or with the respective class to create an image classification dataset. Although the manual labelling effort required to annotate the images with bounding boxes could be reduced by having the previously trained algorithm pre-label the unlabelled data and then manually inspect it, however, the manual inspection of the video recordings will be difficult to substitute.

5. Current results

The results for both the object detection as well as image classification approach are summarized in Tab. 1. So, far, the results show that none of the examined approaches can produce convincing results. In the image classification approach, the best model achieves an F1-score of 67,06% on the test set, which is clearly insufficient for operational usage. The same can be observed with the object detection approach, where an AP of 0.577 and an overall mAP 0.246 of can be achieved. Compared to the precision, the low recall also shows that the model has difficulties in detecting actual positive samples in the test set. Although the EfficientNet has by far the smallest number of parameters compared to the other models, it achieved the best accuracy on the test set with a value of 66.20% in the classification task. Considering the much higher number of parameters, the SwinTransformers perform on average worse compared to the other models. The deficient performance of both the image classification approach and object detection approach can be explained as follows:

- Insufficient data basis.
- Both the image classification approach and the object detection approach are inadequate.
- A combination of both.

Image Classification							
Model	**Inference Time (s)**		**Parameters (Mio.)**	**Accuracy**	**Precision**	**Recall**	**F1-Score**
	GPU	CPU					
ResNet50	0.004	0.035	23.51	64.93 %	71.41 %	63.22 %	67.06 %
EfficientNet-B0	0.007	0.018	4.01	66.20 %	59.72 %	68.61 %	63.86 %
SwinTransformer	0.012	0.121	86.74	64.51 %	64.22 %	64.59 %	64.41 %
Object Detection							
Model	**Inference Time (s)**		**Parameters (Mio.)**	**Precision**	**Recall**	**$AP^{IoU=0.5}$**	**mAP**
	GPU	CPU					
YOLOv5	0.002	-	7.01	0.856	0.463	0.577	0.246

Tab. 1: Overview of results

The former could be confirmed by the fact that all models evaluated in this paper showed signs of overfitting. This could be a signal for an insufficient data basis. The second could be confirmed by the fact that the considered use case is too complex to be solved with these simple approaches. The video analyses have shown that certain behavioral patterns can be recognized in the sow shortly before the expulsion of a piglet, e.g., the flapping of the tail or the stretching of the rear legs. The results so far give reason to believe that these patterns, as well as the associated temporal context, need to be considered in the detection of birth events. In the further research development, these aspects will be further investigated and elaborated. The analysis and evaluation of additional farrowing processes will show whether the problem of insufficient performance is due to the data basis or whether novel approaches must be considered in order to effectively detect birth events based on image data. Consideration of other model architectures to capture temporal context and identified behavioral patterns based on, for example, action recognition models could also be explored in this context.

Bibliography

[Al20a] Alameer, A.; Kyriazakis, I.; Bacardit, J.: Automated recognition of postures and drinking behaviour for the detection of compromised health in pigs. Scientific reports 1/10, S. 13665, 2020.

[Al20b] Alexey et al.: AlexeyAB/darknet: YOLOv4 pre-release. Zenodo, 2020.

[Ba12] Banhazi et al.: Precision Livestock Farming: An international review of scientific and commercial aspects. International Journal of Agricultural and Biological Engineering 3/5, S. 1–9, 2012.

[Be14] Berckmans, D.: Precision livestock farming technologies for welfare management in intensive livestock systems. Revue scientifique et technique (International Office of Epizootics) 1/33, S. 189–196, 2014.

[Bo20] Bochkovskiy, A.; Wang, C.-Y.; Liao, H.-Y. M.: YOLOv4: Optimal Speed and Accuracy of Object Detection, 2020.

[Ca19] Cang, Y.; He, H.; Qiao, Y.: An Intelligent Pig Weights Estimate Method Based on Deep Learning in Sow Stall Environments. IEEE Access 99/7, S. 164867–164875, 2019.

[Ch20] Chen, C. et al.: Recognition of aggressive episodes of pigs based on convolutional neural network and long short-term memory. Computers and Electronics in Agriculture 169, S. 105166, 2020.

[Co18] Cowton, J. et al.: A Combined Deep Learning GRU-Autoencoder for the Early Detection of Respiratory Disease in Pigs Using Multiple Environmental Sensors. Sensors (Basel, Switzerland) 8/18, S. 2521, 2018.

[D'18] D'Eath, R. B. et al.: Automatic early warning of tail biting in pigs: 3D cameras can detect lowered tail posture before an outbreak. PloS one 4/13, e0194524, 2018.

[De20] Deutscher Bauernverband e.V.: Situationsbericht 2020/21. Trends und Fakten zur Landwirtschaft. Kapitel 3 Agrarstruktur.

[Fr16] Friard, O.; Gamba, M.: BORIS a free, versatile open-source event-logging software for video/audio coding and live observations. Methods in Ecology and Evolution 11/7, S. 1325–1330, 2016.

[Ga21] Gandorfer, M. et al. Hrsg.: Informatik in der Land-, Forst- und Ernährungswirtschaft. Fokus: Informations- und Kommunikationstechnologie in kritischen Zeiten Referate der 41. GIL-Jahrestagung, 08.-09. März 2021, Leibniz-Institut für Agrartechnik und Bioökonomie e.V., Potsdam. Gesellschaft für Informatik, Bonn, 2021.

[He15] He, K. et al.: Deep Residual Learning for Image Recognition, 2015.

[Ho95] Holyoake, P. K. et al.: Reducing pig mortality through supervision during the perinatal period. Journal of Animal Science 12/73, S. 3543–3551, 1995.

[Hu19] Huang, M.-H.; Lin, E.-C.; Kuo, Y.-F.: Determining the body condition scores of sows using convolutional neural networks. American Society of Agricultural and Biological Engineers, S. 1, 2019.

[Jo21] Jocher, G. e. a.: ultralytics/yolov5: v5.0 - YOLOv5-P6 1280 models, AWS, Supervise.ly

and YouTube integrations. Zenodo, 2021.

[Ka21] Kasani, P. H. et al.: A computer vision-based approach for behavior recognition of gestating sows fed different fiber levels during high ambient temperature. Journal of Animal Science and Technology 2/63, S. 367–379, 2021.

[Ki14] Kingma, D. P.; Ba, J.: Adam: A Method for Stochastic Optimization, 2014.

[Kü20] Küster, S. et al.: Usage of computer vision analysis for automatic detection of activity changes in sows during final gestation. Computers and Electronics in Agriculture 169, S. 105177, 2020.

[La16] Lao, F. et al.: Automatic recognition of lactating sow behaviors through depth image processing. Computers and Electronics in Agriculture 125, S. 56–62, 2016.

[Le19] Lee, S. et al.: Practical Monitoring of Undergrown Pigs for IoT-Based Large-Scale Smart Farm. IEEE Access 7, S. 173796–173810, 2019.

[Li21] Liu, Z. et al.: Swin Transformer: Hierarchical Vision Transformer Using Shifted Windows: ICCV 2021, 2021.

[No19] Norton, T. et al.: Review: Precision livestock farming: building 'digital representations' to bring the animals closer to the farmer. animal 12/13, S. 3009–3017, 2019.

[Oc15] Oczak, M. et al.: Classification of nest-building behaviour in non-crated farrowing sows on the basis of accelerometer data. Biosystems Engineering 140, S. 48–58, 2015.

[Sh16] Shrivastava, A.; Gupta, A.; Girshick, R.: Training Region-Based Object Detectors with Online Hard Example Mining: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), S. 761–769, 2016.

[Sh21] Shao, H.; Pu, J.; Mu, J.: Pig-Posture Recognition Based on Computer Vision: Dataset and Exploration. Animals an open access journal from MDPI 5/11, 2021.

[Si18] Silapachote, P., Srisuphab, A., Banchongthanakit, W.; A. Srisuphab; W. Banchongthanakit: An Embedded System Device to Monitor Farrowing: 2018 5th International Conference on Advanced Informatics: Concept Theory and Applications (ICAICTA), S. 208–213, 2018.

[Ta19] Tan, M.; Le, Q.: EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. In (Chaudhuri, K.; Salakhutdinov, R. Hrsg.): Proceedings of the 36th International Conference on Machine Learning. PMLR, S. 6105–6114, 2019.

[Va17] Vaswani, A. et al.: Attention Is All You Need, 2017.

[Wa16] Wada, K.: labelme: Image Polygonal Annotation with Python, 2016.

[Wa19] Wang, S.; Su, Z.: Metamorphic Testing for Object Detection Systems, 2019.

[Wa21] Wang, M. et al.: A PCA-based frame selection method for applying CNN and LSTM to classify postural behaviour in sows. Computers and Electronics in Agriculture 189, 2021.

[Wh96] White, K. R.; Anderson, D. M.; Bate, L. A.: Increasing piglet survival through an improved farrowing management protocol. Canadian Journal of Animal Science 4/76, S. 491–495, 1996.

[Wi19] Wightman, R.: PyTorch Image Models. GitHub, 2019.

[Ya18] Yang, A. et al.: Automatic recognition of sow nursing behaviour using deep learning-based segmentation and spatial and temporal features. Biosystems Engineering 175, S. 133–145, 2018.

[Zh18] Zheng, C. et al.: Automatic recognition of lactating sow postures from depth images by deep learning detector. Computers and Electronics in Agriculture 147, S. 51–63, 2018.

[Zh19] Zhang, Y. et al.: Real-time sow behavior detection based on deep learning. Computers and Electronics in Agriculture 163, S. 104884, 2019.

[Zh20a] Zhu, X. et al.: Automatic recognition of lactating sow postures by refined two-stream RGB-D faster R-CNN. Biosystems Engineering 189, S. 116–132, 2020.

[Zh20b] Zheng, C. et al.: Automatic posture change analysis of lactating sows by action localisation and tube optimisation from untrimmed depth videos. Biosystems Engineering 194, S. 227–250, 2020.

Remote sensing data analysis via machine learning for land use estimation in the Greater Thessaloniki Area, Greece

Paraskevas Katsalis[1], Evangelos Bagkis[1] and Kostas Karatzas[1]

Abstract: Remote sensing data have been employed for monitoring the differences in land use over time. This information serves as the basis of any further land-related analysis, modelling and decision making. It requires satellite coverage of an area of interest, in various bands, and intense analysis of the data to correctly identify the different land types and associate them to the geographical reality precisely. In this paper, we collect Sentinel 2, level 1C satellite data to extract spectral indices and utilise them as features for land cover classification. The method is based on the use of machine learning for properly mapping the Greater Thessaloniki Area, engaging the random forest algorithm. Two different classification configurations in terms of target labels are tested for their accuracy. The main goal of the study is to present a pipeline for researchers and practitioners that need to define non-generic classes and classify geographical areas accordingly. Results, evaluated with the confusion matrix, suggest excellent performance on the test set and bring to surface limitations of the approach concerning the lack of proper high-quality data for algorithm training.

Keywords: remote sensing, satellite data, land use, machine learning, Normalized Vegetation Index

Addresses Sustainable Development Goal 13: Climate action

1. Introduction

Accurate and timely land cover (LC) information helps with land use (LU) management and LU monitoring. Identifying vegetative areas can also help detect potential landslides or forest fires [KC22]. Moreover, LC maps can help other monitoring and modelling activities that have dependence on land utilization. For example, air quality modelling is a very complex task given that sources and sinks are transient and depend partly on industrial and human activities and thus, LC data can help relevant models to associate a specific LC type with increased or decreased air quality [Jo22]. In recent years remote sensing has been a key information source for LC identification due to fine spatial resolution and great geographical coverage. The most common applications of remote sensing products aim at the identification of the changes in LU over time, mapping plant health and floods, and forest fire detection [EC17]. Furthermore, there is a plethora of spectral indices that can be derived by combining different optical bands from Sentinel 2 such as the normalized difference vegetation index [XS17]. LC and LU identification is

[1] Aristotle University, Environmental Informatics Research Group, School of Mechanical Engineering, Thessaloniki, Greece, 54124, kkara@auth.gr https://orcid.org/0000-0002-1033-5985

necessary for any spatially oriented environmental management and decision-making process. While the former maps the different types of land types of an area under study (forest, wetland, low vegetation etc.), the latter focuses on the way that various types of land are being utilized by people (i.e. a low vegetation area that is being used for agriculture or as grazing land). It is therefore essential to properly map LC as it is the basis of any further land-related analysis, modelling and decision making in scientific areas like spatial planning, microclimate zone definition, urban heat island identification, urban air quality modelling and others.

LC maps are nowadays derived from remote sensing data with a few meters resolution. Products such as Corine land cover (CLC) and Copernicus global land service are produced every six years and annually respectively. Nowadays, both products are mainly based on Sentinel 2 spectral imagery with spatial resolution close to 10 m. However, Sentinel 2 has a revisit time of approximately 5 days and thus, offers the opportunity to extract LC maps in a much finer temporal scale. For example, localization of construction sites, state of the vegetation, snow coverage and others can be beneficial in air quality modelling as they represent processes related with the production and the deposition of air pollutants [VL18, No11].

In [Ru17], the authors used multi-date spectral images and focused on classifying the vegatation with the support vector machine algorithm. Sampling training data from mutiple dates improved the robustnes of the algorithm. Similar to our study, the authors of [TV20] developed a random forest model for Thessaloniki but they concentrated on the build-up areas inside the metropolitan area. In conjuction, we focus on classifying the greater area of Thessaloniki including water bodies and agricalture areas. The authors of [Wa22] review and compare machine learning (ML) approaches such as convolutional neural networks, self-organizing maps with more traditional techniques such as cellular automata. Furthermore, [Jo19] performed a comprehensive comparison of a variety of deep learning (multilater perceptron, variational autoencoders) as well as ensemble algorithms and concluded that ensembles are more verstile and accurate than deep learning algorithms. [Hu20] concentrated in integrating remote sensing with socioeconomic data for improved land use classification. [CGR19] demostrated that convolutional neural networks can outperform traditional ML algorithms and tree-based ensembles but with increased upfront computational cost. Thus, a clear consensus of the best modeling algorithm cannot be reached as it is evident that the volume and quality of data drastically affect the performance of the models. Therefore, we choose the RF as the modeling algorithm, as it is simple to use, can run in parallel if needed and in many cases outperforms even the most advanced deep learning approaches.

Our end goal is to develop a method that will complement air quality modelling systems and provide with an updated LC map with every revisit of the satellite. We concentrated in employing already defined spectral indices related with air pollution production and deposition mechanisms. Specifically, we calculate the spectral indices associated with soil, vegetation, build-up areas, water and moisture and employ them as features to classify the study area into eight classes with the RF algorithm.

2. Materials and methods

2.1 Study area

Central Macedonia region is populated by 1.564 million people accounting to 15% of the total population of Greece, while Thessaloniki is the biggest city in the region and the second largest of Greece [Of22]. A unique characteristic of the study area is that it has a variety of LC classes consisting of industrial, urban, rural, forest, sea, lake, rivers, and agricultural fields as shown in Fig. 1 (Greater Thessaloniki Area-GTA).

Fig. 1: Area of study (using cropped Stentinel 2 Image) : the Greater Thessalobiki Area

2.2 Satellite Data

Remote sensing data were collected from the Sentinel 2 product. Sentinel 2 is equipped with 12 optical detectors that provide data in 13 bands, covering the visible, near infrared and short-wave infrared part of the spectrum. Its spatial resolution varies between 10 m, 20 m and 60 m, depending on the spectral band. This is an Earth observation satellite developed and operated by the European Space Agency (ESA). Each satellite pass is orthorectified by ESA to fit the same coordinates and is matched pixel by pixel. The geographical area of interest is identified with the aid of Sentinel's tiling grid and identified that the GTA corresponds to tile 34TFl [SE22]. This area is then cropped using a custom shapefile for better analysis of the smaller area of interest. Total dimensions of the study area are 27370 x 38730 meters.

The main and only instrument of the Sentinel 2 satellites is a multispectral instrument that generates optical images [MS22]. Using an external sensor, the assembly estimates the attitude and pointing reference to insure a 20 m pointing accuracy and then the image is taken using an optical configuration based on a Three-Mirror Anastigmat (TMA) telecentric telescope [Ca17].

There are two Sentinel 2 products that can facilitate the modelling. For this application, the Level 1C product was selected. The measurements correspond to the top of the atmosphere conditions and thus, atmospheric correction is avoided. On the other hand, Level 2A data, depict the optical properties of the Earth's surface. Both products have been tested and provided similar results. SNAP (Sentinel Application Platform) is a toolkit created by the European Space Agency for users to analyse and project the data from the Sentinel program satellites. It is a GIS application with a targeted use for those missions [SN22]. SNAP was utilised to read the data, create the main shapefile for cropping the region of interest, and to identify the coordinates and "pixel" values for training.

2.3 Spectral Indices

Apart from the 13 spectral bands taken from the satellite, another 7 spectral indices were incorporated to help the algorithm distinguish between the classes of interest. Therefore, a total of 20 features defines the input of the model. Near infrared (NIR), short-wave infrared (SWIR), RED, GREEN and BLUE are involved in the calculation of the spectral indices defined below.

- **Normalized Difference Vegetation Index (NDVI):** This index quantifies the amount of vegetation in an area by comparing near-infrared (that vegetation reflects) and red light (which vegetation strongly absorbs). NDVI is calculated as follows: NDVI= (NIR-RED)/(NIR+RED) [Pe13].

- **Normalized Difference Water Index (NDWI):** This index extracts water body information using Green and Near Infrared bands. Using this method non-water bodies have low reflectance and water bodies have high. NDWI is calculated as follows: NDWI=(GREEN-NIR)/(GREEN+NIR) [Mc96].

- **Normalized Difference Built-up Index (NDBI):** This Index highlights urban areas that have a greater reflectance in the SWIR spectral range in comparison to the NIR. NDBI is calculated as follows: NDBI=(SWIR-NIR)/(SWIR+NIR) [KC19].

- **Built-Up Index (BU):** This combines NDBI and NDVI, to automatically map built-up areas resulting in a map where only built-up and barren area pixels have a positive value. BU is calculated as follows: BU=NDBI-NDVI [Ch10].

- **Bare Soil Index (BSI) :** This index is used to determine the bare soil tiles in a digital satellite image. It uses 4 spectral bands (SWIR, NIR, RED, BLUE). It is a

normalized index with higher values indicating higher chance of bare soil: BSI=[(SWIR+RED)(NIR+BLUE)]/[(SWIR+RED)+(NIR+BLUE)]. [Di17].

- **Green Chlorophyll Index (GCI):** The use of this index is for estimation of the chlorophyll across a wide range of plant species. Here it is used to help the algorithm as a second index for high vegetation areas GCI=(NIR/GREEN)-1[Gi03].

- **Moisture Stress Index (MSI):** This Index is used for canopy stress analysis using the SWIR over NIR ratio. Higher values indicate greater plant stress while lower show less moisture content. It is calculated as follows: MSI=(SWIR/NIR) [HR89].

The normalized indices and the BSI range between -1 and 1. The range for BU is [-1.307, 0.767], for GCI is [-0.832, 10.99] and for MSI is [0.196, 10].

2.4 Methodology

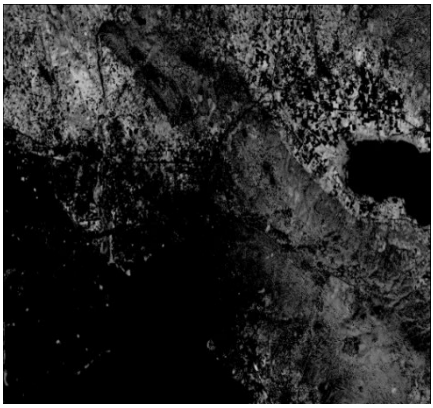
The input of the RF classifier was constructed by concatenating the initial 13 bands with the 7 extracted spectral indices, creating a 20-feature input vector for each "pixel". Eight major classes were identified and included as labels for the classification namely, water, shallow water, concrete, brick, dry soil and crop, wet oil, low vegetation, and tree cover. For the water and shallow water 250 locations were annotated. For concrete and brick classes 574 locations, 1374 for dry soil and crop, 1052 for wet soil and 1120 for low vegetation and 1000 for tree cover with the SNAP tool. The total number of pixels in the study area for the algorithm to classify was 10.600.401. A total number of 6194 locations were manually classified into their respective categories. To overcome the fact that not all bands are in the desired resolution, an up-sampling inter-area interpolation method was used. Thus, we obtained all the channels in the 10-meter resolution.

To move on with the modelling, we selected the RF algorithm as it is robust and accurate. RF [Br01] is an ensemble of decision trees that provides estimations with reduced variance compared to each individual tree. That's because decision trees suffer from overfitting with high variance however, when the estimations are combined with the RF algorithm this variance is reduced leading to improved estimations. Each tree is grown on a randomized subset (random feature selection and bootstrap) of the initial dataset and tries to estimate the same target. During inference, the RF combines the individual estimations via majority voting to produce the final estimation. The number of estimators was set to 100, maximum depth to 5 and all the other parameters were kept to their default values as defined in the scikit-learn python implementation of the algorithm. Finally, the classification map was saved in .tiff format using the metadata of the RED band. For the visualization a custom colormap was created to better represent the natural colours of each category for better visual representation. The performance of the model was evaluated on a total of 619 locations of the manually annotated data. More specifically, a number of 75 locations per class were selected (i.e. 75 x 8 = 700 locations), as well as with 19 additional

key locations, which were selected because of the difficulty that the algorithm had in categorizing them. The 619 locations amount to 10% of the training locations.

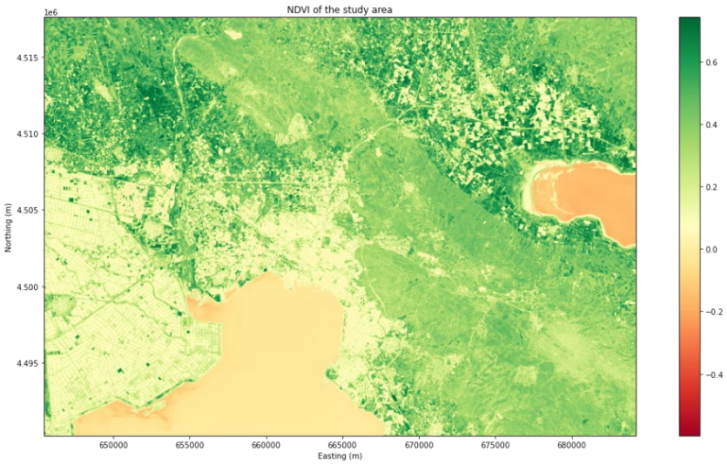
3. Results and discussion

In Fig. 2, three of the calculated spectral indices are depicted. With regards to the green coverage, incorporating the NDVI (panel c) and GCI (panel a) indices light up areas with chlorophyll, help to better differentiate between vegetation and non-vegetated areas. Furthermore, NDWI (panel b) offers valuable information for the model to distinguish between land and water.



(a)

(b)



(c)

Fig. 2: Spectral indices visualization, a) GCI, b) NDWI, c) NDVI

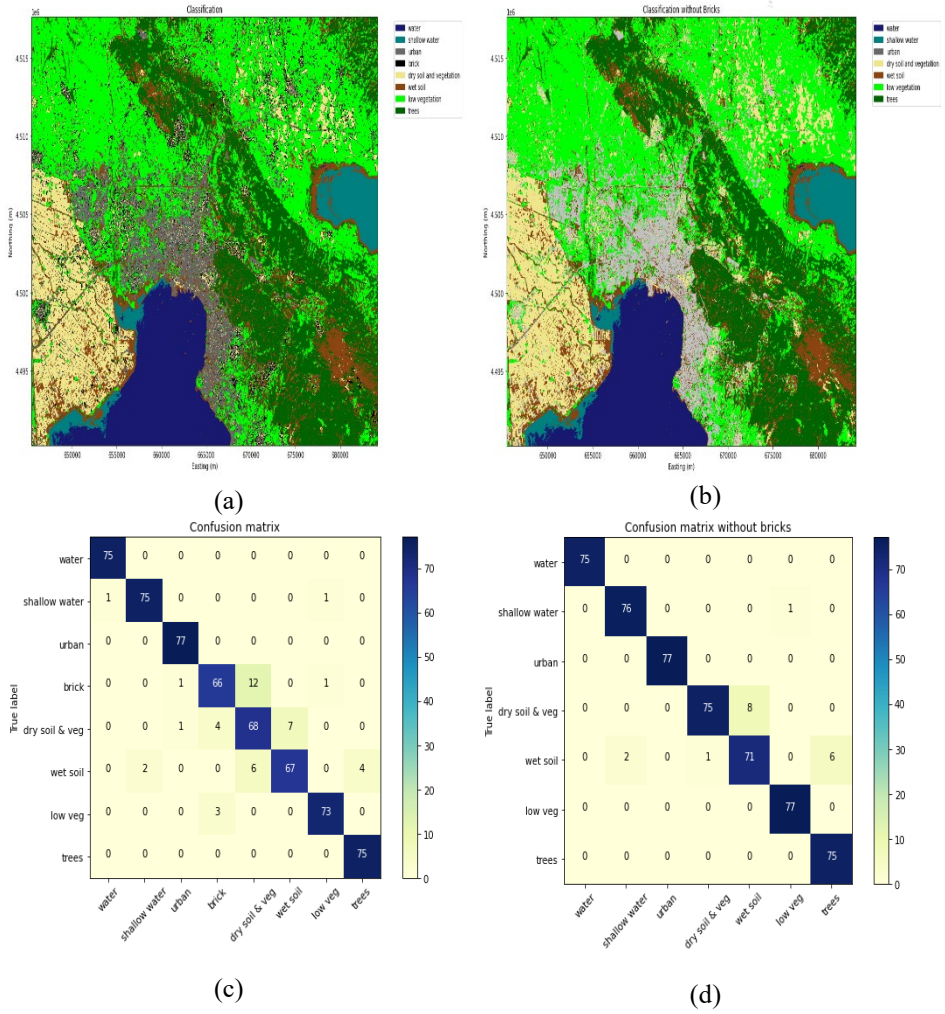


Fig. 3: RF classification results, a) depicts the land-cover map with 8 classes including the Brick class, b) shows the same map but without the Brick class, c) concfusion matrix with Brick class, d) confusion matrix without Brick class.

Overall, the accuracy, calculated on the test set, shows good agreement. However, brick, dry soil, and wet soil categories have very similar reflectance values in almost all the bands and RF struggles to differentiate between them as is evident in Fig. 3 (c). The algorithm can easily identify concrete urban areas due the high reflectance of construction materials and the geometrical shapes that makes these areas stand out. Wet soil areas can be identified around the sores, regarding the city's port on the northern part and the Koroneia

lake on the eastern part. Some wet soil areas can be seen in the mountainous region and are thought to be successfully categorized. Low vegetation fields and dry vegetation fields can be differentiated. Highways even though they have a relatively small width of one to two pixels are easily identified. In the confusion matrix it can be observed that 5 out of 8 classes have almost perfect classification with the problem being in the three classes that have similar reflectance in the spectral bands and indices. When the brick class is removed, the accuracy increases from 93% to 96% but this might lead to the misclassification of the brick covered locations. In future studies we plan to adopt a more robust validation scheme to increase the trust in the results and to compare RF with potentially better algorithms.

4. Conclusions

The main goal of this study was to present a methodology that can help researchers produce LC and LU maps every time the Sentinel 2 revisits a specific area with minimal (>400 locations as a rule of thumb) manual annotation with the aid of ready to use tools. A reproducible machine learning pipeline was presented for LC classification. We proposed to include seven spectral indices that are easily derived from Sentinel 2 L1C as features during modelling to improve the classification error. Furthermore, we identified that brick, dry soil, and wet soil have similar reflectance properties and the misclassifications occur for these classes whereas all other classes were reconstructed almost perfectly. One important aspect related to the increase in accuracy observed after removing the brick class, is that care should be given in the choice of LC classes, and we propose that the classes should be considered with the end goal in mind. It is preferable to identify the classes that play the most important role for the specific problem and build a classifier thereafter than using generic LC maps with irrelevant classes. Therefore, having a machine learning pipeline, like the one proposed here, at hand, can potentially help users create more custom maps for specific use cases. Finally, the proposed method makes use of graphical tools and is low-code, making it easy to work for practitioners in need of accurate and specific user defined classes.

Bibliography

[Br01] Brienman S.: Random Forests. Machine Learning 45, 5–32, 2001.

[CGR19] Carranza-García M.; García-Gutiérrez J.; Riquelme J.: A Framework for Evaluating Land Use and Land Cover Classification Using Convolutional Neural Networks. *Remote Sensing*, *11*(3), 274, 2019.

[Ca17] Cazaubiel V.; Chorvalli V.; Miesch C.: The multispectral instrument of the Sentinel2 program, Proc. SPIE 10566, International Conference on Space Optics — ICSO 2008, 105660H, 21 November 2017.

[Ch10] He, C.; Shi, P.; Xie, D.; Zhao, Y.: Improving the normalized difference built-up index to map urban built-up areas using a semiautomatic segmentation approach. *Remote Sensing Letters*, *1*(4), 213-221, 2010.

[Di17] Diek S.; Fornallaz F.; Schaepman ME.; De Jong R.: Barest Pixel Composite for Agricultural Areas Using Landsat Time Series. Remote Sensing. Remote Sens. 2017, 9, 1245, 2017.

[EC17] Emery W.; Camps A.: Introduction to Satellite Remote Sensing, Chapter 10 - Land Applications p. 701-766 , Elsevier, 2017.

[Gi03] Gitelson, A.; Gritz, Y.; Merzlyak, M.: Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. *Journal Of Plant Physiology*, *160*(3), 271-282, 2003.

[Hu20] Huang Z.; Qi H.; Kang C.; Su Y.; Liu Y.: An Ensemble Learning Approach for Urban Land Use Mapping Based on Remote Sensing Imagery and Social Sensing Data. *Remote Sensing*, *12*(19), 3254, 2020.

[HR89] Hunt J. E.; ROCK, B.: Detection of changes in leaf water content using Near- and Middle-Infrared reflectances. *Remote Sensing Of Environment*, *30*(1), 43-54, 1989

[Jo22] Johansson L.; Karppinen A.; Kurppa M.; Kousa A.; Niemi J.; Kukkonen J.: An operational urban air quality model ENFUSER, based on dispersion modelling and data assimilation, Environmental Modelling and Software, to appear.

[Jo19] Jozdani S.; Johnson B.; Chen D.: Comparing Deep Neural Networks, Ensemble Classifiers, and Support Vector Machine Algorithms for Object-Based Urban Land Use/Land Cover Classification. *Remote Sensing*, *11*(14), 1713, 2019.

[KC22] Kalsnes, B.; Capobianco, V.: Use of Vegetation for Landslide Risk Mitigation. *Springer Climate*, 77-85, 2022.

[KC19] Kuc, G.; Chormanski, J.: SENTINEL-2 IMAGERY FOR MAPPING AND MONITORING IMPERVIOUSNESS IN URBAN AREAS. ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. XLII-1/W2. 43-47. 10.5194/isprs-archives-XLII-1-W2-43-2019, 2019.

[Mc96] McFEETERS, S.: The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *International Journal Of Remote Sensing*, *17*(7), 1425-1432, 1996.

[MS22] MSI instrument Overview https://sentinels.copernicus.eu/web/sentinel/technical-guides/sentinel-2-msi/msi-instrument, last accessed 2022/5/24

[No11] Novotny E.; Bechle M.; Millet D.; Marshall J.: National Satellite-Based Land-Use Regression: NO2 in the United States , Environmental Science & Technology 2011 45(10), p. 4404-4414, 2011.

[Of22] Official General Secretariat of National Statistical Service of Greece Web site https://www.statistics.gr/el/statistics/-/publication/SAM03, last accessed 2022/5/06.

[Pe13] Pettorelli, N.: The normalized difference vegetation index, Oxford University Press, 2013.

[Ru17] Rujoiu-Mare M.; Olariu B.; Mihai B.; Nistor C.; Săvulescu I.: Land cover classification in Romanian Carpathians and Subcarpathians using multi-date Sentinel-2 remote sensing imagery, European Journal of Remote Sensing, 50:1, 496-50, 2017.

[SE22] Sentinel 2 User Guide https://sentinels.copernicus.eu/web/sentinel/user-guides/sentinel-2-msi, last accessed 2022/5/06.

[SN22] SNAP website https://step.esa.int/main/toolboxes/snap/, last accessed 2022/5/06.

[TV20] Tsolakidis I.; Vafiadis M.: Urban land cover mapping, using open satellite data. Case study of the municipality of Thessaloniki. OP Conf. Ser.: Earth Environ. Sci. 410 012062, 2020.

[VL18] Vizcaino P.; Lavalle C.: Development of European NO2 Land Use Regression Model for present and future exposure assessment: Implications for policy analysis, Environmental Pollution, Volume 240, p. 140-154, 2018.

[Wa22] Wang J.; Bretz M.; Dewan M.; Delavar M.: Machine learning in modelling land-use and land cover-change (LULCC): Current status, challenges and prospects. *Science Of The Total Environment*, *822*, 153559, 2022.

[XS17] Xue, J.; Su, B.: Significant Remote Sensing Vegetation Indices: A Review of Developments and Applications, Journal of Sensors , vol 2017, 2017

CIRCULAR ECONOMY AND INDUSTRIAL SYMBIOSYS

Volker Wohlgemuth et al. (Hrsg.): EnviroInfo 2022,
Lecture Notes in Informatics (LNI), Gesellschaft für Informatik, Bonn 2022 167

The application of image recognition methods to improve the performance of waste-to-energy plants

Fenja Schwark[1], Henriette Garmatter[2], Maria Davila[3], Lisa Dawel[4], Alexandra Pehlken[5], Fabian Cyris[6] and Roland Scharf[7]

Abstract: In this paper, we present an image recognition method to improve the performance of waste-to-energy plants. Thermal treatment of waste in waste-to-energy plants is central for the treatment of municipal solid waste. The heterogeneous nature of municipal solid waste results in a fluctuating lower calorific value to which plant operation must be adapted. Compensating for drastic changes in the lower calorific value is challenging for plant operation and can require short-term interventions. Estimating the lower calorific value prior to the combustion process should reduce the number of short-term interventions. In this work, we propose a process-engineering approach to estimate the lower calorific value of waste as a new application of image recognition in waste-to-energy plants. The method is implemented using videos and sensor data from a case study in a real waste-to-energy plant in Germany.

Keywords: waste-to-energy; image recognition; waste properties; process modeling

Addresses Sustainable Development Goal 9: Industry, innovation and infrastructure

1. Introduction

Waste-to-energy (WtE) is the process which uses residual materials as primary energy to generate electricity and/or heat. The key to understand its importance is that it mainly concerns municipal solid waste (MSW), i. e. trash from cities. MSW comes from households, commerce, trade, office buildings, small institutions, garden and street sweepings, among others. The utilization of MSW is not trivial. Ideally, in a circular economy, the amount of recycled material should be maximized and residual waste minimized. Even in an ideal scenario, residual waste could never be completely eliminated. Therefore, policies exist which aim to regulate the utilization. If this is not done adequately, the impacts to human health and environment are harmful. Just to set

[1] OFFIS Institute for Information Technology, Oldenburg, Germany fenja.schwark@offis.de
[2] Institut für Kraftwerkstechnik und Wärmeübertragung, Leibniz Universität, Hannover, Germany garmatter@ikw.uni-hannover.de
[3] OFFIS Institute for Information Technology, Oldenburg, Germany maria.davila@offis.de
[4] OFFIS Institute for Information Technology, Oldenburg, Germany lisa.dawel@offis.de
[5] OFFIS Institute for Information Technology, Oldenburg, Germany alexandra.pehlken@offis.de
[6] EEW Energy from Waste GmbH, Helmstedt, Germany fabian.cyris@eew-energyfromwaste.com
[7] Institut für Kraftwerkstechnik und Wärmeübertragung, Leibniz Universität, Hannover, Germany scharf@ikw.uni-hannover.de

one example, MSW decomposition in landfills produces methane, which, according to the IPCC (Intern-governmental Panel on Climate Change), has a global warming potential 28 times higher than CO_2 [UN16]. For this reason, waste generation and recycling is one of the 231 indicators for monitoring the United Nations' sustainability strategy [DeS20].

The amount of MSW generated is strongly driven by population and urbanization. The problem is that, on the one hand, more people produce more trash and, on the other hand, waste production per capita has increased significantly. From the year 2010 to 2015, annual production of garbage in China, Thailand, Vietnam, India and Pakistan combined grew from 60 Mt to over 300 Mt. It is also estimated that it could more than double during 2015-2025, resulting in over 600 Mt of MSW annually by 2023 [IE19].

As a consequence of the growth, new waste management solutions need to be implemented. Using WtE plants has become essential to provide two solutions: waste treatment and clean energy supply [MH20]. In Germany, WtE plants are well developed offering a very high environmental standard. Current digital technologies have an immense potential to further optimize the performance of the plants, and support the decision-making of the operators.

Artificial intelligence (AI) is widely implemented for waste applications such as waste production forecasting and waste management or classification. Waste classification is used when individual objects can be identified. Ruiz et al. and Chu et al. both present methods using image recognition combined with deep learning to automatically detect objects in waste for recycling [Ch18; Ru19]. Image processing has also been used, but mostly using spectral analysis. In 2013, Vijayakumar et al. implemented infrared spectral analysis to classify the 15 different types of PET bottles [VR13]. Bonifazi et al. used similar methods to propose a fast classification strategy [Bo22]. This, however, cannot be applied in WtE plants since MSW is completely mixed and objects generally cannot be detected.

In addition, machine learning (ML) models have great potential for predicting the thermal properties of waste. Taki et al. used four different types of artificial neural networks (ANNs) to predict the higher heating value of MSW based on the initial materials [TR22]. Genuino et al. used ANNs to investigate the extraction of humic substances from waste during chemical activation [Ge17]. Yet, these studies are based on data sets which include characteristics of the waste, which cannot be identified for MSW in a WtE plant.

Our previously published work has focused on analyzing whether image recognition methods can be used to characterize waste as fuel for WtE plants, using different transfer learning methods and pre-trained datasets [Pe21]. Additionally, ML methods have been implemented to make operationally relevant predictions based on the historical sensor data [Pe22].

The scope of the present research is to introduce a self-developed image recognition method to identify parameters important for the characterization of waste as fuel for WtE plants. In this new approach, image processing is used to determine physical waste properties, and AI is applied to estimate the characteristics of the waste as fuel. In this

way, a proactive component may be added to the combustion control process which will facilitate plant operation. This method's limitations are determined by the heterogeneity of the waste and its properties such as the lower calorific value (LCV), water content or porosity, which strongly influence the combustion process. To provide optimum combustion conditions, as much information as possible should be gained on these properties prior to combusting the waste. In the future, image recognition is to be used to analyze the waste as it is fed to the WtE plant to obtain such information. Such an AI model will only be able to predict the range in which the LCV varies. However, this is sufficient since for an optimized operation of the incinerator it is not the absolute values of the LCV but the changes in LCV that are of importance for the decision process and configuration of the combustion.

The presented method is implemented in a case study, which has been carried out in a WtE plant in Hannover, Germany. The plant is operated by EEW and treats approximately 280000 t MSW per year [EE22]. Fig. 1 shows a schematic overview of the plant. The plant uses forward acting grate firing technology. The waste is temporarily stored in a bunker, where it is mixed for homogenization. The operator uses a crane to charge MSW into a funnel, which then passes onto the grate through. Air is added through the grate for combustion. The hot flue gases from the combustion pass through a steam generator, where the heat is transfered to a water-steam-cycle. The resulting steam is used to generate electricity in a turbine and generator as well as for district heating in the city of Hannover. The gases receive a flue gas treatment, which ensures that all exhaust gases comply to the legally specified emission limits.

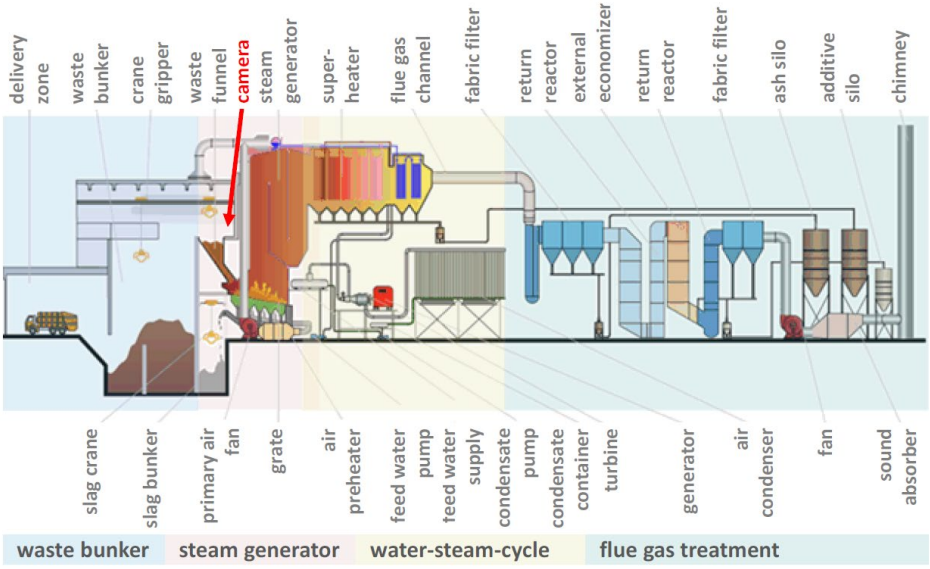


Fig. 1: Schematic layout of a waste-to-energy plant with forward acting grate firing technology [Cy21].

The paper is organized as follows: in section 2 the method developed to estimate physical waste properties with regard to the combustion process by using measurement data is presented. In section 3 the image processing method is shown. Section 4 combines sections 2 and 3, in order to describe how the data architecture of a holistic training set looks like. Finally, section 5 contains the discussion and conclusion.

2. Obtaining physical waste properties with regard to the combustion

Information on the varying waste properties during combustion are required to provide target information for image recognition. It is important that the images of the waste in the training dataset are accompanied by high quality waste property information to ensure the best possible result of the image recognition.

2.1 Lower calorific value

Waste is characterized by different properties. The most important property for combustion processes is the lower calorific value (LCV). It quantifies the energy released during combustion. LCV is a property which may be measured using calorimeters. Such an analysis, however, is unfeasible inside a full-scale WtE plant. Instead of calorimeters, energy balances across the steam generator of the WtE plant are used to derive LCV from measurements obtained inside the steam generator [Ho07]. Hence, the value of LCV is not measured directly but indirectly through temperature, pressure, concentration and mass flow rate measurements.

2.2 Residence time

Using measurements inside the steam generator to derive properties of the waste requires knowledge of the time delay between waste feeding and combustion of the waste on the grate in the steam generator, i. e. the residence time of the waste. The residence time depends on the operating conditions. A varying time delay has to be considered therefore when matching images and measurements.

An experimental investigation of the waste residence times in the feeding system was conducted to obtain information on the time delay [Ga21]. Regression equations were derived which provide the basis to compute the current residence time based on operating parameters of the plant. To estimate for example the residence time in the funnel and chute Δt in min at time t, the velocity of the ram u_{ram} in mm/min may be used according to the equation

$$\widehat{\Delta t}(t) = 97.35 - 0.14 \int_{t}^{t+1h} u_{ram}(\tau)\, d\tau \tag{1}$$

The coefficient of determination R^2 of this regression is 0.32 with a root mean squared error (RMSE) of 8.25 min. The residence times obtained using eq. (1) are due to the comparatively large RMSE not suitable to predict exact values of Δt. However, they may be used to assess the trend of Δt.

Different approaches are possible to use the result of eq. (1) to match the waste images to the computed waste properties. Fig. 2 presents the experimentally obtained residence times with three different methods for residence times quantification.

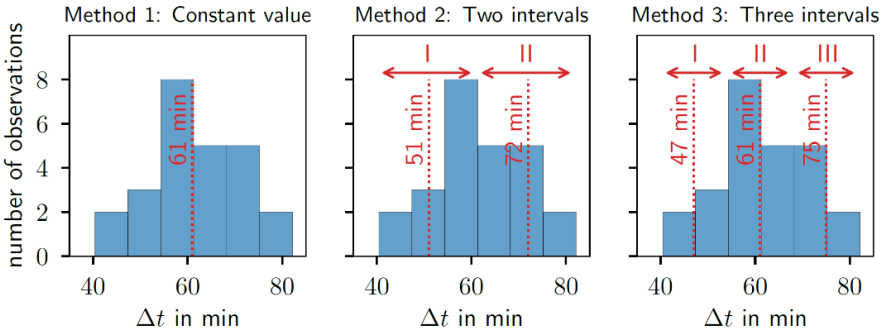


Fig. 2: Different options to assess the residence time.

The first method uses a constant value for Δt. The mean value of all measured residence times, 61 min, is a suitable value. The second method divides the measured residence times into two intervals I and II. Depending on $\widehat{\Delta t}$ computed using eq. (1), the value of Δt is set to either 51 min or 72 min according to

$$\widehat{\Delta t} \leq 61 \rightarrow \Delta t = 51 \text{ min} \quad \text{(interval I)} \tag{2a}$$

$$\widehat{\Delta t} > 61 \rightarrow \Delta t = 72 \text{ min} \quad \text{(interval II)} \tag{2b}$$

The third method in Fig. 2 defines three intervals I, II and II. The value of Δt is set to 47 min, 61 min, or 75 min depending on Δt based on

$$\widehat{\Delta t} \leq 54 \rightarrow \Delta t = 47 \text{ min} \quad \text{(interval I)} \tag{3a}$$

$$54 \leq \widehat{\Delta t} \leq 68 \rightarrow \Delta t = 61 \text{ min} \quad \text{(interval II)} \tag{3b}$$

$$\widehat{\Delta t} > 68 \rightarrow \Delta t = 75 \text{ min} \quad \text{(interval III).} \tag{3c}$$

The most suitable method to match images and waste properties has to be selected in an iterative process. It is expected that the chosen method will vary depending on the examined aspect.

3. Obtaining physical waste properties during waste feeding using image processing

In general, all power plants operate on the same operation principle: generating electricity from a source of primary energy. In the case of WtE plants, the fuel used is waste produced by the citizens. The characterization of the fuel in a power plant, i. e. knowledge of the physical properties of the fuel, is crucial for the power generation. In contrast to other power plants, WtE plants handle a very heterogeneous fuel. While a continuous chemical analysis of the waste could provide detailed information about the material composition in the waste, such a method is not practical at all for in-plant applications. Instead, image analysis is used to gain valuable information about the waste before it is combusted.

3.1 Camera system

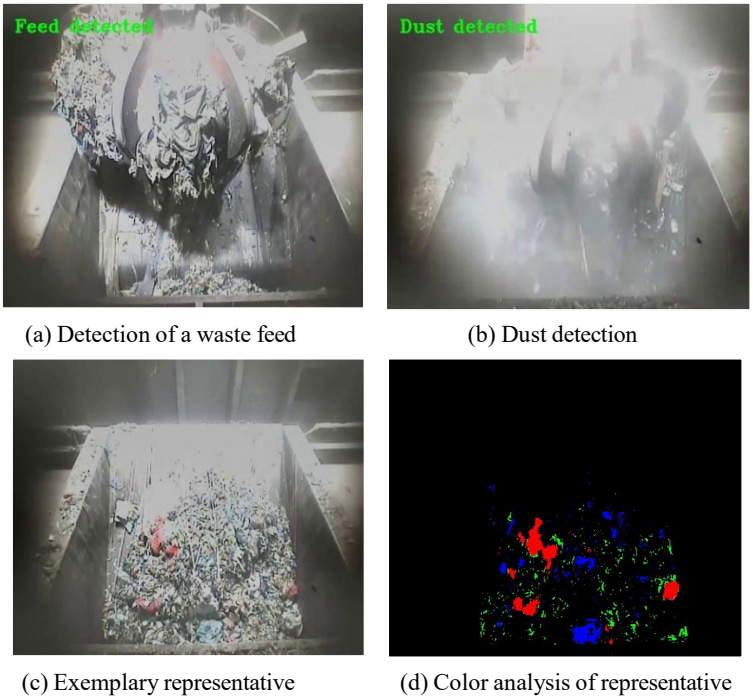
The choice of the cameras is an important aspect in the development of the new technology. Technologies such as multispectral imaging would allow to specify the chemical composition of single objects. This type of technologies are successfully used in applications, where the waste is organized and close to the camera [Bo22; TR22], a scenario not given in a WtE plant. In addition, the waste must be illuminated properly to apply multispectral imaging. Due to safety concerns regarding the flammability of the waste, this requirement cannot be met in the waste bunker. RGB cameras are therefore used to monitor the MSW at the funnel. These cameras are not a hazard and do not affect the plant operation. The location of the camera system in the plant is indicated in Fig. 1.

3.2 Analyze waste feeding

With the camera system installed, it is neither possible to identify individual waste components nor to determine the overall composition of the components, except in a few special cases, e. g. feeds containing large objects such as mattresses or barriers, among others. Yet, experienced crane operators can use visually obtained information to estimate the potential LCV of a feed. Using their expert knowledge, it was possible to identify several physical properties related to the feeds that can be extracted using image processing. These properties include the weight, color, feed duration and properties related to water content.

In general, the procedure of waste feeding can be divided into several phases. In the first phase, the crane operator grabs waste from the waste bunker using the crane gripper and navigates the waste-filled gripper to the funnel. The crane scale can be used to determine the weight of the currently grabbed waste. When the gripper is positioned over the funnel, it is opened and the waste is released. Dust may occur during this second phase. According to the experienced operators, dust formation is an indicator for the water content of the waste. Thus, the dust duration is determined using image processing. After the dust subsides, the waste becomes visible. In the third phase of the feed, the waste gradually

slides into the waste chute. During this phase, more information about the physical properties of the current feed is extracted from the videos. By tracking structures detected in the waste, the movement of the waste can be monitored. In this way, the time it takes for the waste to slide into the chute, i. e. the feed duration, can be determined. In addition, the evolution of the velocity can be analyzed. The velocity at which the waste slides into the chute might provide useful information about the water content of the waste, i. e. moist waste might slide more quickly. At the same time, one must take into account that external circumstances such as the filling level of the chute also affect the velocity at which the waste slides. While the waste slides, representative images of the waste that characterize the feed are extracted. From these representative images, further physical properties to characterize the current feed can be determined. For example, the color of the waste can be analyzed. These properties combined represent the visual cues used by experienced operators to classify the waste as fuel and to estimate its LCV. The decision process of an experienced operator will therefore be simulated by extracting these properties using image processing. In Fig. 3, exemplary frames from a waste feed and the information gained by image processing are shown.



(a) Detection of a waste feed (b) Dust detection

(c) Exemplary representative (d) Color analysis of representative

Fig. 3: Exemplary frames from the videos recorded at the waste funnel. The top panels (a) and (b) show feed and dust detection, respectively. In the bottom panels (c) and (d), an exemplary representative image and the corresponding color analysis as extracted for each feed are shown.

4. Data architecture for image recognition

The methods presented in Section 2 and 3 should be implemented in the WtE plant case study in Hannover, Germany. As the project has not been completed, this is still work inprogress. Fig. 4 illustrates the data architecture that should be used for processing image and sensor data.

Due to the heterogeneity of waste, the LCV in WtE plants varies significantly over time. The WtE plant can operate within a certain range of LCVs. Yet, on the one hand, the lower heating value must be high enough for the waste to inflame, but on the other hand, it must not be so high that the thermal load on the plant's components exceeds the allowed limits. Moreover, compensation of unexpected drastic changes in the LCV is problematic as the effects of the fluctuating LCV must be compensated by intervening in the combustion control, e. g. by changing the combustion air supplied or by introducing auxiliary fuels (light heating oil). Finding relations between the information on the feed gained by the image processing and the calculated LCV would enable to estimate the LCV in the future. In this way, it could be possible to adjust the processing parameters in advanced and reduce the number of short-term interventions.

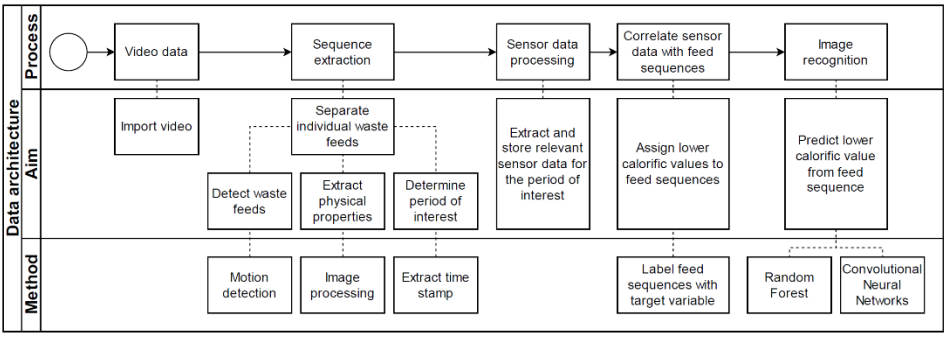


Fig. 4: Flow diagram illustrating the architecture used to relate information gained by image processing to the sensor data.

Within our new process-engineering approach, the information gained by image processing should be related to the LCV observed in the steam generator. Given a video sequence from the camera at the funnel, in a first step the feed sequences within the video are detected and separated from one another. Subsequently, image processing is performed to extract the information characterizing each feed, as described in Section 3. In order to extract the relevant sensor data for each feed, the operating parameters of the plant are analyzed and the residence time of each feed is determined by one of the quantification approaches presented in Section 2. The combustion process of the waste in the steam generator is a continuous process during which the energy of the waste is released. Hence, the LCV of each waste feed is stated as an average value across the combustion period. Accordingly, the sensor data needed to calculate the LCV is extracted from the plants operating

parameters for the period of interest. In a next step, each feed sequence will be labeled with the average LCV as target variable. In this way, a training data set for image recognition will be generated. This data set can be used to train an AI model for predicting the LCV based on the information gained by image processing in the future.

5. Discussion and Conclusion

The presented work exhibits that image recognition of waste is a challenging problem. Waste is a very heterogeneous fuel and cannot be compared for example to coal. Nevertheless, WtE plants are an essential technology to handle the remains of developed societies. Image recognition at the waste funnel may provide a new tool to detect rapidly changing physical properties. Due to the mixed, unordered nature of the waste treated in the WtE plant, multispectral imaging, as known from applications such as waste recycling for analyzing the material composition of single objects, is not applicable for this purpose. However, a process-engineering approach that uses image processing to analyze the physical properties of the feeds can be applied to characterize the waste feeds. A challenge is to relate the properties obtained by image processing to the operating parameters of the WtE plant, since the residence time of the waste is not constant. Therefore, three different options for integrating the residence time into the model are proposed. In addition, the continuity of the combustion process in the steam generator further impedes the assignment of one LCV to each waste feed.

In this work, we presented a new process-engineering approach for integrating image recognition into WtE plant operation. The aim of this approach is to establish an AI model capable of estimating the LCV of the waste such that the number of short-term interventions in the plant operation can be reduced. In a next step, we will develop a demonstrator for applying the data architecture presented in this paper to the real operating WtE plant in Hannover. As the research is still ongoing, there are no final results available. The combination with physical properties and process flows might be an ideal support for the image recognition.

Acknowledgement

This work is supported by the Federal Ministry for Economic Affairs and Climate Action on the basis of a decision by the German Bundestag (FKZ 03EE5038).

Bibliography

[Bo22] Bonifazi, G.; Capobianco, G.; Cucuzza, P.; Serranti, S.; Uzzo, A.: Recycling- oriented characterization of PET waste stream by SWIR hyperspectral imaging and variable selection methods. Detritus/, pp. 42–49, Mar. 2022.

[Ch18] Chu, Y.; Huang, C.; Xie, X.; Tan, B.; Kamal, S.; Xiong, X.: Multilayer Hybrid Deep-Learning Method for Waste Classification and Recycling. Computational Intelligence and Neuroscience 2018/, pp. 1–9, Nov. 2018.

[Cy21] Cyris, F.: Prozessgütemonitoring der Wärmekraftumwandlung von Abfallverbrennungskraftwerken. TEWISS-Technik und Wissen GmbH, 2021.

[DeS20] DeStatis Statistisches Bundesamt - Abfallaufkommen in Deutschland, visited on: 06/02/2020.

[EE22] EEW: Standorte Hannover, 2022, url: https://www.eew-energyfromwaste.com/de/standorte/hannover, visited on: 05/19/2022.

[Ga21] Garmatter, H.; Bauer, L.; Cyris, F.; Scharf, R.: Experimentelle Unter- suchung der Verweilzeit von Siedlungsabfällen im Abfalltrichter und auf dem Vorschubrost eines Abfallverbrennungskraftwerks. In (Beckmann, M.; Hur- tado, A., eds.): Kraftwerkstechnik 2021. SAXONIA Standortentwicklungs- und -verwaltungsgesellschaft mbH, Freiberg, pp. 71–82, 2021, isbn: 978-3-949169-01-4.

[Ge17] Genuino, D. A. D.; Bataller, B. G.; Capareda, S. C.; De Luna, M. D. G.: Appli- cation of artificial neural network in the modeling and optimization of humic acid extraction from municipal solid waste biochar. Journal of environmental chemical engineering 5/4, pp. 4101–4107, 2017, issn: 2213-3437.

[Ho07] Horeni, M.; Beckmann, M.; Fleischmann, H.; Barth, E.: Ermittlung von Betriebsparametern in Abfallverbrennungsanlagen als Voraussetzung für die weitere Optimierung. In (Thomé-Kozmiensky, K. J.; Beckmann, M., eds.): Energie aus Abfall Band 2. Neuruppin, 2007.

[IE19] IEA: Will energy from waste become the key form of bioenergy in Asia? In. Paris, 2019.

[MH20] Mohammadi, M.; Harjunkoski, I.: Performance analysis of waste-to-energy technologies for sustainable energy generation in integrated supply chains. Computers & Chemical Engineering 140/, 2020.

[Pe21] Pehlken, A.; Eschemann, P.; Garmatter, H.; Cyris, F.; Nieße, A.: Einsatz von Künstlicher Intelligenz in der Digitalisierung von Abfallverbrennungskraftwerken. INFORMATIK 2021/, 2021.

[Pe22] Pehlken, A.; Garmatter, H.; Dawel, L.; Cyris, F.; Beck, H.; Schwark, F.; Nieße, A.: How can machine learning improve waste-to-energy plant operation. In. 2022.

[Ru19] Ruiz-Parrado, V.; Sanchez, A.; Vélez, J.; Raducanu, B.: Automatic Image-Based Waste Classification. In. Pp. 422–431, May 2019, isbn: 978-981-10-4912-5.

[TR22] Taki, M.; Rohani, A.: Machine learning models for prediction the Higher Heating Value (HHV) of Municipal Solid Waste (MSW) for waste-to-energy evaluation. Case Studies in Thermal Engineering 31/, p. 101823, 2022, issn: 2214-157X.

[UN16] Intergovernmental Panel on Climate Change (IPCC), visited on: 2016.

[VR13] Vijayakumar, S.; RAJAKUMAR, P. R.: Infrared Spectral Analysis of Waste Pet Samples. International Letters of Chemistry, Physics and Astronomy 4/, pp. 58–65, Sept. 2013.

Optimization paper production through digitalization by developing an assistance system for machine operators including quality forecast: a concept

Moritz Schroth[1], Felix Hake[2], Konstantin Merker[3], Alexander Becher[4]
Tilman Klaeger[5], Robin Huesmann[6], Detlef Eichhorn[7] and Lukas Oehm[8]

Abstract: Nowadays cross-industry ranging challenges include the reduction of greenhouse gas emission and enabling a circular economy. However, the production of paper from waste paper is still a highly resource intensive task, especially in terms of energy consumption. While paper machines produce a lot of data, we have identified a lack of utilization of it and implement a concept using an operator assistance system and state-of-the-art machine learning techniques, e.g., classification, forecasting and alarm flood handling algorithms, to support daily operator tasks. Our main objective is to provide situation-specific knowledge to machine operators utilizing available data. We expect this will result in better adjusted parameters and therefore a lower footprint of the paper machines.

Keywords: Operator assistance, AI, circular economy, paper production, industrial big data

Addresses Sustainable Development Goal 12: Responsible consumption and production

1. Introduction

Germany's paper industry, as an elementary part of the process industry, has achieved great success in sustainability in recent years with the continuous increase in the use of waste paper. This is a big step for achieving circular economy. Paper production only

[1] Fraunhofer Institute for Process Engineering and Packaging, Heidelberger Str. 20, 01159 Dresden, Germany, moritz.schroth@ivv-dd.fraunhofer.de
[2] Consulting Talents AG, Altrottstrasse 31, D-69190 Walldorf, felix.hake@consultingtalents.com
[3] Fraunhofer Institute for Process Engineering and Packaging, Heidelberger Str. 20, 01159 Dresden, Germany, konstantin.merker@ivv-dd.fraunhofer.de, http://orcid.org/0000-0002-3395-6266
[4] University Siegen, Chair of International Production Engineering and Management, Siegener Str. 152, 57223 Kreuztal, alexander.becher@uni-siegen.de
[5] Fraunhofer Institute for Process Engineering and Packaging, Heidelberger Str. 20, 01159 Dresden, Germany, tilman.klaeger@ivv-dd.fraunhofer.de, http://orcid.org/0000-0002-6062-7353
[6] LEIPA Group GmbH, Torgauer Str. 12-15, 10829 Berlin, robin.huesmann@leipa.com
[7] Consulting Talents AG, Altrottstrasse 31, D-69190 Walldorf, detlef.eichhorn@consultingtalents.com
[8] Fraunhofer Institute for Process Engineering and Packaging, Heidelberger Str. 20, 01159 Dresden, Germany, lukas.oehm@ivv-dd.fraunhofer.de, http://orcid.org/0000-0002-4102-8697

using waste paper is a complex process with several production steps using a variety of additives and chemicals as well as the monitoring and adjusting machine parameters due to fluctuating waste paper qualities. The current shortage of recovered paper exacerbates these difficulties.

The process of paper production therefore requires a lot of expert knowledge. Currently operators need a lot of personal experience to produce paper in an adequate quality. But fluctuations in waste paper quality bring major uncertainties to the process stability and also results in fluctuating quality of produced paper. Using process data has great potential for harmonizing paper quality. The operator's communication about daily troubleshooting and the acquired knowledge is currently not systemized. There is a catalogue of actions, but the knowledge is stored analogue and a quick response to machine faults and downtimes is thus not possible. This problem is exacerbated by many relevant parameters to be monitored and changing input material qualities leading to a flood of alarms in the process control system (PCS) to be prioritised by operator.

Paper production is a high resource and energy consuming process. Recovered paper is the most important raw material for the paper industry. Producing one ton of paper additionally requires 17 kg process chemicals, 150 kg other additives, 9.000 l water and 2644 kWh energy. The specific CO_2 emission per ton of paper is 610 kg CO_2 on average [Ve19]. Using LEIPA as an example, process experts estimate that the use of chemicals could be reduced by 3 % and energy requirements by 5 % using intelligent data utilization. This corresponds to annual savings of approx. 750 metric tons of process chemicals, 4,800 metric tons of other additives, and 108,000 MWh of energy.

For optimizing recovery paper cycle this project aims to develop and deploy an AI-based operator assistance system (AS) using linked data from different companies and processes within the paper value chain combined with implicit process expert knowledge. First prototype of operator AS is deployed on one paper machine, to show functionality and benefits in real production. The envisaged solution contributes directly to the United Nations' Sustainable Development Goals (SDG), especially to SDG 9, 11, 12 and 13.

2. State of the art

2.1 Data transfer in circular economy of paper

Concepts and examples for data transfer within the paper industry value chain exist, yet they are not broadly applied. Specific data of samples taken from the paper mother reel is, in some cases, transferred analogously to packaging producers. Process optimization using process-data from paper mills is not applied.

Sorting plants aim to sort waste paper within specification of the EN643 (definitions of paper classes). Data of sorting plants is exclusively used to optimize the sorting process.

Sorting process data includes Near-infrared spectroscopy (NIRS), visual data (VIS) and sometimes object recognition data. Paper mills buy paper according to the EN643, which only vaguely defines included paper fractions. The paper production process has three main leverage options to impact product quality. First is the quality of sorted paper. There are developments to support the manual inspection with NIRS and VIS. Second is the stock preparation with recovered paper being provided and pre-processed. Third option is the paper production machine. Sensor Data is collected in stock preparation and in the paper machine.

Opposing business models hinder data transfer between sorting plants and paper mills. Currently there is no benefit in selling sorted paper above specification. Mutual benefits in data transparency need to be examined. We aim to exchange data between companies and use information from sorting plants to optimize the value chain of paper.

2.2 Quality forecasting systems in paper production

Processes to secure product specifications are often based on laboratory measurements after production of a mother reel. Currently numerous parameters are controlled, and actions are taken when thresholds are exceeded. Selection of supervised parameters and according actions are based on knowledge of process experts. Concepts to predict machine events as i.e., web breaks are offered by companies [Ke21]. How well these systems work cannot be estimated. Prediction of stock preparation parameters using multiple linear regression models based on process and pulp data yielded promising results [Ek20]. Also, the prediction of the Canadian Standard freeness has been studied using convolutional neuronal networks [Ka21]. Neuronal networks have shown a prediction Mean Squared Error of ±8 % when used to predict laboratory quality parameters of produced paper as e.g., Scott bond and tensile strength [Ek20]. Classical mathematical models have been developed to predict relevant quality parameters of recovered paper if the composition of recycled paper is accurately described [Ka21].

Described approaches focus on the papermill as an enclosed dynamic and complex system. Yet the underlying recovered paper quality has an impact on all steps of production in a paper mill and the quality of produced paper, as it is a main resource used in the production process.

2.3 State estimation in machines for error detection and user assistance systems

Looking at a typical production estimating the current (error) state is one of the key factors for providing decision support. Based on the data available various options are possible. Looking at state models and comparing the outputs of the model to the actual process has been used for a long time. Here especially methods like Hidden Markov Models or timed automata have gained interest in detecting malfunctions [BN17, Wi15]. A more black box style approach, or phenomenological approach, as opposed to the former model-based

ones can be used by applying process data directly to some kind of machine learning model with little to heavy pre-processing [Kl19, Le16]. A mixed approach querying ontologies for the process in combination with black box models is proposed in [Ra18].

Looking at data preparation of feature extraction using the process data as is (looking at a time spot) is possible, others tend to build a digital twin of the product while being build [Kl19]. Especially for single, repeating processes analysing the data in other domains like a frequency domain is a common way of data preparation to speedup model training [Ka19].

2.4 Usage of assistance systems in paper production

The idea of user assistance systems for industrial use cases has been coming up in the last years with the thoughts of Industry 4.0 [Re17]. The two major tracks are either physical assistance or mental assistance. This paper focuses on the later. Few of the AS focusing on assist the operator at processing machines in their daily tasks [Ko22, Sc18, Pl17, Ra18]. Approaches here range from using a dialog system over identifying machine states using process data to learning user interaction on the machine panel. Whereas many papers concentrate on one aspect fewer papers show the actual and positive impact of utilizing such systems [Do20, PK22]. It is further known that there are task differences in open systems (e.g. packaging industry) and closed systems (e.g. process industry) [MO19]. The operator tasks in paper production are in many aspects in-between. While the tasks in open systems are characterized by facing a lot of machine stops, in process industry the parameter handling is more important. AS in paper industry has to support both kinds of operator tasks. Current solutions help with different approaches for analysing and controlling the processes, store knowledge and helps to create reports. Examples are Valmet DNA Paper Reel Quality Monitoring and ABB Industrial^{IT} [La08]. Those systems provide a good start but often fail when handling the “automation bias” and a proper user interaction to support the actual task and not disturb with e.g. to many warnings [Ba83, Cu17].

2.5 Alarm flood

In paper production a lot of alarms and warnings arrive in a control room next to the paper machine. Some of these alarms are very important for the operator and require taking action, others are simple information. Sometimes 50 % of the alarms are meaningless for the operator [Mu00]. Alarm floods can result in stress or inattention. There are different approaches to narrow down the number of incoming alarms. One approach to reduce alarm chattering is to group alarms. Another way is to find the root-cause of the alarm flood. It is also possible to find alarm sequences in historical alarm data [Ro16]. Last here mentioned approach for alarm flood filtering in process industry is Signed direct graphs (SDG) [YX12]. Most alarm flood filters are developed for chemical process industry, not for alarm flood handling in paper production. Explorative data analysis on industrial alarm

and warning data combined with visualization has current and futural potential [Be19].

3. Concept of the machine operator assistance system

The mentioned challenges in operator tasks and the resource consuming process require a solid solution. Moisture reduction of 1 % in the predryer section of paper machine is approximated to save 2 to 4 % of the total energy [Ze19]. We focus on paper production in paper circle economy as this step is the most resource consuming. To exploit this potential, we propose an operator assistance system, that helps handling the necessary tasks, even if the employee is not a trained expert. Process knowledge is explicitly stored in the AS. Proper machine handling results in a better product quality, a reduction of web breaks, less reject and thus the required resources.

3.1 Concept overview

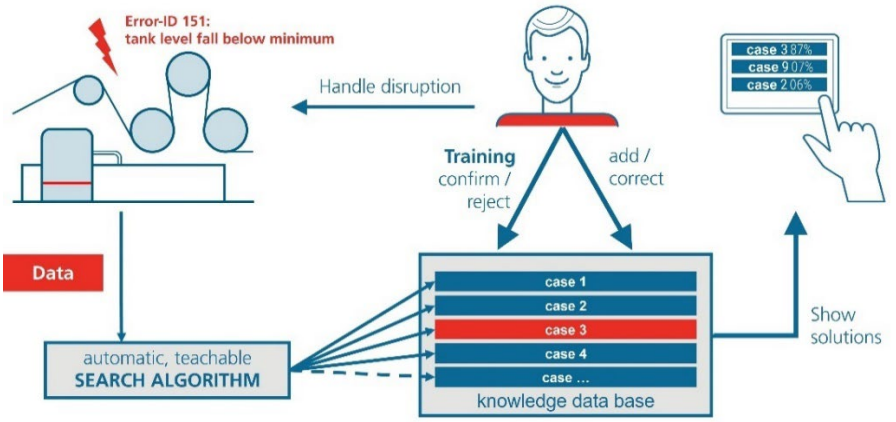


Fig. 1: concept of the assistance system

The assistance system provides information (instructions and procedures, tools required, constraints to be observed, checklists etc.) on known events. To save time-consuming manual searches, available process data is analysed by an intelligent search algorithm using machine learning. The trained model uses alarms and warnings from the process control system (PCS), quality forecast, and sensor data. When an event, relevant for manual examination, is detected, the AS suggests appropriate actions. With user feedback by means of confirmation, rejection, correction and / or supplementation, the AS learns with every event. The action is performed by operators by changing settings or working at the machinery. The concept of the AS is shown in Fig. 1. The implementation will be conducted using a microservice architecture as proposed in [Me22].

3.2 Knowledge base

The implicit process knowledge can be structured into components for different tasks. These components are called "knowledge cards" and are structured into three separate sections: malfunction, solution, and comment section. The malfunction section contains information on the machine error and what caused the situation. The solution section proposes different instructions how to solve the problem. The comment section can be used for further discussion or to propose changes. Because knowledge is linked, the knowledge cards can have causal relationships within the system.

To meet high quality assurance aspects, each change on a knowledge card must be approved by a card editor. Machine operators can only propose changes or discuss them in the comment section. This is critical for the correct operation of the machine as unproper actions based on wrong knowledge can cause significant costs.

3.3 Data sources

Many control systems are generally implemented at paper mills. Distinct control systems handle data differently and it can be a challenge to get access to these systems. In the present scenario MOPS™ [9] is used as a manufacturing information warehouse system. Data is collected from many sources i.e., from the sorting process, the production process, the quality control system and laboratory data [Ek20].

The **sorting data** of the machines for the sorting process is provided via a cloud-based solution TOMRA Insight[10]. A challenge concerning the sorting data is to match it with the data of paper deliveries as there are practical limitations in tracking deliveries causing significant process overhead when implemented. So, machine learning models need to work with a non-deterministic delay from paper delivery to the sorting process.

The **process data** of the production process consists mostly of sensor data. Due to technical aspects, the data sample rate can vary and approximately every 5 to 15 seconds a new sensor value is collected for each sensor.

After a paper reel is produced, samples are taken from the production process and sent to the laboratory. Depending on the priority of a quality parameter the samples will be taken of every mother reel, every n^{th} reel, or irregularly on demand. The **laboratory data** corresponding to one quality parameter has approximately up to 50 values per day.

The PCS raises warnings, alarms, and errors when sensor values breach operator defined thresholds. These **machine events** are used as trigger events to start querying the machine learning models.

[9] https://www.mopssys.com/products/how-we-fit/ last access: 11.07.2022
[10] https://content.tomra.com/insight last access: 11.07.2022

3.4 Classification and forecasting algorithms

Most quality measurements of produced paper are manually taken in a laboratory after production of a mother reel. The machine operator therefore encounters a significant time delay when taking corrective actions to ensure product quality. Thus, forecasting of current paper quality is applied and shows first promising results. Currently a prediction accuracy of ±10 % on data with removed outliners concerning tensile strength has been reached using Random Forest Models, CatBoost and ExtraTrees Regressors. As expected, the model provides better performance when the production is running within the specifications as there are many example values to train the model. The performance decreases when predicting rare values, as for example an especially low or high value of a target variable. One approach to further optimize the overall accuracy is by applying Neuronal Networks. Another focus is to achieve better results with extreme values, therefore the problem is translated to a classification problem, with labels as “low”, “in-specification” and “high”. Using methods for anomaly detection, as the dataset for very low and high values are rare in comparison to the whole dataset, seems a promising option. Within the project good results for the prediction of laboratory measured quality parameters as e.g., bursting strength and Scott bond have been achieved using Random Forest Models. The goal is also to supervise selected parameters and to detect and inform the machine operator when system changes occur.

Presenting the best hint to the user relies on the correct identification of the current situation. Looking in the future (forecasting) the situation based in historic data can be an option. Looking at the current state is another option. As a (mostly discrete) state model is unlikely to work for a continuous process a black box model will be built using algorithms like Random Forest, proven to work in such settings [Kl19]. For the first attempt we propose to analyse a fixed state / the current process image at the very specific moment a user or an alarm triggers the AS using a location identifier to provide knowledge cards. A possible extension is to have the data virtually flow and grow with the paper. In this case faults induced at a certain spot of the paper web causing faults later in the process can internally be tracked. With knowledge of the current production speed, it is therefore possible to build a digital twin of discrete parts of the web. If first approach is precise enough must be proven and heavily depends on the view of the process as a discrete or a continuous process and the time response of errors in the overall system. By using specific data of input material the dosage of chemicals can adaptively adjusted.

3.5 Alarm flood filter

As the PCS of a paper machine generates many alarms, e.g., on simple excess of a single threshold, different methods for alarm flood filter will be compared and brought into action. Alarm notifications can also be reduced by only passing ones with knowledge available in the knowledge base. Here linking via error codes is applied. Another approach is to filter by smart repetition. Important notification should be repeated, but normal or not relevant notifications should not be repeated. Furthermore, patterns in alarm sequences are

analysed and grouped, so that concatenated alarms are summarized in only one event.

3.6 Human machine interfaces

A big issue is the interaction of the AS with the operator. The operator must get all information at the correct moment. Therefore, the development of the graphical user interface (GUI) is implemented according to the Ecological Interface Design [VR92]. For visual simplifying also images and videos are used.

Knowledge to the operator is provided in one of the following situations: a web break, alarms and warnings from the PCS, recognized machine situations based on sensor and process data and predicted quality deviations. Because situation recognition and interpreting are ML-based and therefore probabilistic, the AS can only make recommendations for operators' actions. This is intended, because the AS should neither substitute operator nor inducing the operator to not thinking his- or herself. So, the operator has to make the (final) decision what to do in a certain situation.

4. Summary and further work

In this work we have shown the concept of an operator assistance system for paper machines using modern techniques of the field of machine learning. We detailed how we use forecasting techniques, alarm flood filtering algorithms besides other machine learning models and combine them in an AS for machine operators to provide situational support. This should enable machine operators to apply fixes faster, gain a better understanding of the machine, have a structured knowledge transfer and therefore quicker learning curves. This results in higher machine uptime, better paper quality and a lower resources footprint, so that both environmental and economic aspects are improved.

Currently, the AS is running as a prototype. In the next year, we will continue deploying and adding more features, listen to the feedback of the users and will validate the AS. We see the potential of further work for more creative using of data and quicker implementing of assistance systems in paper production.

Acknowledgment

The underlying project of this contribution was sponsored by the German Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung, BMBF), under the sponsorship reference number 02WDG019D. The authors are responsible for the contents of this contribution.

Bibliography

[Ba83] Bainbridge, L. 1983. Ironies of automation. Proceedings of the IFAC/IFIP/IFORS/IEA Conference, Baden-Baden, 129–135.

[Be19] Bezerra, A. et al. 2019. Extracting Value from Industrial Alarms and Events: A Data-Driven Approach Based on Exploratory Data Analysis. Sensors. 19, 12, 2772. DOI:10.3390/s19122772.

[BN17] Birgelen, A. von and Niggemann, O. 2017. Using self-organizing maps to learn hybrid timed automata in absence of discrete events. 2017 22nd IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), 1–8. DOI:10.1109/ETFA.2017.8247695

[Cu17] Cummings, M.L. 2017. Automation bias in intelligent time critical decision support systems. Decision making in aviation. Routledge. 289–294.

[Do20] Doltsinis, S. et al. 2020. A Decision Support System for rapid ramp-up of industry 4.0 enabled production systems. Computers in Industry. 116, 103190. DOI:10.1016/j.compind.2020.103190.

[Ek20] Ekbåge, D. 2020. Process modelling in pulp and paper manufacture. PhD-Thesis, Karlstad University.

[Ka19] Kammerer, K. et al. 2019. Anomaly Detections for Manufacturing Systems Based on Sensor Data — Insights into Two Challenging Real-World Production Settings. Sensors. 19, 24, DOI:10.3390/s19245370.

[Ka21] Karsimus, M. 2021. Applying machine learning based soft sensors to a pulp refining process. Master Thesis, Aalto University.

[Ke21] Keitel, N. 2021. Voith expands Papermaking 4.0 portfolio.

[Kl19] Klaeger, T. et al. 2019. Using anomaly detection to support classification of fast running packaging processes. 2019 IEEE 17th International Conference on Industrial Informatics (INDIN'19) (Helsinki-Espoo, Finland), 343–348, DOI:10.1109/INDIN41052.2019.8972081

[Ko22] Kott, M. et al. 2022. Methodological approach for the development of an operator assistance system for the press shop. The International Journal of Advanced Manufacturing Technology. 119, 3, 2409–2428. DOI:10.1007/s00170-021-08199-5.

[La08] Laukkanen, I. 2008. Knowledge Transfer and Competence Development in Complex Paper Production Environments. PhD-Thesis, Aalto University.

[Le16] Lei, Y. et al. 2016. An Intelligent Fault Diagnosis Method Using Unsupervised Feature Learning Towards Mechanical Big Data. IEEE Transactions on Industrial Electronics. 63, 5, 3137–3147. DOI:10.1109/TIE.2016.2519325.

[Me22] Merker, K. et al. 2022. A holistic approach for rapid development of IIoT systems. arXiv:2201.13243 [cs, eess]. DOI:10.48550/arXiv.2201.13243.

[MO19] Müller, R. and Oehm, L. 2019. Process industries versus discrete processing: how system characteristics affect operator tasks. Cognition, Technology & Work. 21, 337–

356. DOI:10.1007/s10111-018-0511-1.

[Mu00] Mumaw, R.J. et al. 2000. There Is More to Monitoring a Nuclear Power Plant than Meets the Eye. Human Factors. 42, 1, 36–55. DOI:10.1518/001872000779656651.

[PK22] Psarommatis, F. and Kiritsis, D. 2022. A hybrid Decision Support System for automating decision making in the event of defects in the era of Zero Defect Manufacturing. Journal of Industrial Information Integration. 26, 100263. DOI:10.1016/j.jii.2021.100263.

[Pl17] Plenk, V. et al. 2017. Scoring of Machine-Learning Algorithms for Providing User Guidance in Special Purpose Machines. 10th International Conference on Advances in Human-oriented and Personalized Mechanisms, Technologies, and Services CENTRIC 2017, Athens, Greece, 7–13.

[Ra18] Rahm, J. et al. 2018. KoMMDia: Dialogue-Driven Assistance System for Fault Diagnosis and Correction in Cyber-Physical Production Systems. 2018 IEEE 23rd International Conference on Emerging Technologies and Factory Automation (ETFA), DOI:10.1109/ETFA.2018.8502615

[Re17] Reinhart, G. et al. 2017. Der Mensch in der Produktion von Morgen. Handbuch Industrie 4.0. Carl Hanser Verlag GmbH & Co. KG. 51–88.

[Ro16] Rodrigo, V. et al. 2016. Causal analysis for alarm flood reduction. IFAC-PapersOnLine. 49, 7, DOI:10.1016/j.ifacol.2016.07.269.

[Sc18] Schult, A. et al. 2018. Selbstlernende Bedienerassistenzsysteme in Verarbeitungsmaschinen. Tagung Verarbeitungs- und Verpackungstechnik Dresden, Dresden/ Radebeul.

[Ve19] Verband der deutschen Papierindustrie e. V. ed. 2019. Papier 2019 - Ein Leistungsbericht.

[VR92] Vicente, K.J. and Rasmussen, J. 1992. Ecological interface design: theoretical foundations. IEEE Transactions on Systems, Man, and Cybernetics. 22, 4, 589–606. DOI:10.1109/21.156574.

[Wi15] Windmann, S. et al. 2015. A HMM-based fault detection method for piecewise stationary industrial processes. 2015 IEEE 20th Conference on Emerging Technologies Factory Automation (ETFA), 1–6. DOI: 10.1109/ETFA.2015.7301465

[YX12] Yang, F. and Xiao, D. 2012. Progress in Root Cause and Fault Propagation Analysis of Large-Scale Industrial Processes. Journal of Control Science and Engineering. 2012, e478373. DOI:10.1155/2012/478373.

[Ze19] Zerler, D. et al. 2019. Papier- und Kartonmaschine. Papiermacherschule, Gernsbach.

GREEN CODING

Volker Wohlgemuth et al. (Hrsg.): EnviroInfo 2022,
Lecture Notes in Informatics (LNI), Gesellschaft für Informatik, Bonn 2022 189

Analysis and evaluation of mobile apps with regard to resource efficiency and data volumes

Methodologies and tools

Kira Obergöker[1]

Abstract: The impact of software on energy and resource consumption is receiving more and more attention. While the examination of desktop software already provides initial sresults and criteria for its evaluation, the consideration of mobile apps is not quite as advanced. This paper is a first step to get an overview of which methods and tools can be used to analyse the resource and data consumption of mobile apps and to evaluate their sustainability.
First, I present the previous criteria for desktop software products. In the next step, I present an existing measurement environment for determining the data volume of mobile apps. I created simple environments to identify and test components that can be used to build new measurement environments. I evaluate and compare the measurement environments based on their results. This showed variations between the environments, but an internally equal proportionality. Finally, I used the results obtained to consider how mobile apps can be analysed in terms of their resource consumption, as well as the steps to be taken from now on and the difficulties to be overcome for further investments.

Keywords: mobile apps; resource efficiency; data volumes

Addresses Sustainable Development Goal 12: Responsible consumption and production

1. Introduction

As a result of society's growing environmental awareness, Information Technology (IT) and its resource efficiency are also increasingly coming into focus regarding their environmental footprint. But it is not only the hardware that consumes energy. Information and Communications Technology (ICT) comprises a large number of components, all of which consume resources. This chain of consumption initially includes the manufacture of the hardware. These are end devices, servers that provide data and computation, and networks that transmit the data between the end devices and the servers. But all these components also need to be supplied with energy during use. The amount of this consumption depends on the software running on them. There are already several studies that prove the higher energy consumption of inefficient app programming code [Bu18], [Pa19]. The aim of this

[1] Trier University of Applied Sciences, Umwelt-Campus Birkenfeld, PO-Box 1380, 55761 Birkenfeld, Germany
s18c53@umwelt-campus.de

project is to identify possible criteria with which mobile apps can be evaluated in terms of their sustainability, building on the investigation of the resource consumption of mobile apps.

One of the differences between mobile apps and desktop apps is their increasing usage. The study by "Perficient" gives a comprehensive overview of the use of mobile phones and tablets compared to desktop computers on the internet. It states that "Mobile's share of total visits continues to grow at a steady pace, but desktop devices still have the most total time on site."[2].

In addition, mobile apps are often clients for distributed systems, so here not only the end devices consume resources, but also the network and the servers. This is because most apps store data and process information on servers and this data is transferred back and forth over the network.

One goal could be to ensure that each component consumes as little energy as possible. In the case of the network, this could be achieved by avoiding data transfers that do not serve the app's functionality (Advertisement and Tracking Services (ATS)) [UBS21]. Thus, no advertising and no unnecessary personal data should be sent. This also brings the protection of the user's privacy into the focus of these investigations.

This paper is a summary of my bachelor thesis and examines how the resource consumption of mobile apps can be determined. To this end, I discuss and compare various approaches, identify useful and required tools as well as the functions they need to fulfil. When analysing a mobile app, there are different influencing factors that I take into account. For example, the difference between using an emulator and an end device as well as the influence of automation tools in the measurement setup. I focused on network traffic and elaborated on the potential and feasibility of mobile app evaluation approaches. The questions I will answer are: How can mobile apps be studied? Can mobile app consumption be compared, and if so, how? What options are there to measure consumption and what are their advantages and disadvantages? What needs to be investigated further to enable the evaluation of mobile apps? The focus is thus on the development of a suitable evaluation procedure, which is why the apps considered and the measurement results only play a subordinate role.

There already is some research that aims to operationalise resource efficiency, like the "Blue Angel for resource and energy efficient software products" [RA20] which I used as an example catalogue in this work. It should be noted that the criteria contained therein were developed to assess desktop software. However, since mobile apps are special software for certain types of hardware, the question arose as to the transferability of the criteria to mobile apps. I discuss the extent to which these criteria can be adopted and what adjustments would have to be made in section 4.

[2] https://www.perficient.com/insights/research-hub/mobile-vs-desktop-usage [2022-05-19]]

2. Test execution

2.1 Elements of a test environments

With the aim of determining the resource consumption of mobile apps, I had to create a suitable test environment. I considered using emulators and end devices to execute the app under test. Emulators create virtual mobile devices on the PC, but are partly limited in their possibilities.
A particularly important element was the measuring instrument that I used to collect the desired data. This could either be a mobile app that measures the data directly on the device, a desktop software that has access to the end device, or an external measuring device that is connected to the end device. There were several options that were tested to find the most suitable and practicable solution that would provide reliable values. I had decided to use a mobile measurement app because it is available to everyone and doesn't require much setup. Moreover, they measure directly on the device itself, without detours via the network or other influencing factors. The more accurate solution would probably be to measure the power consumption directly, however this did not fit the scope of the thesis. For this purpose, various mobile apps[3] that output CPU or network traffic were tried out, as well as various macro recorders[4] as automation tools. The automation is necessary because apps have to be measured not only once, but repeatedly in order to get the most reliable results possible and to detect outliers.
So, in the ideal case, I would have had a environment that automatically runs tests based on standard usage scenarios and documents, processes, and graphically displays the results. Already available ad-free mobile apps did not meet the requirements, so my own test environments were built from the available components. Even though these environments are rather improvised and will not represent definitive methods, the measured values can be compared with each other so that conclusive statements can be made.

2.2 General conditions and preparatory measures

As a preparatory measure, I formed app groups that contained mobile apps that should be comparable in principle. Mobile apps with similar complexity and functionality were grouped together. For some groups, such as "weather" and "navigation", I selected four exemplary mobile apps for which I formulated standard usage scenarios. These document

[3] Simple System Monitor https://play.google.com/store/apps/details?id=com.dp.sysmonitor.app&gl=US [2022-02-11]
System Monitor Float Free https://play.google.com/store/apps/details?id=a98apps.monitorfree&hl=en_US&gl=US [2022-02-11]

[4] atbswp http://github.com/rmpr/atbswp [2022-02-16]
JitBit Macro Recorder https://www.jitbit.com/macro-recorder/ [2022-02-16]
Mini Mouse Macro https://sourceforge.net/projects/minimousemacro/ [2022-02-16] Bartels Media GmbH Macro Recorder https://www.macrorecorder.com/de/ [2022-01-16]

the actions of a scenario in detail and thus serve to make the tests comprehensible and repeatable. On the basis of these scenarios, I carried out the test investigations. In general, the selected mobile apps should cover a broad spectrum in order to get an overview of energy consumption and possible influencing factors based on the measurement results. The questions I want to answer with them relate in particular to how the data volumes can be interpreted and analysed and work as a proof of concept.

2.3 Test environments and first impressions

I was able to use the test system of the Institute for Technology and Journalism e.V. (ITUJ)5 to take some measurements. This system records incoming and outgoing network traffic and is based on the "AppChecker"[6] system. It works with smartphones connected to the measuring computer. The data is evaluated with internal databases that contain information about known URLs from owners and third-party providers. At the ITUJ, I measured the selected mobile apps from the navigation and weather groups. The results showed that the traffic within the groups was very different, which could be explained by the range of functions, which varied for each mobile app. The first test runs were performed manually before the system was further developed and automated tests were possible. This allows the repetition of the test runs for statistically relevant data with a smaller amount of work, which makes testing more feasible. Because of this possibility to use both methods, the results can also be used to compare automated and manual tests.

In my own test environments, I examined only the two mobile apps "WarnWetter" from the German Federal Weather Service[7] and "ADAC Spritpreise" from ADAC (General German Automobile Club)[8]. The measurements were intended to compare emulators and end devices, so looking at the test results of two samples was sufficient for a first impression. I chose these randomly from the set of already examined ad-free mobile apps. For all of my test environments I used the "System Monitor Float Free" from 98 Apps, which shows CPU frequency and network traffic. This is an outdated tool, but the only one I could find that does not contain ads and offers the ability to display the measurement data and the measured mobile app on one screen at the same time. This works by allowing the windows that display CPU, network, etc to overlay other mobile apps. Other suitable ad-free mobile apps I found could not do this and did not support the split-screen feature and were therefore not usable. For the automation I used the "Bartels Media GmbH Macro Recorder".
For the end device test, I had to use the Screen Mirroring tool "srcpy"[9] as a supplement to realize the automation with the macro recorder on the computer. The end device was connected to the computer via a USB cable, so there was a secure, direct connection that generated no additional network traffic. Every action, whether entered on the computer or

[5] See https://ituj.de [2022-05-19]
[6] See https://appcheck.mobilsicher.de/ [2022-05-19]
[7] See https://play.google.com/store/apps/details?id=de.dwd.warnapp&hl=de&gl=US [2022-01-13]
[8] See https://play.google.com/store/apps/details?id=com-ptvag.android.adacgasprices [2022-02-11]]
[9] See https://github.com/Genymobile/scrcpy [2022-01-16]

the device, was performed on both screens simultaneously. This also gave me the ability to record the screen on the computer, as I did with the emulator test environment, since I didn't have the ability to write the measured data directly to a document with the measurement mobile app. So after each measurement, I manually extracted the values from the recorded video to analyse them.

3. Results

3.1 Automated vs. manual testing

Using the ITUJ test system, I first ran some manual tests and later retested them together with the ITUJ using the automation tool. However, due to limited time, I initially performed only five test repetitions. My first finding was that most mobile apps that contain advertising cannot be tested automatically, because the ad banners vary. They have different sizes and ways of closing them, and this could not be accounted in the fixed automatic test sequence. Also, this probably leads to a change in consumption within the test runs, rendering the measurements unusable. The same applied to captchas within mobile apps, which always require different inputs. In general, I observed that the values determined for the network traffic within a test series of an app were significantly closer to each other when the automation tool was used. The fluctuations could therefore be reduced, but outliers occurred equally with both test methods. As can be seen in Tab. 1, sometimes the manual tests provide values that are close to the automated test results, but sometimes they provide completely different ones. The same happens between different automated test runs, as can be seen at "wetter.com". With many repetitions, these deviations, which probably occurred due to the small number of test runs, should be compensated. However, these results should be sufficient for a first assessment. Manual tests are therefore quite meaningful, but they are much more labour-intensive if the desired number of 30 tests per run is to be achieved. Nevertheless, mobile apps that are not suitable for automation can be tested and deliver meaningful results.

3.2 Mobile app group comparison

The studies show how much network traffic can vary between mobile apps that are supposed to serve the same purpose, as defined by the app groups such as "weather" or "navigation". I attribute this to the fact that mobile apps have a different range of functions. But, even with the same range of functions, no uniform consumption was found in further investigations by the ITUJ. To what extent this was related to the duration of the standard usage scenarios still needs to be investigated. The ITUJ analysed the collected traffic data and identified the third parties responsible for the ATS traffic based on their URLs. This ATS-share, was also compared between the mobile apps. In terms of evaluating potential mobile applications, I believe it is a reasonable and feasible strategy to evaluate mobile

applications within their groups based on their undesirable traffic. After all, the mobile app providers should not be motivated to remove their useful functionality to reduce traffic.

navigation			
mobile app	**mode**	**Øtotal traffic (MB)**	**ØATS-share**
ADAC Spritpreise	manual	2,032441	3,07%
Komoot	manual	77,4321022	0,13%
Waze	manual	92,0981954	0,07%
OsmAnd	manual automatic	561,55542114 549,15754225	0,00%
weather			
mobile app	**mode**	**Øtotal traffic (MB)**	**ØATS-share**
RegenRadar	manual automatic	73,3119336 36,3866598	5,14%
WarnWetter	manual	0,596894	4,53%
Wetter.com	manual automatic automatic	27,169031 17,7921598 21,1942408	1,17%
Wetter Online	manual	93,7289994	25,11%

Tab. 1: Comparison of the app groups

3.3 Emulator vs. end device test

The advantage of an end device is that all hardware components are available. This is generally not the case with emulators. The emulator "Android Studio" always provides the latest Android version. Thus, there are no costs for hardware devices. However, "Android Studio" is the only emulator providing the recent Android version. For end devices, the update of the operating system either takes longer or is no longer guaranteed after a certain time, which would result in a replacement of the end device. However, it should be noted that there is always the possibility to get the latest Android version in case of the end device, even if it involves costs. It is also problematic that there are hardly any ad-free emulators, so the choice is limited. However, there are dependencies on both sides, either to the emulator developer or to the device manufacturer[10].

For this analysis, I started a test series, with the mobile device "Huawei Mate 10 pro" and an Android 10 version and with the "Android Studio" emulator using the simulated smartphone "Pixel 3" and an Android 10 version. Since the CPU cores were not simulated in the emulator, only the network traffic was compared, which was logged during mobile

[10] See https://www.bluestacks.com/download.html [2022-01-18], https:/memuplay.com/blog/what-is-memu.html [2022-01-18], https://developer.android.com/studio/features [2022-01-18], https://fjsoft.at/en/ [2022-01-18]

app usage and then summed up.
For more detailed results, network traffic values should be logged along with the timing of the scenario performed to capture the duration and timing of a particular transmission. In this way, a specific transmission could be linked to an executed function of the mobile app, which would allow statements about individual functions and, for example, a comparison of the same functions of different mobile apps. I neglected to do this here because the effort involved would not fit the purpose of the study. But for this reason, a monitoring tool that accurately logs traffic is very important.

In the "ADAC" test with the emulator, the network values were very similar within the test runs of a mobile app. When using the end device, the results fluctuated much more and were also generally higher than with the emulator. The results provided by the emulator method were 30% smaller for the "ADAC" mobile app than for the end device method. Among other things, this could be due to the fact that GPS positioning did not work with the emulator and it was therefore not possible to go beyond loading the maps and navigation. With the "WarnWetter" mobile app, there were larger fluctuations in the values within the test runs of a mobile app for both methods. Overall, the determined network traffic through the emulator was 20% lower than with the end device. I could observe that despite the use of an automation tool, the results were subject to large fluctuations for all test methods. To put the relationship between the values obtained in a broader context, I drew on the results of the ITUJ. Tab. 2 and 3 show a comparison of the results.

ØTraffic (MB/s)	**measurement 1 End Device (ITUJ)**	**measurement 2 End Device**	**measurement 3 Emulator**
ADAC Spritpreise	2,032	1,203	0,842
WarnWetter	0,597	0,315	0,257

Tab. 2: Comparison of the test results

	measurement 1 End Device (ITUJ)	**measurement 2 End Device**	**measurement 3 Emulator**
ADAC Spritpreise	100%	ca. 59% of measurement 1	ca. 41% of measurement 1 ca. 70% of measurement 2
WarnWetter	100%	ca. 53% of measurement 1	ca. 43% of measurement 1 ca. 82% of measurement 2

Tab. 3: Percentage distribution of network traffic

The ratio between the results of the emulator test (measurement 3) and the results of the ITUJ test (measurement 1) agreed with just over 40% for both mobile apps. The ratio between the results of the end device test (measurement 2) and the measurements from the ITUJ showed a small difference of about 6% between the mobile apps. Only the difference in the ratio between the results of the end device test and the results of the emulator test was relatively high at about 12%. However, the sample size would have to be increased for further findings. In addition, other influencing factors need to be identified, as the differences between the test environments could have various causes.

In summary, I found that the fluctuations of the values could only be counteracted with a larger number of tests. The results of different test environments will likely always show differences. Therefore, it is important to always use the same test environment when comparing mobile apps. When I compared the two mobile apps tested, the "ADAC" app was more energy intensive in terms of network traffic in each test method. In percentage terms, "WarnWetter" consumed about 30% of the traffic used by the "ADAC" app in all methods. I took this as an indication that the ratio in which the mobile apps could be placed was on target for each method.

4. New evaluation criteria of the Blue Angel criteria to mobile apps

The requirements for desktop software are divided into three areas: Energy and resource efficiency of the software product, hardware useful life and user autonomy [NGK21].
In examining the measurement options, I considered the first area, "energy and resource efficiency of the software product". Furthermore, I examined the quality criteria "hardware useful life" and "user autonomy" of the Blue Angel for desktop applications for their transferability to mobile apps. I consider the criteria transferable, even if they have to be adapted for mobile apps. Especially because of the wide range of mobile apps, the specification is very difficult. Special attention must be paid to health apps, for example, which must not be restricted in a way that could potentially harm patients. Other mobile apps that serve entertainment could likely be restricted to minimise energy consumption. In general, a good balance between functionality and energy saving needs to be found. To take all these cases into account, the criteria must be particularly well thought out. In addition, further criteria are useful. I recommend adding the following criteria as new criteria in the catalog for mobile apps.

Energy and resource consumption
Depending on the influence of the components (end device, network, server) on the overall consumption of a mobile app, these should be weighted differently in the evaluation of a mobile app. The extent to which the servers used should be included as an evaluation criterion at all must be shown by further investigations. However, since there are already some approaches to assessing server consumption [Gr21], it would be interesting to wait for further development of assessment methods for data centers.

Updates
Updates to the operating systems of end devices have a major impact on resource efficiency and service life. Since continuous operating system updates for all devices are rather unrealistic due to the high effort and the large number of different end devices, the most realistic solution is to provide reliable updates of mobile apps that support at least the operating system versions of the last 5 years. The guarantee of updates can be covered by the "continuity" criterion [RA20].

Energy saving management and default settings
Energy saving management could have much more impact, especially if users are aware

of the options to set energy saving settings. This could be achieved if the most energy-saving settings of a mobile app are the default. It is important to use a setup wizard that points out the current settings and allows changes to achieve the best balance between functionality and energy savings for users.

Privacy protection
In order to avoid unnecessary data traffic and protect users, data protection could be included as a criterion. For this purpose, the ITUJ's privacy score could be used, which rates mobile apps in terms of the personal data they transmit to third parties.

5. Summary and outlook

5.1 Summary

This study analysed different measurement strategies. The emulator is more controllable because it can be tested in a closed environment, so unknown influencing factors can be avoided. The end device offers the possibility to test a mobile app under conditions that are closer to real-life use. In general, it is necessary to use the same test environment to compare apps. For more meaningful and better evaluable results, an automation tool should be used. For mobile apps that are compared within app groups, it must be taken into account that the energy consumption also depends on the variety of functions and the complexity of a mobile app as well as the distribution of consumption on the respective client and server side. Due to the diversity of mobile apps, it is difficult to compare them, even with app groups. The ATS share, for example, offers a comparable value in terms of network traffic. In general, it is important to create universally valid and defined standard usage scenarios that provide a comparable test scenario to be able to make comparisons between apps at all.

New possible criteria for the Blue Angel rating of mobile apps take into account the components of the resource consumption chain, which must be weighted differently. An update guarantee should be included to ensure the longest possible service life of the end devices. A powerful tool of the operating system is the internal energy-saving management, which should be supported by the mobile app. In addition, the mobile apps themselves should use energy-saving settings. And even if mobile apps have to collect personal data, because they are medical apps, for example, these must be adequately protected.

5.2 Outlook and conclusion

When evaluating energy and resource consumption, it is necessary to further investigate how the components (end device, network, server) can be weighted and taken into account.

Furthermore, it is also necessary to determine the test conditions that influence the test environment and test execution, such as the duration of the entire standard usage scenarios, as well as the involved functions of a mobile app.

I assume that relative values can be used to make statements about energy and resource consumption. This could be done by comparing energy consumption within groups. How practicable this is is questionable, since there would have to be enough mobile apps for comparison purposes to be able to evaluate a mobile app. So a representative test environment must be determined and a method for evaluating the measurement results must be developed. I see the integration of ratings from other institutions as realistic. Since there are no official solutions to fall back on yet, it remains to be seen whether a procedure for server evaluation will become established in the future. Furthermore, the results of the ITUJ could also be used for the assessment of network traffic. This would mean that the workload for the Blue Angel would only include the measurement of local consumption and the testing of the other quality criteria.

In general, it is important to know exactly which are the most influential components in terms of resource consumption of a mobile app and then define appropriate measurement or assessment methods for them. The assessment must always be practicable so that a high dissemination of the Blue Angel can be achieved. Compromises must therefore be found so that economically operating companies can also fulfil the criteria. These points are currently being further investigated.

Bibliography

[Bu18] Bunse, C.: On the impact of code obfuscation to software energy consumption. In: From Science to Society. Springer, pp. 239–249, 2018.

[Gr21] Gröger, J.; Liu, R.; Stobbe, D. L.; Druschke, J.; Richter, N.: Green Cloud Computing - Lebenszyklusbasierte Datenerhebung zu Umweltwirkungen des Cloud Computing./, 2021.

[NGK21] Naumann, S.; Guldner, A.; Kern, E.: The Eco-label Blue Angel for Soft- ware— Development and Components. In: Advances and New Trends in Envi- ronmental Informatics. Springer, pp. 79–89, 2021.

[Pa19] Palomba, F.; Di Nucci, D.; Panichella, A.; Zaidman, A.; De Lucia, A.: On the impact of code smells on the energy consumption of mobile applications. Information and Software Technology 105/, pp. 43–55, 2019.

[RA20] RAL gGmbH; Umweltbundesamt; Jury Umweltzeichen; Bundesministerium für Umwelt, Naturschutz und nukleare Sicherheit: Blauer Engel Das Umweltze- ichen - Ressourcen- und energieeffiziente Softwareprodukte (DE-UZ 215) - Vergabekriterien./, Jan. 2020, url: https://www.blauer-engel.de/de/produktwelt/ressourcen-und-energieeffiziente-softwareprodukte.

[UBS21] Uijttewaal, M.; Bergsma, G.; Scholten, T.: Carbon footprint of unwanted data-use by

smartphones - An analysis for the EU./, July 2021.

Volker Wohlgemuth et al. (Hrsg.): EnviroInfo 2022,
Lecture Notes in Informatics (LNI), Gesellschaft für Informatik, Bonn 2022 200

Architecting for Sustainability[1]

Patricia Lago[2], Danny Greefhorst[3] and Eoin Woods[4]

Abstract: Sustainability is becoming an increasingly important topic. Information Technology (IT) is an important factor for sustainability; it consumes a substantial, and growing, part of the world supply of energy, but it can also enable significant insights and improvements related to sustainability. These factors need to be taken into account in the design of IT systems, meaning that we need to architect for sustainability. This paper provides insights into the experience and beliefs of IT practitioners and researchers into current and desired practices of architecting for sustainability. It reports on the results of three workshop sessions with practitioners and researchers, providing insight into the state of research and practice.

Keywords: sustainability; software architecture; software design; Information Technology; energy efficiency; maintainability.

Addresses Sustainable Development Goal 12: Responsible consumption and production

1. Introduction

The subject of sustainability, in Information Technology (IT) and software engineering in general, and software architecture in particular, has been receiving increasing attention in both practice and research. In spite of that, the field is still missing knowledge on how to design for software to address the sustainability goals it is meant to help achieve.

As discussed in Andrikopoulos et al. [An22], research in sustainable software engineering has focused on topics from software energy efficiency e. g. [Hi16] to other dimensions as discussed in secondary studies e. g. [Pe12]. This holds for research in the intersection between software architecture and sustainability, too. Some studies [Ve18, An22] provide overviews, while others zoom into specific software architecture topics such as metrics [Ko11], reference architectures [Vo17], technical and economic sustainability in architectural technical debt [VML18], and the implications of software architecture on the social sustainability dimension [Gr21].

In spite of the frequent research efforts, practice is missing consolidated and reusable knowledge that helps attain the sustainable development goals. In this direction, with this

[1] We thank the study participants for sharing their knowledge and experience.
[2] Vrije Universiteit Amsterdam, Netherlands, p.lago@vu.nl
[3] ArchiXL, Netherlands, dgreefhorst@archixl.nl
[4] Endava, UK, eoin.woods@endava.com

work we aim at getting insight into the experience and beliefs related to architecting IT-systems that take sustainability into account. This study is the result of a common activity, in which two workshops were organised to collect data on the subject.

2. Study Design and Execution

The study presented in this paper is illustrated in Fig. 1. After the preliminary Study Design, it is organised in two types of sessions: a Practitioners Workshop and an Academics Workshop, the latter split in two parts for data collection and reflection, respectively. Finally, we analysed the collected data which resulted in a preliminary list of solutions, open problems, impediments, drivers for change, and open research topics/questions, which we used as a taxonomy. These results will hopefully trigger the conversation with interested researchers and practitioners, on effective collaborations to embed sustainability in architecting.

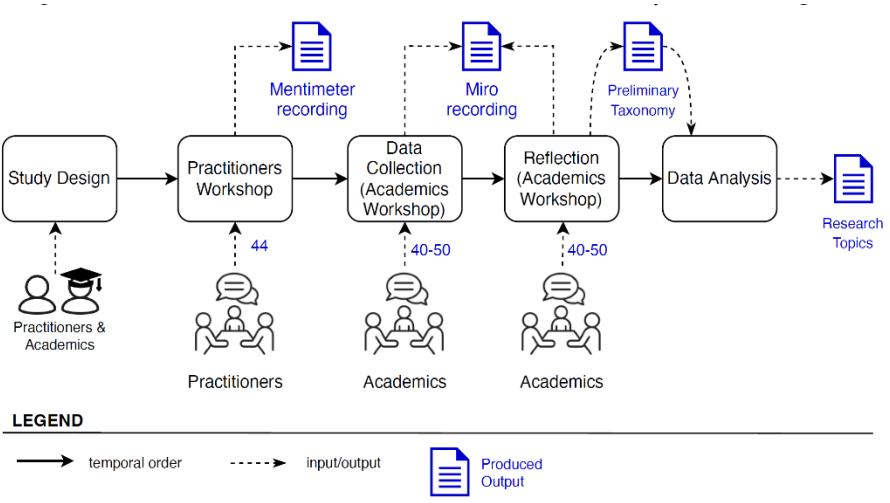


Fig. 1: Overview: Study Design and Execution

2.1 Practitioners Workshop

In January 2021 we have organised a digital workshop with IT-professionals in the Netherlands. There were 44 participants all from different organisations, of which around half of them work as IT-architects or IT-consultants and the remaining in other IT-related professions. The participants are active in a broad range of industry sectors, private and public. The aim of the session was to gain a better understanding of how sustainability can be incorporated into system design. The session was organised in a semi-structured way,

with a collection of predefined questions that the participants answered via Mentimeter[5], with each question followed by a short reflection carried out collectively. We also asked the participants for their ideas about sustainability in general, for themselves as individuals, for their organisation, for their design practice and for their profession. Most questions were open questions, with the possibility to provide multiple answers. Not all participants answered all questions.

2.2 Academics Workshop

In March 2021, we organised a digital workshop in the context of the International Conference on Software Architecture[6] with the aim of engaging the SA research community to reflect on the research directions needed in software architecture to address sustainability. There were 23 participants, all active in software architecture research, practice, or both. The workshop started with a short presentation introducing software sustainability and how software architecture can address sustainability goals. The workshop was then organised in two parts.

The first part of the workshop created a shared understanding of the state of the art, the state of practice and the open research challenges in the intersection between software architecture and sustainability.

In the second part of the workshop we briefly reviewed the results of the first part and then asked the participants identify resources available to assist practitioners with achieving sustainability in their systems.

Both workshop parts were organised in a virtual digital setting, and were very well attended, each counting between 40 and 50 participants.

3. Results

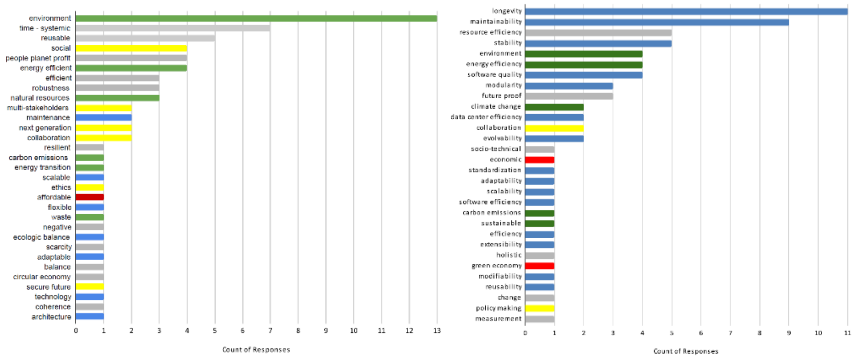
3.1 Results from Practitioners Workshop

Notion of sustainability. The term sustainability does not have the same meaning for everyone. We asked the participants which terms they associate with sustainability and found that associations with ecological goals (green, environmental, planet) and social goals (yellow, social, people) predominate. Economic goals (red, economic, prosperity) are mentioned less often and the association with technology as a goal of sustainability is not often mentioned (blue, technology). The concept of sustainability also evokes associations with robustness, resilience and adaptability (grey, neutral).

[5] www.mentimeter.com
[6] Icsa-conferences.org

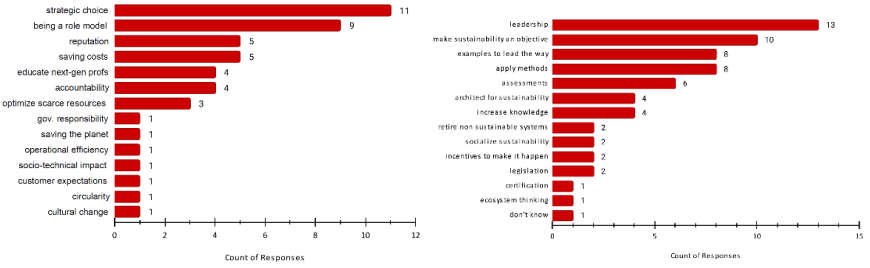
Importance of sustainability. When asked whether it is important to explicitly include sustainability in a design, 29 participants answered positively and 6 negatively. So the importance of sustainability in design was generally recognised by the participants.

Organisational drivers. We then asked the participants what they thought were the main motivations for organisations to be active in the field of sustainability. Their answers are summarised in Fig. 3a. According to the participants, this mainly has to do with strategic choices made by organisations (11 out of 38 responses) and that such organisations want to set an example (9 out of 38).

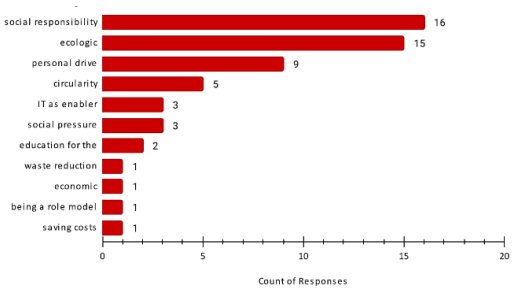


(a) Practitioners Workshop (b) Academics workshop

Fig. 2: Word associated with the notion of sustainability



(a) Organizational drivers: importance for own company (b) personal drivers: importance for oneself



(c) Aids for making sustainability actionable

Fig. 3: Practitioners workshop: specific perspectives

Personal drivers. We asked the participants why sustainability is important for them personally. A classification of their answers (reported in Fig. 3b) shows that a lot of them feel a strong social and ecological responsibility (16 answers). Another class of answers showed a personal drive for the subject (15 answers). The remaining 17 answers were more specific.

Obstacles. We asked what the main obstacles are to embedding sustainability in the design practice of the participants. The answers are summarised in Fig. 4a, which shows that short-term thinking is the biggest obstacle. In addition, lack of knowledge and the financial-economic model of the organisation are explicitly mentioned as obstacles to giving sustainability significant attention.

Making it actionable. We asked how sustainability can be made more actionable and the answers are summarised in Fig. 3c, which shows that the participants believe that more leadership, more examples, concrete sustainability objectives and the application of methods would help. The attention given to the (often abstract) notion that *sustainability is important* can be made more concrete by formulating specific objectives in one's own context. The participants felt that development of working methods would help to give sustainability a place in daily practice.

Incentives. We asked which incentives would provide motivation to make sustainability an integral part of design. The main classes of answers show that goals for sustainability need to be translated into concrete objectives and KPI's. Also the benefits of sustainability need to be made more concrete for organisations and more attention is needed for governance of sustainability.

Closing the gap. We also asked what participants can do themselves to close the gap between designers and sustainability. They indicated that as IT-professionals they can at least explicitly put the subject on the design table. Furthermore, IT-professionals can also show leadership themselves by drawing explicit attention to the subject. Also, the importance of examples was described as helpful for this question.

Role of professional organisations. We finally asked what IT-professional organisations could do to address sustainability in design. According to the participants these organisations can contribute by sharing knowledge, joining forces and communicating about it.

It appears that practitioners are personally motivated to strive for sustainability and see that they can show leadership by incorporating sustainability in design. In contrast, the organisations that they work for seem to be less motivated, often prioritising bottom-line financial results over sustainability. More senior management focus on sustainability is needed, along with concrete objectives for sustainability and governance of the results.

Takeaways Practitioners Workshop:

- Organisational and Personal drivers can accelerate sustainability by complementing each other, e. g. by defining strategies that align with the organisation's goals and motivate employees by leveraging on their personal drivers.
- The business motivation for sustainability needs to get more attention. Many organisations lack a business driver for sustainability, leading to insufficient focus, awareness and understanding, resulting in sustainability not being a priority.
- IT practitioners have trouble translating sustainability to their own work. They feel that sustainability is mostly related to aspects outside of their sphere of influence. They also miss concrete guidance to embed sustainability in IT design.

3.2 Results from academics workshop (Part 1)

The interaction with the participants focused on gathering information about the following aspects.

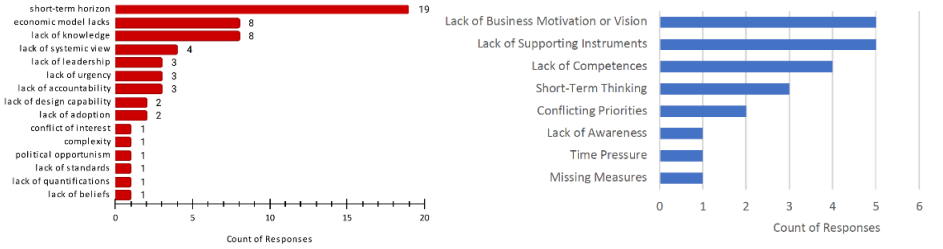
Notion of sustainability. We then asked the participants which words they associate with sustainability. We classified the results as illustrated in Fig. 2b. The top terms were related to longevity, maintainability stability and resource efficiency. We also color-coded the results to relate them to the sustainability aspects we discern: blue for technology, green for environment, yellow for social, red for economic and grey for terms that cannot be clearly positioned in one of the aspects above. The results show that most of the terms are related to technology, followed by environment.

Obstacles. We asked the participants what they think are the main obstacles for companies to embed sustainability in architecture design. We classified the answers, and show the results in Fig. 4b. The top three obstacles that participants identified were lack of supporting instruments, lack of business motivation/vision, short term vision and lack of competence. This can be summarised as: sustainability is not a driver for companies and companies do not know how to operationalise sustainability.

Existing software architecture research for sustainability. We asked what software archi- tecture research already exists for sustainability and the answers are summarised in

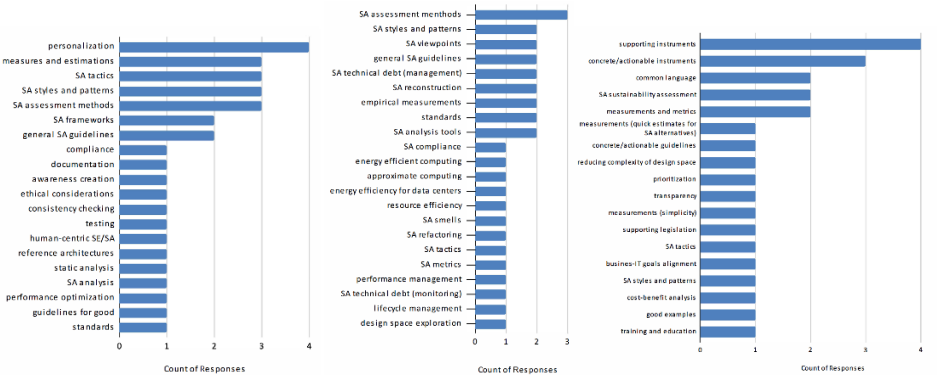
Fig. 5b. The most important type of research identified was related to software architecture assess- ment methods, software architecture viewpoints, general software architecture guidelines, software architecture debt (management), software architecture styles and patterns, software architecture reconstruction, https://www.overleaf.com/project/62b8c8721afe4822090ebc3a empirical measurements, standards and software architecture analysis tools.

Existing software architecture practices for sustainability. We asked the participants which software architecture practices to support sustainability already exist. The results are shown in Fig. 5a. The most important practices are related to collaboration (personalisation mechanisms). The other top classes of answers are measures and estimations, software



(a) Practitioners (Obstacles encountered) (b) Academics workshop

Fig 4: Obstacles for industry to embed sustainability in the design practice



(a) Reported SA practices (b) Reported SA research (c) Reported open problems

Fig. 5: Academics workshop: specific perspectives

architecture tactics e. g. [LL15], software architecture styles and patterns and software architecture assessment methods.

Open problems. We ended the first part of the workshop with a question on the open problems that participants are aware of. The results are summarised in Fig. 5c. The top problem areas that the participants identified were: supporting and concrete/actionable instruments, measurements and metrics, a common language and a software architecture sustainability assessment. The time left for the question was short, so they chose to elaborate on the open problems in the second part of the workshop.

Takeaways Academics Workshop (Part 1):

- The focus of researchers with respect to sustainability in software architecture is mainly on software maintainability and software energy efficiency.
- Two key aspects of embedding sustainability in software architecture are: identify the right measures (patterns, guidelines, frameworks, styles, tactics) and assess the impact of these measures on sustainability.
- Identifying measures and assessing their impact are the overarching themes for future research, too.

3.3 Results from academics workshop (Part 2)

The workshop continued into a second session where we explored solutions to the problems that had been identified in Part 1 and then the drivers for change that are opportunities to improve sustainability in this field and open research questions for future research work.

Solutions. When asked to consider *solutions* to the problems from Part 1, 34 candidate solutions were suggested, which were then classified by the leaders and participants into 7 themes of varying strengths. The top three solutions were: architectural knowledge, implementation analysis and human aspects. Architectural knowledge includes artefacts such as reference architectures, tactics and patterns. Implementation analysis includes various forms of analysis such as static analysis and testing. Human aspects includes aspects such as ethics, compliance and impact on processes and societies.

Open problems. Moving on to *open problems*, the participants listed 48 open problems that they identified as being significant barriers to practice in architecture for sustainability. When these ideas had been classified into themes, the most important of these were: “industrial perspectives and motivation”, “definition” and “techniques and measurements”. Industrial perspectives and motivation includes raising awareness, business prioritisation and cost-benefit analyses. Definition mostly refers to common terminology. Techniques and measurements mostly refers to techniques and metrics for measuring sustainability.

Impediments. The next area that the participants considered was the impediments that they believe exist to implementing possible solutions to the problems facing sustainability in architecture. The participants identified 31 possible impediments to solution implementation and when analysed and sorted into groups, the most important impediments mentioned were “definitions, awareness, competence and business impact and prioritisation”.

The final part of the workshop focused on exploring what the future of this field is likely to be. With this aim, we addressed two questions, (i) What are the drivers for change in the current industrial environment?, and (ii) What are the open research questions which academic researchers should be exploring? The results are summarised below.

Driversforchange. When we asked the participants to identify the likely drivers for change, we identified 19 proposals which resulted in the themes "personal factors", "government action", "industry and academic collaboration", "definition and education" and "awareness and education".

Open research questions. When we moved on to identification of the open research questions, 9 suggestions were made for research topics, which we did not attempt to classify into themes, as we wished to preserve the intent of each suggestion. The suggested research topics/questions are:

- (Assessing the) efficacy of the existing techniques
- Sustainability analysis and assessment framework
- New architectural paradigms (for sustainable architecture)
- What business culture needs to be in place for sustainable software to be successful?
- (Creating a) clear definition of sustainability that we can agree about
- How to make (the) sustainability impact of architectural decisions visible?
- How to trade-off different sustainability dimensions and aspects?
- (Creating) metrics and associated tools for sustainability assessment
- (Creating) a clear definition framework for sustainability of software (should take main- tainability and long-lived software but also the second and third impact into account)

These research questions reveal that there are some fundamental aspects of this topic which still need investigation but that there is also a need for practical guidance and assistance for those trying to improve the sustainability of their software through software architecture.

Takeaways Academics Workshop (Part 2):

- Researchers perceive industry awareness and prioritisation as critical for research to be effective.
- There is a lack of common language on sustainability in software architecture; more (and better) definitions are needed.
- Collaborative research and education are needed to increase and spread the knowledge on sustainable software architecture.

4. Reflections and Conclusions

Through two workshops we have been able to establish an insight into the experience and beliefs of practitioners and researchers into current and desired practices of architecting for sustainability.

From our analysis, we see that both researchers and practitioners consider business motivation and short term thinking as obstacles. Also, both groups agree that key obstacles are the lack of agreement on why sustainability is important, and the absence of knowledge on what concrete measures can be taken.

The lack of concrete measures is a possible bridge between these groups, if they work together on to address it. Solutions and techniques that are available from industry can be assessed by researchers to validate their effectiveness. In turn, researchers can develop new techniques which can be tested in practice.

Research can also provide useful input to create business motivation by providing a clear definition of sustainability, making the impact of sustainability measures explicit and determining the business culture that needs to be in place for sustainable software to be successful. In a more general sense, research can provide analysis and assessment techniques to help organisations achieve their sustainability goals.

From this work we conclude that sustainability requires contributions and commitment on all levels, from senior management to practitioners to research and education.

Bibliography

[An22] Andrikopoulos, Vasilios; Boza, Rares-Dorian; Perales, Carlos; Lago, Patricia: Sustainability in Software Architecture: A Systematic Mapping Study. In: Euromicro Conference Series on Software Engineering and Advanced Applications (SEAA). September 2022.

[Gr21] Grua, Eoin Martino; De Sanctis, Martina; Malavolta, Ivano; Hoogendoorn, Mark; Lago, Patricia: Social Sustainability in the e-Health Domain via Personalized and Self-Adaptive Mobile Apps. In: Software Sustainability, pp. 301–328. Springer, 2021.

[Hi16] Hindle, Abram: Green Software Engineering: The Curse of Methodology. In: 2016 IEEE 23rd International Conference on Software Analysis, Evolution, and Reengineering (SANER). Institute of Electrical and Electronics Engineers (IEEE), pp. 46–55, May 2016.

[Ko11] Koziolek, Heiko: Sustainability evaluation of software architectures: a systematic review. In: Joint ACM SIGSOFT conference on Quality of software architectures–QoSA and architecting critical systems–ISARCS. pp. 3–12, 2011.

[LL15] Lewis, G; Lago, P: Architectural tactics for cyber-foraging: Results of a systematic literature review. The Journal of systems and software, 107(September):158–186, 2015.

[Pe12] Penzenstadler, Birgit; Bauer, Veronika; Calero, Coral; Franch, Xavier: Sustainability in software engineering: A systematic literature review. In: International Conference on Evaluation & Assessment in Software Engineering (EASE). IET, pp. 32–41, 2012.

[Ve18] Venters, Colin C; Capilla, Rafael; Betz, Stefanie; Penzenstadler, Birgit; Crick, Tom; Crouch, Steve; Nakagawa, Elisa Yumi; Becker, Christoph; Carrillo, Carlos: Software sustainability: Research and practice from a software architecture viewpoint. The Journal of Systems and Software, 138:174–188, 2018.

[VML18] Verdecchia, Roberto; Malavolta, Ivano; Lago, Patricia: Architectural technical debt identification: The research landscape. In: 2018 IEEE/ACM International Conference on Technical Debt (TechDebt). IEEE, pp. 11–20, 2018.

[Vo17] Volpato, Tiago; Oliveira, Brauner RN; Garcés, Lina; Capilla, Rafael; Nakagawa, Elisa Yumi: Two perspectives on reference architecture sustainability. In: 11th European Conference on Software Architecture (Companion). Springer, pp. 188–194, 2017.

Conception and test of a measuring station for the analysis of the resource and energy consumption of material flow-oriented environmental management information systems (EMIS)

Dennis Junger[1], Volker Wohlgemuth[2] and Eike Kammer[3]

Abstract: The aim of this work is to design a measurement setting that enables the analysis of the resource and energy consumption of material flow oriented corporate environmental information systems (BUIS). For this purpose, the focus is put on the Life Cycle Assessment (LCA) software. LCA software supports LCA mainly in Life Cycle Inventory (LCI) and in Life Cycle Impact Analysis (LCIA). For the impact assessments, through which potential impacts on the environment are calculated, the LCA software requires many substance-specific life cycle impact data sets, which are stored in databases. In this work, we mainly refer to LCA software that runs in the form of desktop applications. In order to create a measurement configuration suitable for LCA software, the software to be tested and the necessary methodology for desktop solutions are searched. The found solution for sustainable evaluation of software products is adapted to the goal of comparing the efficiency LCA software. The determined measurement standard will be applied, modified and extended. The measurement configuration will be built and tested. Problems of a theoretical nature and problems in setting up the tests are addressed. Any problems of the comparability basis that may arise when performing tests on the selected LCA in the software in the future will be addressed, and a possible solution path will be shown. An outlook on the further course of the measurements is given.

Keywords: Green IT, Energy Consumption, Green Coding, Life Cycle Assessment Software, Desktop Software

Addresses Sustainable Development Goal 12: Responsible consumption and production

1. Introduction

The sustainable design of material and substance flows is inevitable in order to best counter the growing number of environmental crises, such as the hole in the ozone layer,

[1] HTW Berlin, University of Applied Sciences, Industrial Environmental Informatics Unit, Wilhelminenhofstr. 75A, 12459 Berlin, jungerdm@htw-berlin.de, https://orcid.org/0000-0002-9271-2261
[2] HTW Berlin, University of Applied Sciences, Industrial Environmental Informatics Unit, Wilhelminenhofstr. 75A, 12459 Berlin, Volker.Wohlgemuth@HTW-Berlin.de
[3] HTW Berlin, University of Applied Sciences, Electrical Engineering Unit, Wilhelminenhofstr. 75A, 12459 Berlin, eike@kammer-cloud.de

acid rain or global warming. In response to previous environmental crises, the german Enquête Commission adopted the Agreement on the "Protection of Man and the Environment" in 1992. Since then, there has been an obligation for governmental and company actors to carry out material flow analyses. Material flow analyses, which are prepared on a unit or product basis, are product life cycle assessments (LCA) or, more generally, life cycle assessments, which are defined under the ISO standard 14040 and 14044. Such LCAs can be used to derive product environmental declarations and environmental impact assessments (LCIA), such as a carbon footprint. The first production system to be LCA's in Germany was that of beverage packaging in the 1970s. [TU22] The ISO standard 14040 defines the principles and methods of the same, whereas the standard 14044 specifies the steps of the LCA. According to ISO 14040, a typical LCA is composed of four phases: Goal and Scope of Investigation, Life Cycle Inventory (LCI), Impact Assessment (LCIA), and the Interpretation.

The environmental impact of software depends on the hardware used. As described in Fig.1, the hardware basically consumes natural resources, primarily in the form of electrical energy during the time of usage. This is consumed mainly by the components listed under "capacities used", or operating resources. In addition to uninstallation, these are mainly required for the production and use of the software.

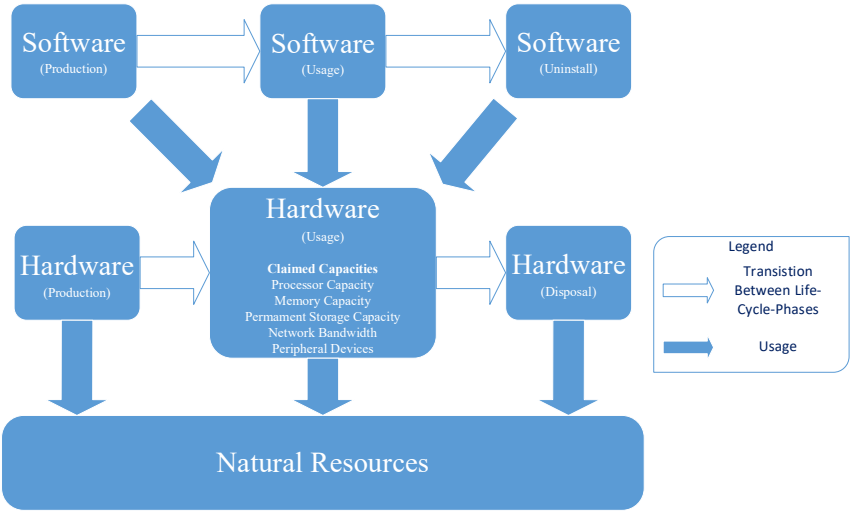


Fig. 1: Life Cycle of Hardware and Software modified [Gr18]

The chosen Software is usually used to create LCAs of products. Here, manufacturing and transport steps as well as their individual impacts of products are quantified as precisely as possible and calculated by means of software solutions that take into account as many circumstances of the entire process as possible with the help of large (impact) databases. The production paths to be balanced, which can be represented as Petri nets, are modeled

and calculated in the software. Thousands of upstream and downstream data sets are usually required to map production paths, and detailed material and energy flows must be described for each one. [LC22] In order to query these LCI data sets in the LCA software in an energy-saving and high-performance way, efficient databases on the one hand and efficient query algorithms on the other hand are required. In addition to data management, data retrieval and data storage, LCA also differs in the algorithms of the mathematical procedures. [BS22] Once LCA models gain in size and number of elements examined, computation time and memory requirements can increase rapidly, with varying resource requirements also noticeable when computing linear than recursive models. Currently available LCA software is implemented in a wide variety of programming and scripting languages and their frameworks. Examples would be OpenLCA in Java [OL22], the Activity Browser with Brightway backend in Python [AB22], and Umberto LCA+ in .NET (C#) [IS22]. The framework chosen can also mean a difference in the efficiency and resource consumption of the software. In the course of this work, a measurement framework and setting is created to test already produced LCA software products in the usage phase of their life cycle. The aim is to get an accurate picture of the energy consumption and power distribution within the hardware in order to measure relevant operating parameters of LCA software and to verify the functionality of the tested in general.

2. Conception of a test setup

2.1 Categorization of the LCA software

To create a suitable test scenario, the software to be tested must be characterized according to its properties. It is categorized according to the location of the data processing and the location of the data storage. A distinction is made between the following four software classes: Local application, remote data storage application, remote processing application, and server service. [BE22] In order to choose the software and the necessary measurement setup correctly, it is necessary to classify the corresponding architecture pattern of the software.

There are currently several different LCA software on the market. In addition to purely web-based software, the main focus in this work is on locally executed software. Characteristic for this type of software is that the user interface, application logic and data management are mainly executed on the local system.

The category chosen for the selected LCA software corresponds here to that of the "local application". The concept of the measuring station is thus based on existing measuring stations for local desktop applications.

2.2 The designed measurement setup

In the measurement setup shown in Fig. 2, a distinction is made between "on-site" and "remote" settings. Here, attention was paid to extending the measurement setup presented in the Blue Angel [BE22] to include a remote measurement method. Especially in the Covid-19 pandemic, access to offices and universities was difficult due to restrictions. In the remote setting, the System Under Test (SUT) and the energy data aggregator are accessed via the control & evaluation station. Here, the Workload Generator (WG), the resource measurement, and the energy data collection are started. It is necessary to start the WG via the control and evaluation station so that there is still an active remote connection to the SUT if the connection to the test network is interrupted. This way the signal resolution of the SUT is constant. Alternatively, the WG and the resource measurement of the SUT can be started manually on site. The energy measurement is started at the control & evaluation station as with the remote measurement. This is currently necessary because of the different operating systems. An NTP server provides the current time for the entire measurement environment and is used for the subsequent assignment of the measurement data in the evaluation. The energy data aggregator queries the energy data of the energy meter via SNMP. These are stored on the aggregator in the form of a log file. After completion of the measurements and data collection, the control & evaluation station evaluates the data using the evaluation software "Open Source Software Consumption Analysis and Reporting" (OSCAR) from Trier University of Applied Sciences and creates an energy report.

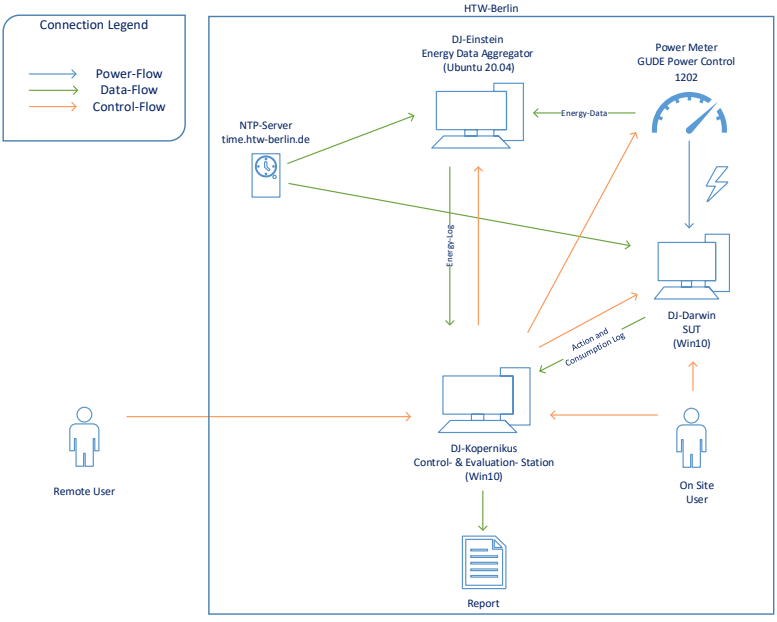


Fig. 2: Information flow chart of the measurement setup

2.3 Problem of comparability of measurements with solution approach

The Blue Angel for software products [BE22] aims to cyclically optimize measurements of a software solution in a continuous improvement process. When different software is compared, the question arises as to what extent the measurement processes are comparable. Tasks and actions to be performed are defined and executed. In order to be able to repeat and compare measurement sequences, they must be recorded and replayed using macros. The same macro cannot be used for different software because different software interfaces, dialogs or operating modes are implemented. In order to be able to play the macros require these after inputs a delay time in form of a pause, so that the software can react to the transacted control instructions. Different numbers of input and output loops now lead to a different number of pauses and thus to a different measurement duration of the software. The length of the measurement pauses is also decisive for the measurement result. The approach of ISO/IEC 14756 [UC22] is that these delays are defined for different usage types and are run through, weighted and averaged by means of a measurement plan. The intersection of the solutions presented is chosen. When building LCA, a very large part of the modeling consists of analyzing and understanding the model being built. This usually takes place with the PC turned on without user interaction and is tied to work time rather than usability. Thus, consistent measurement length is prioritized. Measurements are divided into tasks and actions with the same content and time (e.g. 60 seconds). Furthermore, the total duration of the measurements is defined (e.g. 600 seconds). The advantage is that the human component is not considered. A better measurement result is obtained, with which also conclusions can be made about the efficiency of the programming, about the efficiency of used algorithms, frameworks and databases. In order to avoid measurement outliers, a statistical mean value is formed over 30 measurement processes as specified in [BE22].

2.4 System Under Test (SUT) hardware

The hardware selection is based on the recommendations of the blue angel. In "Development and application of evaluation bases for resource-efficient software taking into account existing methodology", it is described how a reference system for software measurements is created on the basis of the ‘bwPC’, a desktop computer defined once a year for the public service in Baden-Württemberg as the state of the art. Here, the “Öko-Institut e.V.” proposed a SUT for the years 2010 to 2017. In order to select a current SUT for the year 2021, the hardware configuration was based on the hardware noted in the "bwPC Formular v6.2.7" of the computer center of the University of Freiburg at the time of procurement. An almost identical system was ordered from the current HTW-Berlin contract partner Dell.

2.5 Special features of the display unit and the display settings of the SUT

Since the energy consumption of the display can vary greatly depending on the display

technology used, the display device is not measured when designing the measuring station. However, even if the display device is not part of the system to be tested, it has a considerable influence on the measurement. Among other things, the display resolution, signal resolution, color depth and also the repetition rate have an influence on the calculation effort of the display. Since the Blue Angel does not define a display device for software, the test setup assumes that the system under test runs at the standard resolution recommended for Windows 10 for the monitor used. The default refresh rate of 60 hertz and the color depth of 32 bits are selected. The default Windows resolution depends on the screen size and aspect ratio. In order to ensure traceability of future tests, the bwMonitor of the computer center of the University of Freiburg was chosen analogous to the bwPC. This is currently a 24-inch widescreen Full HD monitor (Philips 241B8QJEB) with a 16:9 resolution of 1920 x1080 pixels.

It is recommended to use this standardization also in the Blue Angel for software products.

2.6 The energy data aggregator

As recommended in the Blue Angel, a measuring device is used that aggregates the measured power consumption over one second and can save it together with a time stamp in a log file. [BE22] The "Gude Expert Power Control 1202" was chosen, which is read out via SNMP (here from the energy data aggregator). The script used was obtained from Trier University of Applied Sciences and extended by the functionality of a time synchronization. The current version is provided in [GL22].

2.7 Image of the measuring station

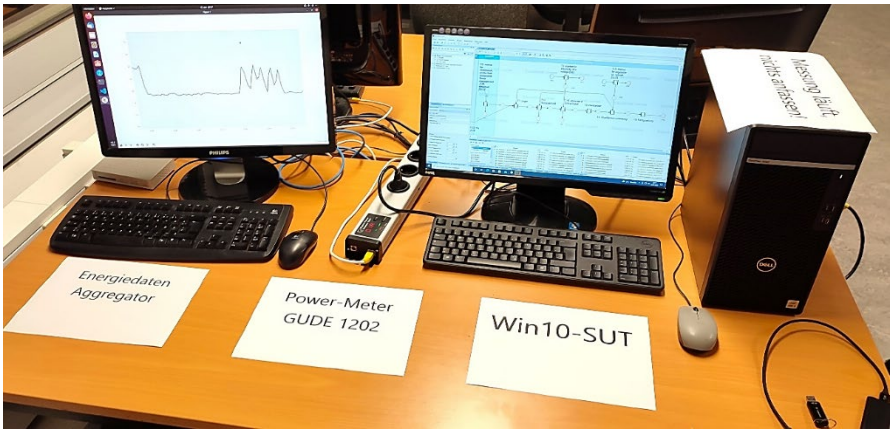


Fig.3: The measuring station with loaded Umberto LCA+ model

2.8 Selection of the SUT equipment devices to be monitored

Apart from the Blue Angel, there is yet no measurement methodology with corresponding measurement standards for carrying out resource- and energy-related software measurement. [BE22] The unique selling point of the Blue Angel software measurement methodology, compared to existing hardware measurements for energy consumption evaluation, is the inclusion of the operating equipment consideration for finding the cause of additional consumption. The choice of resources was based on the evaluation software "OSCAR" version 0.190404 published by the Environmental Campus of the University of Trier (resources: "data carrier physical" (written and read), "processor time", "working memory assured" and "swap file"). The resources are recorded like the energy data in intervals of 1 second each by means of the Windows-own and preinstalled program "performance monitoring" (perfmon.exe). The network traffic is recorded by means of "network adapter" (send and receive) in the measurement instructions of Trier University. In this test setup the "network interface" (send and receive) is also recorded to not only record the physical signal amount, but also signal amounts that take place through connections with "localhost"(loopbacks). Recording facilitates the detection of bottlenecks or energy wastage.

3. Test measurements

In order to test the functionality of the measuring station, test measurements were carried out (600 seconds, 30 repetitions, full screen mode, LCIA data sets according to Tab. 1). In each case, an idle measurement of the listed software and an idle measurement without software (baseline) were made for the evaluation. The LCIA data of Ecoinvent 3.6 was chosen as database basis. Due to the use of different data formats, import difficulties of these as well as also license problems not the same LCIA data sets could be imported into the tested software. By Tab. 1 the following conclusions are obvious:

- OpenLCA requires less power in idle mode than Umberto LCA+ or Brightway2 with Activity Browser frontend
- Software frameworks (.NET, Java, Python) can create differences
- It’s possible that different number of data sets can have an impact on RAM usage.
- Umberto LCA+ required a larger RAM usage and permanent memory usage than OpenLCA and Activity Browser (maybe due different user interface or technological stack)

For further qualitative statements, further measurements according to the measurement plan are necessary.

Test measurement with data sets Idle, On site	**Umberto LCA+**	**Activity-Browser**	**OpenLCA**
Imported LCIA data	Ecoinvent: APOS 3.6 Consequential 3.6 CutOff 3.6	Ecoinvent: APOS 3.6 Consequential 3.8 CutOff 3.6	ELCD 3.2 Agribalyse
Mean el. power	8,24 W	9,53 W	7,91 W
Mean el. work	1,37 Wh	1,59 Wh	1,32 Wh
Average CPU usage	0,55%	4,02%	0,87%
Average RAM usage	40,07%	39,81%	36,73%
Amount of data transferred via network	781,0612269 MByte	444,7875595 Mbyte	736,2177884 MByte
Permanent storage usage	7, 5396711 × 10^5 MByte	1, 9882778 × 10^5 Mbyte	2, 1451082 × 10^5 Mbyte

Tab.1: Preliminary energy and resource comparison Idle on site

4. Conclusion and Outlook

It was possible to identify the most significant criteria for the measurement site. The software to be tested was categorized as desktop software and thus a measurement concept was sought. The Blue Angel for resource and energy efficient software products was found. Based on the setting proposed here, a setup of a measurement station was designed. The possibility of remote measurements was added to the measurement concept. In order to validate this, further measurements are necessary.

Current weaknesses of the SUT configuration of the existing methodology are the non-standardized display settings. Different resolutions, color depths and refresh rates affect the power and resource consumption. Based on the specification of the bwMonitor, defined by the computer center of the University of Freiburg for public authorities, a possible standard could be found. This can now be recommended as an extension of the Blue Angel criteria to couple the screen resolution with the device information and Windows standard settings.

Basic considerations of a measuring place and measuring procedure were conceived. User-independent measurement scenarios with normalized times were chosen to compare existing software in the LCI and LCIA phases. To highlight the differences of the linear and recursive calculation, suitable usage scenarios have to be defined. Commonly used LCIA methods that determine, store and export the potential fossil energy input (KEA_{fossil})

or the potential global warming potential (GWP_{100}) are suitable.

For the exclusion of performance differences in the data set construction, the measurements are performed with LCIA data from the same databases. These should have the same data sets imported. Current LCIA data base is the Ecoinvent 3.6, because they are already integrated in Umberto LCA+ and a license is available at the university. The OpenLCA software does not import the selected LCIA data in their original form, but only pre-processed and are extra to purchase. Alternative data can be found in a variety of different formats and cannot be imported in every software. Looking ahead, standardized formats of LCIA data would simplify the import into any LCA software and therefore will improve the measurement results.

During the conception of the test phase, it was noticed that it was difficult to impossible to obtain appropriate licenses for the software to be tested without having to purchase them. The sales department of the companies could not be convinced, despite clear target communication, to allow a free comparison with their software. It would be advisable for procedures such as the Blue Angel for sustainable software products to be implemented not only as an eco-label, but also, for example, in the form of directives, e.g. within the European Ecodesign Directive.

Bibliography

[BE22] https://www.umweltbundesamt.de/sites/default/files/medien/479/publikationen/texte_119-2021_umweltzeichen_blauer_engel_fuer_ressourcenund_energieeffiziente_softwareprodukte.pdf accessed: 19/05/2022

[BS22] Dr. Bernhard Steubing, Universität Leiden - mail 2022

[GL22] https://gitlab.rlp.net/green-software-engineering/mobiles-messgerat/-/blob/master/get_energy_utilization_mit_Sync.py, accessed: 19/05/2022

[Gr18] Gröger et al. 2018 (Universität Zürich - UZH)

[TU22] https://www.semanticscholar.org/paper/Vergleich-von-Softwarel%C3%B6sungen-f%C3%BCr-die-%E2%80%93-eine-L%C3%BCdemann-Feig/cfb699d68e9e9c1f8dea22fbc92fdfd1de13fddf , accessed: 19/05/2022

[UC22] https://www.umwelt-campus.de/fileadmin/Umwelt-Campus/Greensoft/measurement_of_software-induced_energy_consumption_web.pdf, accessed: 19/05/2022

GREEN DATA CENTRES

Transient numerical simulation for optimization of a water-cooled high-performance computing center with dynamic cooling circuit temperatures – Work-in-progress

Nils Bayer[1], Henner Kerskes[2] and Konstantinos Stergiaropoulos[3]

Abstract: Due to the high consumption of electrical energy, data center operators are increasingly aiming for a data center that is as energy-optimized and efficient as possible. The cooling system has a crucial role here. In this context, an innovative control strategy with dynamic cooling circuit temperatures will be investigated. For this purpose, a detailed numerical simulation model of an existing water-cooled HPC data center cooling system was developed. The model was validated using real operating data. It includes the complete heat flow from the rack cooling distribution units to the cooling towers and the district cooling supply. The model can correctly simulate the thermal behavior in detail and is the basis for further investigations and optimization in terms of various criteria like energy or cost efficiency. The modeling and validation approach is presented in this paper.

Keywords: data center cooling, high-performance computing, simulation, energy efficiency

Addresses Sustainable Development Goal 13: Climate action

1. Introduction

In recent years, there has already been great progress in terms of data center efficiency. Still, the consumption of electrical energy by data centers is predicted to increase due to the growing data volumes and their processing. In [KW21], an energy consumption of up to 1287 TWh in 2030 is forecasted in different scenarios, starting from 286 TWh in 2016. However, all scenarios are subject to a high degree of uncertainty. Other studies such as Borderstep [HH19] also predict an energy consumption in this range of slightly below 1000 TWh, whereas Andrae [Aa19] and Belkhir and Elmeligi [BE18] estimate around 2000 TWh for the year 2030.

The High-Performance Computing Center in Stuttgart (HLRS) is one of the three national supercomputing sites in Germany organized in the Gauss Center for Supercomputing

[1] University of Stuttgart, Institute for Building Energetics, Thermotechnology and Energy Storage, Pfaffenwaldring 6, Stuttgart, 70569, Germany, nils.bayer@igte.uni-stuttgart.de
[2] University of Stuttgart, Institute for Building Energetics, Thermotechnology and Energy Storage, Pfaffenwaldring 6, Stuttgart, 70569, Germany, henner.kerskes@igte.uni-stuttgart.de
[3] University of Stuttgart, Institute for Building Energetics, Thermotechnology and Energy Storage, Pfaffenwaldring 6, Stuttgart, 70569, Germany, konstantinos.stergiaropoulos@igte.uni-stuttgart.de

[Ga22]. Fig. 1 shows the long-term trend of increasing energy consumption, resulting from more powerful supercomputers with every new generation. In 2020, the transition from the Hazel Hen computer generation to the new Hawk took place. Normal operation was therefore interrupted so that the annual energy consumption was exceptionally low in that year. In the future, the predicted energy consumption of the Hawk computer is approximately 33 GWh per year. A continuously increasing energy consumption is associated with high CO_2 emissions, as electricity currently often still has a high specific CO_2 emission factor.

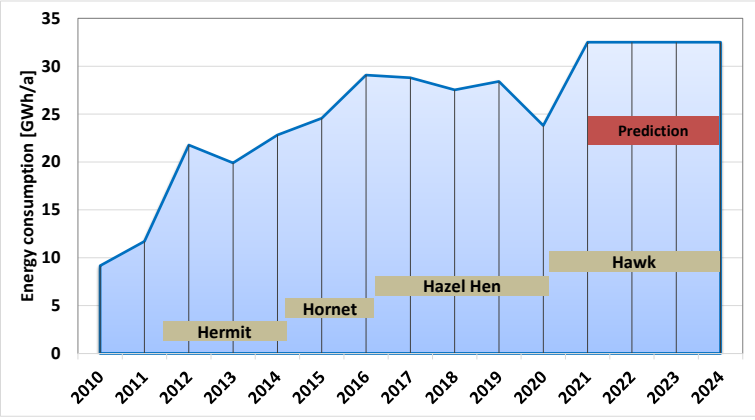


Fig. 1: Trend of energy consumption for HLRS from 2010 to 2024

In order to reduce the high CO_2 emissions, it is necessary to improve the energy efficiency of high-performance computing (HPC) data centers. Although IT components are becoming more efficient, the cooling system must also be considered and improved. This will result in savings of electrical energy and thus in a reduction of CO_2 emissions. Power usage effectiveness (PUE) is a simple and widely used indicator for measuring the energy efficiency of data center infrastructures [AAF12]. The PUE is calculated from the ratio of the used energy of the total facility (IT equipment plus everything that supports the IT equipment) and the IT equipment:

$$PUE = \frac{Total\ Facility\ Energy}{IT\ Equipment\ Energy} \tag{1}$$

The objective is to achieve a PUE as close as possible to 1. For the year 2021, the average annualized PUE of the largest participating data centers is 1.57 according to a survey by the Uptime Institute [Bi21]. The desired future values are PUE < 1.1, which some data center operators such as Google are already achieving with an average PUE of 1.10 for their data centers [Sh16][Go22]. The largest savings potentials in PUE can be achieved through optimized cooling. Using different cooling technologies, the combined PUE of all systems is calculated. The energy efficiency ratio (EER) can be used to determine the efficiency of such cooling systems. The EER of cooling systems is the ratio of the useful cooling $\dot{Q}_{cooling}$ and the consumed electrical power P_{cooling}:

$$EER = \frac{\dot{Q}_{\text{cooling}}}{P_{\text{cooling}}} \tag{2}$$

Free cooling or chiller-less cooling is an effective solution for reducing the power consumption of cooling systems, as vapor-compression chillers consume a lot of electricity and thus have a significantly lower EER [Zh14]. The heat to be dissipated is emitted to the environment via cooling towers, whereby the temperature level of the heat dissipated must be above the ambient conditions. A recent trend in server development is therefore towards higher operating temperatures in order to increase the share of free cooling. However, it needs to be considered that the increase in operating temperatures also has negative effects. Higher operating temperatures result in higher power consumption for numerous components within the data center, especially for semiconductors such as processors. In addition, the computing performance and the lifetime of the IT components can also decline [No16]. Studies were conducted on the temperature dependencies in HPC data centers in terms of power consumption and computing performance [Sh17][NHS20]. Thus, higher temperatures do not only lead to positive effects. One way to counteract this is to implement a winter-summer operation, as reported by the Leibniz Supercomputing Centre in Garching [Wi17]. In summer, the cooling circuit temperatures are increased to a sufficiently high level so that free cooling can be achieved. In winter, in contrast, the cooling circuit temperatures are lowered again.

A new approach is a dynamic cooling circuit temperature depending on the ambient conditions and the IT-load, which could result in much more efficient data centers. The cooling circuit temperatures are raised just high enough to allow free cooling at all times, while at the same time avoiding unnecessarily high cooling circuit temperatures. This aspect is addressed by the work started in this paper.

2. Description of HLRS cooling system

A schematic overview of the HLRS cooling system is shown in Fig.. The dominant system by far is the direct water-cooled Hawk supercomputer with 5632 nodes and 11264 64-core CPUs AMD EPYC 7742 and 1.44 Petabytes of main memory. The nodes are arranged in 44 non-standard racks in three rows. Two cooling distribution units (CDU) per row provide inner circuits for directly cooling CPUs, DIMMs, Chipsets and internal network components. The system has an electrical power consumption of approx. 3.2 MW. The data center also has further server racks, such as the necessary storage racks and other IT components that are air-cooled. The cooling capacity for the air-cooled IT systems and the computer room base load is small in comparison to the water-cooled supercomputer and is not important in the studies on the dynamic cooling circuit temperatures. The air-cooled systems and the associated primary and secondary circuits of the air cooling are therefore not discussed further. Only the primary and secondary circuits of the water-cooled systems are considered in this paper. The six CDUs cooling the racks are responsible for the major heat input into the cooling system. They are coupled to the secondary cooling circuit via

heat exchangers. In addition to the CDUs, a flywheel uninterruptible power supply (UPS) is the second heat input into the secondary circuit of the water cooling system. Three identical 1.7 MW UPS systems compensate frequency and voltage fluctuations and short-term interruptions of the electrical network. Since the UPS does not operate without losses, it must be cooled accordingly. The cooling of the CDUs and UPS in the secondary circuit is mainly provided by wet cooling towers installed in the primary circuit of the water cooling system. At high ambient temperatures and in the event of malfunctions of the cooling towers, the facility is supplied with cold via a connected district cooling network. The district cooling is provided by vapor-compression chillers. A special characteristic is that the UPS has an additional connection to the district cooling network, as lower flow and return temperatures are required for this than for the CDUs. Furthermore, two heat pumps use the secondary circuit of the water cooling system as a heat source. The heat pumps provide space heating for the data center building and a neighboring office building. In terms of energy, the heat pumps have a minor impact, as less than 1 % of the waste heat is required to operate the heat pumps over the year. A total of four identical cooling circuits, each with a wet cooling tower with a cooling capacity of 1.2 MW, and a connection to the district cooling network via a 4 MW heat exchanger are available for cooling. The capacity of each cooling system is sufficient to handle the load independently. The share of free cooling in the water cooling system was between 82 and 86 % for the years 2019 to 2021.

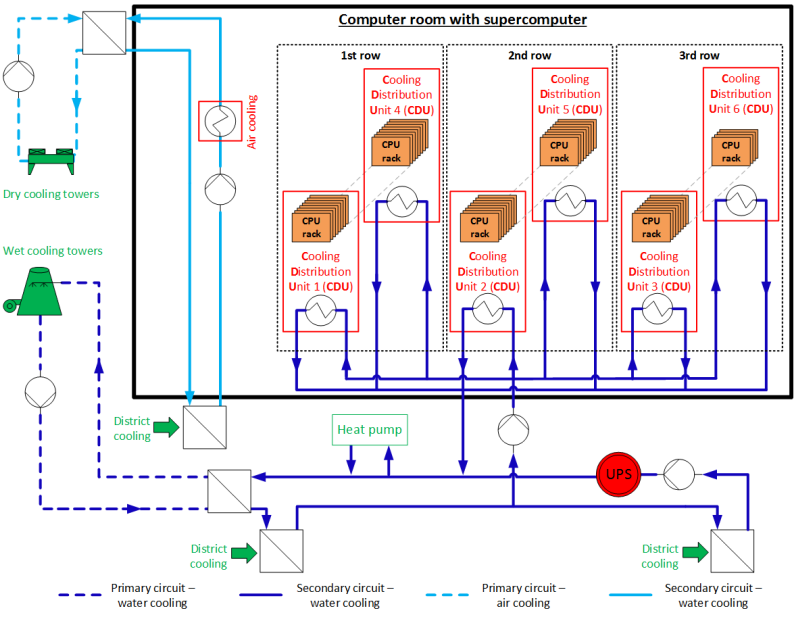


Fig.2: Schematic overview of the HLRS cooling system

3. Modeling of HLRS cooling system

The simulation model design is based on the detailed piping and instrumentation diagram of the HLRS cooling system. Due to its complexity, the system is implemented in a simplified but still detailed way. For the development of the model, the operational characteristics of the individual components of the cooling system were reproduced based on manufacturer and planning documents, measured values and on-site inspections. The simulation environment used is TRNSYS (TRaNsient SYstem Simulation [Wi75]), which provides extensive libraries. These include all necessary components such as pumps, wet cooling towers or heat exchangers, which can be linked together to simulate an overall system.

3.1 Heat inputs

The electrical power consumed by Hawk is converted into heat and must be dissipated to avoid overheating. This is done by the CDUs. Modeling has so far been limited to the heat input of the CDUs into the secondary circuit of the water cooling system. Internal processes of the CDUs are not yet considered in detail. The supercomputer is thus described as a heat source with a variable load, which is required as an input variable. The three UPS units are responsible for the second heat input into the secondary circuit of the water cooling system. For determining the heat transfer into the cooling system, the correlation to the resulting active power of the UPS was determined using measurement data and polynomial regression. The active power is used as the input variable for the UPS model.

3.2 Cooling systems

The four identical cooling circuits, each with a wet cooling tower using the counterflow method, are modeled using a wet cooling tower model with four tower cells. The water in the primary circuit to be cooled is injected into an unsaturated air flow and partially evaporated. Fans supply the required air volume flow. The resulting combination of heat and mass transfer can theoretically reduce the temperature to near the wet bulb temperature of the ambient air [ER13]. Manufacturer data in combination with measured data are used to parameterize the wet cooling tower model. In addition, the model is fed with overall performance data to determine the coefficients of the mass transfer correlation. To determine the required performance data in Tab. 1, measurement data are used on the waterside and measurement data from a weather station in the immediate vicinity are used on the airside. For best results, the performance data cover a range of conditions characteristic of the expected operation of the wet cooling tower. The exact procedure for the calculation can be found in [Tr17]. The water mass flow in the primary circuit is cooled depending on the water and air mass flow conditions at the inlet. The water evaporated in the wet cooling towers is replenished by fresh water. Additional water is needed for

desalination. The output variables of the wet cooling towers are the states of the water and air mass flow as well as the water and energy consumption.

Symbol	Description	Units	Symbol	Description	Units
$\dot{V}_{\mathrm{air}}$	Air volumetric flow rate	m³/h	$\dot{m}_{\mathrm{water}}$	Water mass flow rate	kg/h
$T_{\mathrm{air,in}}$	Air dry bulb temperature	°C	$T_{\mathrm{water,in}}$	Water inlet temperature	°C
$T_{\mathrm{air,in,wb}}$	Air wet bulb temperature	°C	$T_{\mathrm{water,out}}$	Water outlet temperature	°C

Tab. 1: Mandatory performance data of the wet cooling towers

Besides the cooling towers, district cooling is the second cooling source. It is modeled as an ideal cooling device in a district cooling circuit. The district cooling circuit is coupled to the secondary circuit of the water cooling system via two heat exchangers. The maximum cooling capacity of the district cooling is assumed always available, as vapor-compression chillers provide it. For determining the electrical energy consumption of the chillers, the EER is taken into account. According to the information from the district cooling supplier it is approx. 4.5 on average. For the completeness of the energy balance, despite the small quantitative share on an annual basis, the heat pump is included in the model. The heat flow required to operate the heat pump is drawn from the secondary circuit.

3.3 Hydraulic components and control

All pumps and heat exchangers in the cooling system are implemented in the model, with redundant pumps and heat exchangers combined into one element. The pumps are all speed-controlled glanded pumps, so that the heat transfer into the cooling system is minimal. The pumps were parameterized on the basis of their data sheets. The required pump power is modeled with a polynomial as a function of the mass flow rate, using a control signal as the input variable. The heat exchangers are designed on the counterflow principle and are modeled as zero capacitance heat exchangers. The needed input is the overall heat transfer coefficient, which is given for all heat exchangers from the technical data sheets, and the inlet conditions of the mass flows. The heat exchanger effectiveness and the outlet conditions are then computed during the simulation. The heat exchangers are insulated. Furthermore, all necessary pipes in the cooling system are included in the model, but are implemented as adiabatic, as all pipes are also insulated. The energetic influence of the pipes is negligible, confirmed by measurement data. The control of the simulation model has been simplified in comparison to the real system. The setpoint temperatures of the CDUs and the UPS units must be defined. This results in the main control parameter, which is the inlet temperature of the cooling water into the CDUs both in the real cooling system and in the simulation. To maintain the set inlet temperature, the fans of the wet cooling towers, the pump in the primary circuit and the valves for district cooling supply are controlled. Additionally, the other pumps are controlled according to the necessary mass flows in the cooling system. The controllers in the simulation were implemented as PI controllers. The result is an overall simulation model that includes the

complete heat flow from the rack CDUs to the cooling towers and the district cooling supply, including the hydraulic components and the control system.

Fig. 3 summarizes the relevant input and output variables of the simulation model. Weather data, the active power of the UPS, the heat inputs of the CDUs and the definitions of the temperature levels in the secondary circuit are required as inputs. The outputs from the model are the cooling capacity of the wet cooling towers and the district cooling and additionally the electricity consumption of the pumps and fans as well as the water consumption.

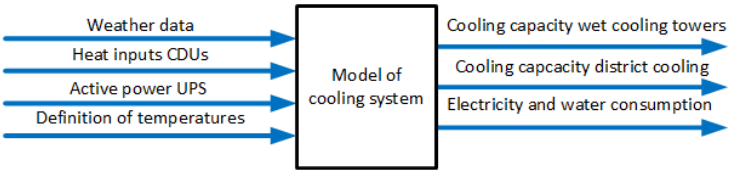


Fig. 32: Input and output variables of the simulation

3.4 Validation

Operational data of the HLRS from 2021 were used to validate the simulation model. The measurement data were extracted from the existing building control system. The time step of the measured data and the simulation was five minutes. The simulation results were compared with the measured data both qualitatively in the form of power curves and scatter plots and quantitatively as monthly energy balances. Fig. 4 shows the averaged cooling capacity of the wet cooling towers from both measurement and simulation over the course of the year 2021 using a moving average filter. Daily and seasonal variations due to varying ambient conditions can be seen.

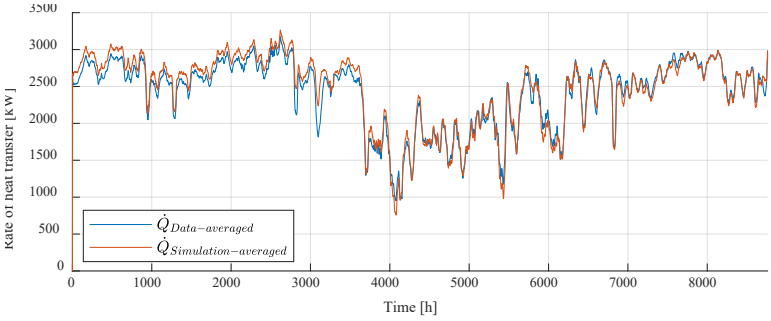


Fig.4: Comparison of averaged cooling capacity of wet cooling towers for the year 2021 (Averaged: MAE = 84.45; RMSE = 106.17; Not averaged: MAE = 198.34; RMSE = 353.44)

The match of the averaged courses of the simulation in comparison to the measurement is good. Considering the non-averaged curves, scattering occurs due to the dynamic system

and thus there are slight deviations in the cooling capacity between the measurement and simulation, see Fig. 5. For both periods, the root mean squared error (RMSE) and mean absolute error (MAE) are given.

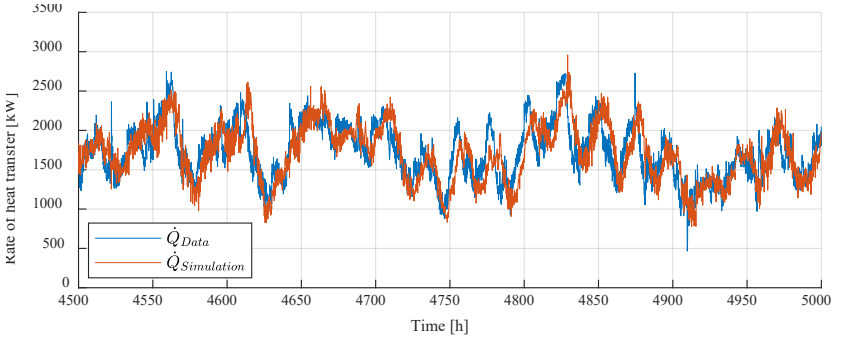


Fig. 53: Comparison of cooling capacity of wet cooling towers between hours 4500 and 5000 (MAE = 202.54; RMSE = 256.13)

The transient behavior of the cooling system combined with a high variability is also shown in Fig. 6. It shows the cooling capacity of the wet cooling towers over the wet bulb temperature for the year 2021 as a scatter plot for the measured and simulated data. In addition, averaged curves are illustrated using a moving average filter. The cooling capacity of the wet cooling towers is dependent on the wet bulb temperature. From a wet bulb temperature of approx. 10 °C, the cooling capacity decreases for rising temperatures. The missing cooling capacity must be covered by district cooling at high wet bulb temperatures, primarily in summer.

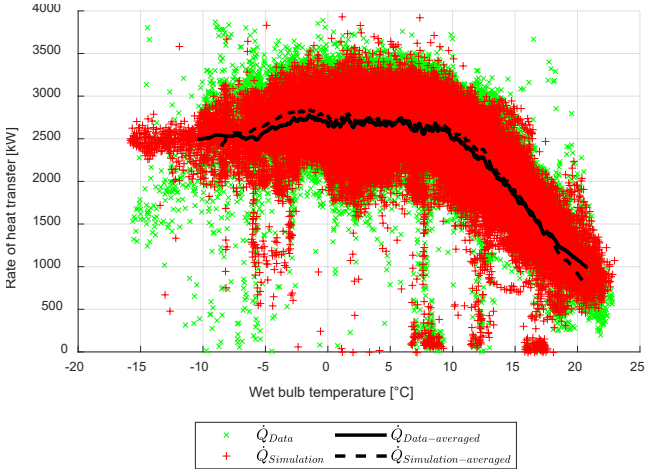


Fig.6: Comparison of cooling capacity of wet cooling towers as a dependency of the wet bulb temperature for the year 2021

This correlation can also be seen in Fig. 7, which shows the monthly energy balances for the year 2021. The percentage of the wet cooling towers and district cooling in terms of cooling is shown for both measurement and simulation. A characteristic annual course results. In summer, in months 6 to 9, a high proportion of district cooling was required, up to a maximum of 42 % on a monthly basis. In the other months, the wet cooling towers almost completely cover the needed cooling. In general, the simulation and the measurement show a high degree of agreement. The dynamic operation of the complete water cooling system of the HPC data center can be simulated successfully.

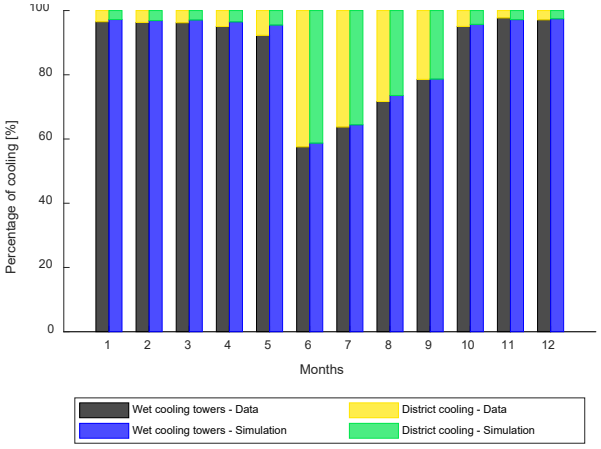


Fig. 7: Comparison of the percentage ratios in terms of cooling for the year 2021 (MAE = 1.01; RMSE = 1.30)

4. Summary and future work

In this paper, a transient simulation model of a water-cooled HPC data center cooling system is presented. The developed model includes the complete heat transfer from the CDUs to the cooling towers and the district cooling network. The model indicates the needed cooling capacity of the wet cooling towers and district cooling depending on the weather data, the heat inputs into the system and the defined temperature levels in the secondary cooling circuit. In addition, the consumption of electrical energy of the pumps and fans as well as the water consumption are calculated. The model is validated and can simulate the real operating behavior of the cooling system in an excellent way.

In the next step of our work, the developed model is used for further investigations of the dynamic cooling circuit temperatures. The problem of the area of conflict between the lowest possible operating temperatures and a maximum of free cooling will be investigated using extensive simulation studies. Cooling circuit temperatures are dynamically adjusted to allow free cooling at all times, while avoiding unnecessarily high cooling circuit temperatures to reduce the negative impact on power consumption,

computing performance and the lifespan of IT components. Due to the temperature dependencies of both the cooling system and the computer performance, it is necessary to consider the entire system. For this purpose, a temperature-dependent processor performance model is being developed that complements the current cooling system simulation model. For the parameterization of the model, benchmarks will be carried out at different temperature levels, which provide information on both the electrical power consumption of the processors and the computing performance. The model will be completed by detailed modeling the CDUs, which are presently still rather simplified.

The resulting numerical model is then used for a simulative analysis and subsequent optimization with regard to various criteria with cooling circuit temperatures that are dynamically controlled. Optimization targets are energy and emissions, costs of operation, which are implemented by cost functions, and the computing performance. The result will be an optimized operating strategy for the data center in terms of various criteria. The work has a model character and serves the transfer of the results to other data centers. These studies are carried out for the first time in the DEGREE project, funded by the German Federal Environmental Foundation, grant agreement number 37006/01-24/0.

Bibliography

[Aa19] Andrae, A.: Projecting the chiaroscuro of the electricity use of communication and computing from 2018 to 2030, Preprint 1-23, 2019.

[AAF12] Avelar, V.; Azevedo, D.; French, A.: PUE: A comprehensive examination of the metric. The Green Grid White Paper No. 49, 2012.

[BE18] Belkhir, L.; Elmeligi, A.: Assessing ICT global emissions footprint: Trends to 2040 & recommendations. Journal of Cleaner Production Vol. 177, pp. 448-463, 2018.

[Bi21] Bizo, D. et al.: Uptime Institute Global Data Center Survey 2021. UI Intelligence Report 51, 2021.

[ER13] Erens, P.; Reuter, H.: Kühltürme, 11th ed. VDI-Wärmeatlas, Berlin, Heidelberg: Springer Verlag, pp. 1655-1674, 2013.

[Ga22] Gauss Centre for Supercomputing, https://www.gauss-centre.eu/, last access: 05/10/2022.

[Go22] Google Data Centers, https://www.google.com/about/datacenters/efficiency/, last access: 05/11/2022.

[HH19] Hintemann, R.: Hinterholzer, S.: Energy consumption of data centers worldwide: How will the Internet become green? CEUR Workshop Proc., 2019.

[KW21] Koot, M.; Wijnhoven.; F.: Usage impact on data center electricity needs: A system dynamic forecasting model. In Applied Energy Vol. 291, 116798, 2021.

[NHS20] Nonaka, J.; Hanawa, T.; Shoji, F.: Analysis of Cooling Water Temperature Impact on Computing Performance and Energy Consumption. In 2020 IEEE International Conference on Cluster Computing, pp. 169-175, 2020.

[No16] Nosayba, E. et al.: Temperature management in data centers: Why some (might) like it hot. ACM SIGMETRICS Performance Evaluation Review Vol. 40, Issue 1, pp. 163-174, 2012.

[Sh16] Shehabi, A. et al.: United States Data Center Energy Usage Report, 2016.

[Sh17] Shoukourian, H. et al.: Analysis of the efficiency characteristics of the first High-Temperature Direct Liquid Cooled Petascale supercomputer and its cooling infrastructure. In J. Parallel Distrib. Comput. Vol. 107, pp. 87-100, 2017.

[Tr17] TRNSYS 17, a TRaNsient System Simulation programm, Volume 4 Math. Reference

[Wi17] Wilde, T. et al.: Simulation und Optimierung des Energiekreislaufs von Rechenzentrums-Klimatisierungsnetzen unter Berücksichtigung von Supercomputer-Betriebsszenarien (SIMOPEK) - Abschlussbericht, 2017.

[Wi75] University of Wisconsin-Madsion, TRNSYS, a Transient Simulation Programm. Madison, Wis.: The Laboratory, 1975.

[Zh14] Zhang, H. et al.: Free cooling of data centers: A review. In Renewable and Sustainable Energy Reviews Vol 35, pp. 171-182, 2014.

SUSTAINABLE MOBILITY

Hydrogen Technology Business Process Management Modeling: Standardization and digitization of processes within the hydrogen infrastructure

Uta Kühne[1], Miriam O'Shea[2], Benjamin Wagner vom Berg[3] and Lars Wöltjen[4]

Abstract: Standardization and digitization are important factors to accelerate the development of hydrogen infrastructure for transport and mobility purposes. The digitization of administrative processes in particular can help to make information-intensive and often complex processes within authorities more efficient and transparent. A uniform standardization of technologies and process descriptions can simplify and accelerate the market launch and approval processes. The project 'Hydrogen Technology Business Process Management Modeling' concentrate on the two focus areas standardization and digitization to ease the application process of hydrogen filling stations.

Keywords: Hydrogen infrastructure processes, Standardization and digitization, Business Process Management Modeling

Addresses Sustainable Development Goal 9: Industry, innovation and infrastructure

1. Introduction

Digital technologies are an important driver for sustainable development and therefore have a special place in the implementation of the United Nations Agenda 2030. Within the 17 Sustainable Development Goals (SDG), information and communication technologies are directly named in several goals, for example in SDG 4 "Quality education" and in SDG 9 "Industry, innovation and infrastructure". Using digitization as a tool can increase resource efficiency, manage the energy transition and improve climate protection [Bm21].

The digitization of administrative processes in particular can help to make information-intensive and often complex processes within authorities more efficient and transparent. This enables the user to obtain quick and targeted information, improves the relationship between citizens and the state and opens up new opportunities for innovation for the economy through accelerated processes. Digitization is therefore an innovation driver that

[1] Hochschule Bremerhaven, Smart Mobility Institute, An der Karlstadt 8, 27568 Bremerhaven, Germany, uta.kuehne@hs-bremerhaven.de
[2] Hochschule Bremerhaven, Smart Mobility Institute, An der Karlstadt 8, 27568 Bremerhaven, Germany, moshea@hs-bremerhaven.de
[3] Hochschule Bremerhaven, Smart Mobility Institute, An der Karlstadt 8, 27568 Bremerhaven, Germany, benjamin.wagnervomberg@hs-bremerhaven.de
[4] Hochschule Bremerhaven, Smart Mobility Institute, An der Karlstadt 8, 27568 Bremerhaven, Germany, lwoeltjen@studenten.hs-bremerhaven.de

can become an innovation accelerator, especially when expanding a green hydrogen infrastructure. An important companion of digitization is standardization. In hydrogen and fuel cell technology in particular, the topic of standardization is a complex, internationally oriented and continuous process. Due to the novelty of the technology carriers and their application, especially in mobility, there are different technical requirements for identical technologies as well as ambiguities in the classification of e.g. hydrogen filling stations in various application scenarios [No22, JD18]. A uniform standardization of technologies and process descriptions can simplify and accelerate the market launch and approval processes [Bm22].

In 2021, we first reported on our project H2BPMM, which stands for 'Hydrogen Technology Business Process Management Modeling'. In this article, we would like to publish the latest findings and open them to discussion. The results are twofold and directly match the two focus areas of the project: (1) standardization and (2) digitization to ease the application process of hydrogen filling stations. In the area of standardization, the analysis of documents and the interviews with two focus groups have supported the development of a general taxonomy of the configurations of filling stations, which we present in a morphological box. Further analysis led to the identification of today's typical filling station. Moreover, through the classification of the parameters of the taxonomy, it became apparent that there are two types of configurational factors, one of them leading directly to the relevant legal procedure for the application. This finding is very important in the pursue to ease the application process. In the second focus area of the project, digitization, those parameters driving the relevant legal procedure have been used to build a tool that implements the relationships between the configurational setup of the filling station and the legal procedure being relevant for it into an automated decision table. Thereby, the first and critical task in the application process is supported and the user is guided to find the right legal procedure which in turns leads to a known set of documents and parameters to be supplied within the application. With the digitization of processes in connection with the development and expansion of hydrogen technology, a contribution should be made to making those processes clearer and more efficient and thus to accelerate infrastructure expansion and climate protection. The H2BPMM project contributes to this by simplifying the complex matter of the approval process for both focus groups, thus saving time and money.

2. Hydrogen Technology Business Process Management Modeling

As in any infrastructure-dependent technology, hydrogen mobility suffers from the typical the-chicken-or-the-egg-problem: Neither transport and mobility providers nor the infrastructure companies are willing to take the first step without financial support [Ma03, pp. 743-744]. To accelerate the development of hydrogen infrastructure for transport and mobility purposes, it is essential to simplify the approval processes involved by digitalization and standardization. Careful attention thereby is paid to digitalization being part of the solution not the problem. Here, digitalization is used to prevent multiple and

unclear requests for information. In respect to the technology implemented, the end-user is not to be overburdened, using standard web browsers. Other stakeholders will enjoy synergy effects with general applications in the area of e-government. There is an inherent complexity, which can be seen, for example, in the approval processes required by German law. Here, different legal bases apply, depending on the specific configuration of the infrastructure design. In general, two main legal bases need to be considered for the approval, dependent on various filling station parameters:

- Regulation on health and safety in the use of work equipment (Betriebssicherheitsverordnung: BetrSichV)
- Federal Immission Control Act (Bundes-Immissionsschutzgesetz: BImSchG)

In addition, there is a large number of different parameters and critical values that influence the approval of the hydrogen filling station and affect the choice of the right permit or approval procedure (see chapter 3). The aim of the H2BPMM project is to support the establishment of the infrastructure required for hydrogen drive technology.

Approval processes, e.g. of hydrogen filling stations, but other processes in mobility and logistics also, are information heavy. Here, the flow of information between different stakeholders and systems is of great importance. The design of processes of that kind must support the various users involved in respect to the provision of information adequately. For this purpose, the corresponding process must be designed with sufficient precision especially concerning the information flow. This is often not the case yet. First, the use cases of hydrogen related procedures are rather new to e-government, mobility and logistics. Second, currently used notations for the design of processes do not detail the information flow to the extent needed in information heavy use cases. Up to now, information is considered in the form of artefacts like documents or similar containers. The life-cycle is applied to these containers as well. However, a document for example contains several information objects, whose life-cycle can vary. To clearly identify what information object is needed or generated in what process step, containers need to be broken up and information objects need to be referenced individually. [cf. SZS04, KR11, SZN06, SZ06, At11] As BPMN is focused on the flow of activities and not the information flow, the notation and methodology needs to be enriched. Here, in a separate stream of analysis, the information flow is gathered and integrated into the activity flow.

The H2BPMM project started on July 1, 2020 and is planned for two and a half years. The project is being carried out by the University of Applied Sciences Bremerhaven and is supported by a large number of partners from companies, associations, societies, cities, municipalities, districts and the two states of Bremen and Lower Saxony.

3. Standardization of processes and methods to ease the standardization of the approval process

With the help of the standardization of processes, a uniform method or a predetermined result is defined for the respective application. Some key benefits of process standardization are:

- An increase in the transparency of the processes and the required documents.
- The same or similar process steps only have to be mapped once, so that duplications are avoided.
- The effort involved in submitting an application is minimized.
- As part of digitization, standardized interfaces to other users, such as authorities and municipalities are created.

An essential first step for the standardization of existing processes for which little experience is available is the recording of the current status. For the H2BPMM project and the approval process for hydrogen filling stations, this means to research different sources: first to name is the analysis of the technical configurations, the analysis of current processes, next an inspection of existing guidelines, as well as the recording of experiences from people being involved in the approval of hydrogen filling stations. Various methods have been used for this within the H2BPMM project. Existing approval processes are recorded via literature research and document analysis. The need for an analysis of the configuration arises from the fact that, depending on the technical and commercial configuration of a hydrogen filling station, a specific approval procedure applies. A so-called morphological analysis was carried out to record the typical characteristics of a hydrogen filling station. This was followed by an exchange with companies and applicants in a workshop and in the form of expert discussions. In the following, the methods to simplify the standardization and results are presented.

3.1 Morphological Analysis

The morphological analysis was developed by the astrophysicist Fritz Zwicky as a creativity technique that enables the users to fully understand complex problems and to represent all possibilities. In this way, it can make a significant contribution to solving complex problems [Zf48, pp. 122-125]. To achieve this, the method uses what is known as a Zwicky-Box, which is a vivid, multi-dimensional matrix. In our application, a subset of the Zwicky-Box called a typological field format lists the properties (influencing variables) that are independent of one another in one dimension and all of their possible characteristics in another. This creates an overview of all conceivable possible combinations of characteristics [Sm14, pp. 618-619].

In order to achieve the project goals as described above, the complexity of the approval process has to be determined first. The properties and characteristics of the hydrogen filling station have to be analyzed and documented, because they significantly influence the legal basis for the approval process. With the findings of our literature review recorded in the typological field format, the multitude of technical properties could be presented objectively and any premature evaluation was prevented (technology openness). Fig. 1 shows an exemplary section of the morphological analysis created in the project and therewith the complexity of the technical properties under study.

An evaluation of the literature review has taken place in form of expert interviews and workshops. Moreover, existing hydrogen filling stations were mapped in the typological field format and thereby most commonly used configurations of hydrogen filling stations were elicited. The latter forms the first step towards the definition of a "standard".

No.	Influencing Factors	Characteristics						
1.	Use	public	only on the company premises	public and on the company premises	others			
2.	Operator	private	service provider	company	federal / state / local authority	other		
3.	Responsibility process of approving	operator	owner	service provider	other			
4.	Compulsory public participation	yes	no					
5.	Fixed installation	yes	no					
6.	Target sector hydrogen filling station	passenger car	bus	lorry	shipping traffic	air traffic	rail traffic	industrial truck
7.	Type of area	industrial area	commercial area	village area	mixed area	urban area	core area	Other Special area
8.	Total hydrogen storage	< 3 tons	> 3 tons	> 5 tons	> 30 tons	> 50 tons	> 200.00 tons	
9.	Hydrogen quantity threshold in the incedent area	< 5 tons	5 tons to < 50 tons	> 50 tons				
10.	State of aggregation when it is provided at the hydrogen fueling station	gaseous	liquid	solid (chemically bound)				
11.	Pressure when releasing hydrogen	200 bar to < 350 bar	350 bar	700 bar				
12.	Hydrogen production process	electrolysis	reforming	coal gasification	bio-, photo-, thermochemical processes	hydrogen as a byproduct	Other	

Fig. 1. Excerpt of the typological field

3.2 Document Analysis

The morphological analysis delivered much information about the potential configuration of hydrogen filling stations. However, information about the scope and specific content of the approval process could not be determined within the morphological analysis. To obtain this kind of information, a comprehensive document analysis of the underlying legal texts and supporting guidelines concerning the approval process was conducted. All available

literature was examined and, if it contained hints to information objects, these were subsequently compiled, organized and structured [Sm14, pp. 545-546; 8, pp. 39-40]. Resulting, a catalogue of the information fields included in the approval processes was derived.

As mentioned in the introduction, different approval processes can be applicable depending on the individual configuration of the hydrogen filling station at hand. Consequently, a variety of different legal documents can be necessary within the approval application procedure. The previous literature review includes more than five supporting guidelines, e. g. both the volumes of the "Portal Green Power to Gas Leitfaden" [Po20] having more than 460 pages, "Genehmigungsverfahren nach dem BImSchG, Leitfaden für Antragsteller" [Ni20] and "NOW-Genehmigungsleitfaden für Wasserstoffstationen" [Na22]. The laws and ordinances were researched on the website of the Federal Ministry of Justice and the Federal Office of Justice with the current entire federal law. The initial results of the document analysis within the H2BPMM project have been discussed with various experts of the approval authorities and of applicants in companies and service providers.

3.3 Workshops and Configurations of hydrogen filling stations

During the course of the project, it became clear that the focus of the project is on two main groups - the authorities responsible for the permit and approval procedures on the one hand and the planners, applicants and operators of hydrogen filling stations on the other. Workshops and expert discussions were therefore held with both focus groups in order to learn more from their previous knowledge for the description of the current situation.

As a result of this exchange, it can be said in summary that the approval procedure to be used in each case depends on several overriding factors. On the one hand, these are parameters and limit values, such as the total storage quantity of gaseous fuels and the on-site generation of hydrogen by electrolysis. With a total stored quantity of less than three tons by one operator or three to five tons by two economically independent operators, the much simpler procedure under the Regulation on health and safety in the use of work equipment applies. If the requested total storage quantity of gaseous fuels increases to more than three tons or if the hydrogen is produced directly on site, then the planned hydrogen filling station must be approved according to the much more complex procedure under the Federal Immission Control Act. The following Fig. 2 shows the process selection based on the parameters and limit values for the Regulation on health and safety in the use of work equipment. With these insights, it is possible to make the approval process more transparent for users. Depending on the use case, an applicant can be advised to stay below the critical storage quantity in his plans. Other factors that can influence the approval process is the configuration of the hydrogen gas station and its location. A distinction is made between the following four configurations:

- The hydrogen pump will be integrated into an existing filling station. The hydrogen dispenser will be subordinate to the existing filling station and will be approved using the same procedure.
- The hydrogen dispenser is self-contained. If this system is classified as a so-called energy system, then it does not require approval according to the Regulation on health and safety in the use of work equipment.
- The hydrogen filling station will be set up on a depot. If classified as an in-house facility and not sold to third parties, this type of gas station does not require approval.
- A mobile hydrogen filling station can be designed as a container or as a truck with a box body. The approval process for this type of hydrogen filling station is still being developed.

For all four configurations as well as for the classifications according to the recorded limit values, knowledge of this makes it much easier to standardize the processes and makes the processes more transparent for applicants and authorities. These relationships form the basis for the next step in the digitization of process flows.

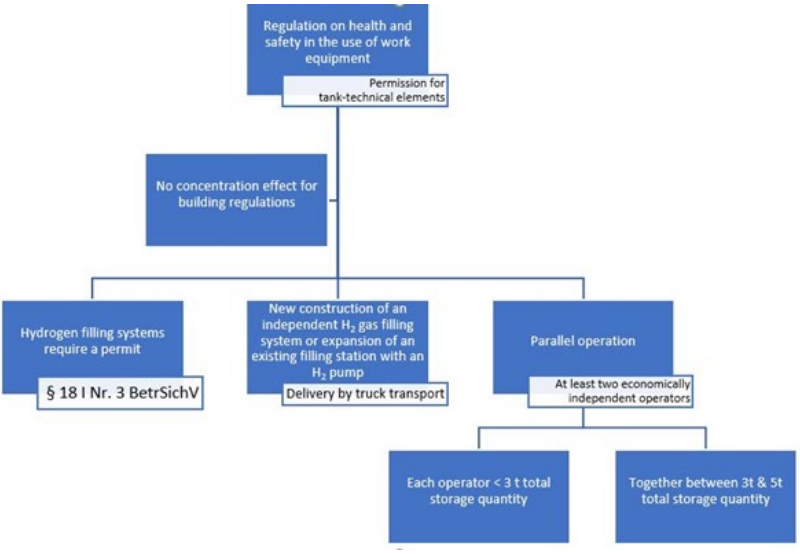


Fig 2. Parameters and limit values of the Regulation on health and safety in the use of work equipment

4. Demonstrator – Decision support for accelerating the expansion of the hydrogen infrastructure

The approval procedures used in Germany, such as the Regulation on health and safety in the use of work equipment or the Federal Immission Control Act, are not only very different in terms of content, but also entail planning and approval phases of different lengths and different costs. In order to make it easier for applicants and authorities to take the first step in selecting the right procedure, a demonstrator was developed in the H2BPMM project, which proposes the approval path to be selected based on the configuration of the respective hydrogen filling station, which is determined by various parameters. Besides the already mentioned prerequisite concerning the knowledge around the approval process in terms of activities, configurations and legal aspects, it is indispensable to identify which information is required at which points within the process. This takes place as a basis for digitization within the standardization work package.

To support the applicant, a tool is needed that provides him or her with information on the procedure based on the planned configuration of the hydrogen filling station and the intended location. This tool was implemented in a "demonstrator" within the H2BPMM project and opened on the project website to public for trial use. Modeling wise, the standard BPMN symbol palette as well as the Decision Model and Notation (DMN) standard were used. An automation of the exemplary process was possible with the help of the CAMUNDA Workflow Engine, i.e. the model could be deployed and executed from the university server. This example prototype proved that (at least partial) automation and therefore support for the applicant is possible with a low investment.

In the following, the notations and systems used to record the current status and to create the demonstrator, as well as the demonstrator itself, are presented. Within the BPMN standard, the workflow (Fig. 3) for decision-making and the issuance of the relevant approval procedure was set up.

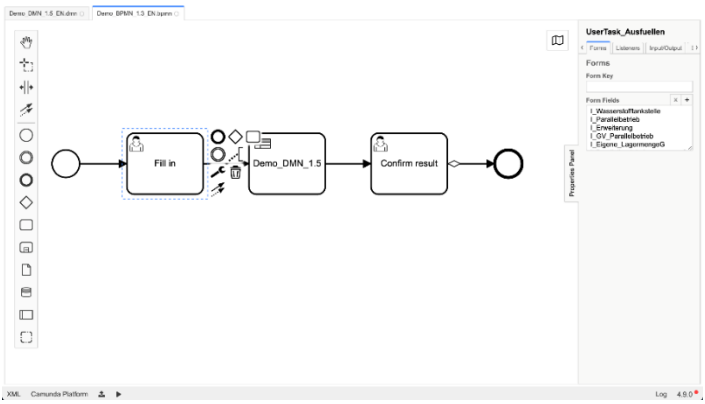


Fig 3. BPMM flow chart demonstrator

The "Fill in" process step is marked here to demonstrate how the definition of parameter queries to be filled out by the user is accomplished. In this process step, the input parameters are determined, which must be described within the decision tables and answered by the applicant. These relate, among other things, on the storage quantity of hydrogen, a possible parallel operation with other operators and the question of the production of hydrogen associated with the filling station. The input parameters are essential for the output process step in which the approval procedure to be used is defined. Within the DMN tool, the parameters are described, the rules for decision-making are defined and the decision modeling is implemented. At the end of the process, the approval procedure defined at the decision level is confirmed as the output parameter and result of the query. The DMN structure of the decision support demonstrator is shown in Fig. 4 below.

The DMN structure includes five decision tables, with the following procedures:

1. Procedure based on the type of integration of the planned hydrogen fuel dispenser together with other fuel station operators.
2. Procedure based on the storage quantities with the limit values for the transition from one procedure to the next procedure.
3. Procedure based on production of hydrogen
4. Procedure due to network connection e. g. with high voltage network
5. Result procedure

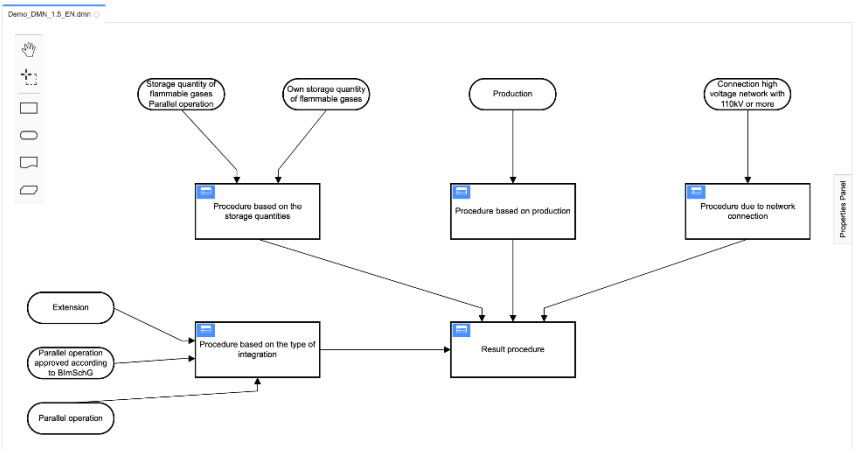


Fig 4. DMN structure of the decision support demonstrator

Fig. 5 shows an example of the queries within the decision tables.

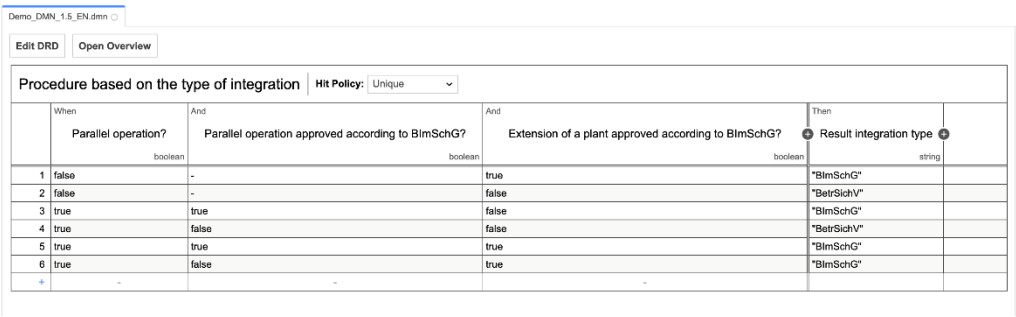


Fig 5. DMN decision table „Procedure based on the type of integration“

Within the “Result procedure” table (see Fig. 6), the results of the four other decision tables are evaluated and, after confirmation, output as the final result for the approval procedure to be applied for.

With the procedure shown, the demonstrator offers applicants the opportunity to classify the time and financial effort of their planned project. In addition, with the help of the demonstrator, the user can take the opportunity to modify the setup of his hydrogen filling station, e.g. by changing the storage capacities, so that it runs into a simplified and less complex approval process. The form of the demonstrator presented can be further expanded and adapted to the various configurations of hydrogen filling stations presented in Chapter 3.3. Furthermore, this type of DMN analysis can be transferred to other applications inside and outside of the hydrogen infrastructure and can be incorporated into the processes and programs of the application authorities as an upstream decision-making process through further modifications. Feedback on the use of the demonstrator referred to its simplicity on the one hand and to suggestions to clearly indicate the critical values in the input step instead of just deriving the resulting legal procedure on the other hand. This suggestion will be implemented in the next version of the demonstrator.

Demo_DMN_1.5_EN.dmn

Edit DRD | Open Overview

Result procedure | Hit Policy: First

	When	And	And	And	Then
	Result integration type	Result stock quantities	Result production	Result grid connection	Result procedure
	string	string	string	string	string
1	-	-	-	"PlanFV"	"Planfeststellungsverfahren"
2	"BImSchG"	-	-	-	"Approval according to Genehmigung nach Bundes-Immissionsschutzgesetz"
3	-	"BImSchG"	-	-	"Approval according to Bundes-Immissionsschutzgesetz"
4	-	-	"BImSchG"	-	"Approval according toBundes-Immissionsschutzgesetz"
5	"BetrSichV"	"BetrSichV"	"BetrSichV"	"BetrSichV"	"Permit according to Betriebssicherheitsverordnung"
6	-	-	-	-	"None!"
+	-	-	-	-	

Fig 6. DMN decision table „Result Procedure

5. Discussion / Conclusion and Future Outlook

The results of the project are twofold. First, aiming at deriving a standardized approach in respect to the approval process of hydrogen filling stations in the model region of Metropolregion Nordwest, a taxonomy delivers a systematic overview regarding potential configurations of hydrogen filling stations. The taxonomy enabled a depiction of the most commonly used setup today. This forms the basis for the ongoing research work to arrive at a standard To-Be process model. Second, looking at our progress in terms of the digitization, the crucial first step in the approval process, in which the main configuration is defined and the resulting legal procedure is identified, can now be supported using an automated decision model. Further research within this project will concentrate on the design of a To-Be process, including a clear representation of the information needed in each process step. The document analysis, therefore is continued and deepened and a mapping with the process steps will finish up the project work. In the area of digitization, the above mentioned mapping that will take place in a two-dimensional table will be translated into a XML-statement that can be directly included into the Camunda-representation of the process model. The modeling approach is planned to be used for the digitization of a variety of other business and administrative processes.

Acknowledgement

The H2BPMM project is funded by the development fund of "Metropolregion Bremen-Oldenburg im Nordwesten e.V.". Substantial co-financing is provided by "Bremerhavener Gesellschaft für Investitionsförderung und Stadtentwicklung mbH (BIS)". The project is currently running from July 2020 to August 2022 and further information can be found at www.wasserstoffprozesse.de.

Bibliography

[At11] Allweyer T. (2011): BPM-Round-Trip: Wunsch oder Wirklichkeit? In: Komus A BPM Best Practice

[Bm21] BMZ (2021): Auswirkungen der Digitalisierung auf die Entwicklungszusammenarbeit. Via: https://www.bmz.de/de/entwicklungspolitik/digitalisierung/auswirkungen-12190 (18.05.2022)

[Bm22] BMWK (2022): Bundesministerium für Wirtschaft und Klimaschutz, Standardisierung. Via: https://www.mittelstand-digital.de/MD/Redaktion/DE/Dossiers/A-Z/standardisierung.html (18.05.2022)

[JD18] Jungmann, T.; Dyck, A. (2018): Standardisierung im Brennstoffzellen- und Wasserstoff-Sektor, HZwei – Das Magazin für Wasserstoff- und Brennstoffzellen, Hydrogeit Verlag, Oktober 2018, Oberkrämer

[KR11] Kaiser, St.; Reichert M. (2011): Datenflussvarianten in Prozessmodellen: Szenarien, Herausforderungen, Ansätze, Institut für Datenbanken und Informationssysteme, Universität Ulm, S. 1 – 67

[Ma03] Marc W. Melaina (2003): Initiating hydrogen infrastructures: preliminary analysis of a sufficient number of initial hydrogen stations in the US, International Journal of Hydrogen Energy, Volume 28, Issue 7, Pages 743-755, ISSN 0360-3199.

[Na22] Nationale Organisation Wasserstoff- und Brennstoffzellentechnologie (2022): Genehmigungsleitfaden für Wasserstoff-Stationen. NOW GmbH, Berlin. Via: https://www.now-gmbh.de/wp-content/uploads/2022/03/NOW_Genehmigungleitfaden_H2-Tankstellen.pdf (15.05.2022).

[Ni20] Niedersächsisches Ministerium für Umwelt, Energie, Bauen und Klimaschutz (2020): Genehmigungsverfahren nach dem Bundes-Immissionsschutzgesetz – Leitfäden für Antragsteller. Via: https://www.umwelt.niedersachsen.de/startseite/themen/technischer_umweltschutz/genehmigungsverfahren/leitfaden_durch_das_bundes_immissionsschutzgesetz/leitfaden-durch-das-bundes-immissionsschutzgesetz-8972.html (07.07.2021).

[No22] NOW GmbH (2022): Regulation, Codes and Standards für den Bereich Wasserstoff- und Brennstoffzellentechnologie. Via: https://www.now-gmbh.de/aktuelles/pressemitteilungen/neue-website-buendelt-wissen-zu-

standardisierung-und-normung-von-wasserstoff-und-brennstoffzellentechnologien (18.05.2022)

[Po20] Portal Green Konsortium (2020): Power-to-Gas-Leitfaden zur Integration Erneuerbarer Energien – Genehmigungsrechtlicher Leitfaden für Power-to-Gas-Anlagen. Via: https://www.dbi-gruppe.de/files/PDFs/Projekte/61_PortalGreen_Genehmigungsleitfaden.pdf (15.05.2022).

[Sm14] Schulte-Zurhausen M. (2014): Organisation 6. Auflage. Verlag Franz Vahlen GmbH, München, Germany.

[SZS04] Sun, S.; Zhao, L.; Sheng, O. (2004): "Data Flow Modeling and Verification in Business Process Management". AMCIS 2004 Proceedings. Paper 508. http://aisel.aisnet.org/amcis 2004/508, S. 4063 – 4073

[SZN06] Sun Sh. X.; Zhao J. L.; Nunamaker J. F.; Liu Sheng O. R. (2006): Formulating the Data-Flow Perspective for Business Process Management. Information Systems Research 17, 4:374-3 91

[SZ06] Sun S. X.; Zhao J. L. (2006): Activity Relations: A Dataflow Approach to Workflow Design. ICIS: 44, S. 660 – 672

[Zf48] Zwicky F. (1948): Morphological Astronomy; In: The Observatory Vol.68 No. 845; NASA Astrophysics Data System; Oxford.

BITS: A Key Performance Indicators (KPIs) supported approach to assess traffic safety for cyclists at intersections in the Netherlands

Johannes Schering [1] and Jorge Marx Gómez[2]

Abstract: Traffic safety is an important factor in the decision process whether people decide to use the bicycle or not. Critical situations that do not lead to an accident are often not reported to the police. To fill this knowledge gap, several regions as the city of Zwolle and the Province of Friesland (Netherlands) have started to detect near accidents at intersections among vehicles and bicycles by 3D camera data to evaluate traffic safety. Four intersections in Friesland and Zwolle were monitored. Different types of intersections (e.g. shared space concept) were considered. Near accidents can be divided into different conflict categories depending on vehicle speed and time to collision (Post-Encroachment Time PET). The preprocessed data including Key Performance Indicators (KPIs) to make cycling safety at the intersections measurable and comparable are provided. Based on the numbers and visualizations, it will be discussed which of the discussed intersections show critical profiles regarding numbers of near accidents, distribution and amount of very critical situations. With the results the intersections can be adjusted to increase traffic safety.

Keywords: Traffic Safety, Near Accident, Post-Encroachment Time, KPIs, Bicycle Data, Cycling

Addresses Sustainable Development Goal 11: Sustainable cities and communities

1. Introduction

Traffic safety is an important factor in the decision process to use the bike or not. The police has a rough overview about situations that lead to severe injured or killed cyclists. As an example, according to official statistics 481 bicycle accidents were registered in the bicycle city of Oldenburg (Germany) in 2020 which is a decrease of 139 accidents compared to 2019 [Po21]. Many accidents especially in the case of not or only slightly injured traffic participants are not registered (e.g. alone accidents [GHS21]). That leads to the problem that the number and the spatial distribution of critical situations that are relevant for the (perceived) safety situation but may not lead to an accident is quite unclear. New knowledge about near accidents could be used to make the traffic situation more safe,

[1] University of Oldenburg, Department of Business Informatics VLBA, Ammerländer Heerstr. 114-118, Oldenburg, 26129 Oldenburg, johannes.schering@uni-oldenburg.de, https://orcid.org/0000-0002-1825-0097 https://uol.de/vlba/personen/mitarbeiterinnen/johannes-schering
[2] University of Oldenburg, Department of Business Informatics VLBA, Ammerländer Heerstr. 114-118, Oldenburg, 26129 Oldenburg, jorge.marx.gomez@uni-oldenburg.de

especially for vulnerable road users (cyclists, pedestrians).

Around 50-60 percent of all inner city bicycle accidents are happening at intersections [SBL80]. These numbers are confirmed by more recently published sources [UV20, Ba13]. Around 25 percent of all killed traffic participants in inner city accidents are cyclists [KLP13]. Accidents and conflicts are often happening when car drivers disregard the priority of a straight moving forward cyclist at a junction, turning to the left or the right (often with severe consequences because of high speeds), traffic participants disregarding red traffic lights or cyclists who suddenly turning to the left [Al88].

As part of this study, the Dutch mobility advisor Mobycon assigned MicroTraffic to make a camera based detection of near accidents at four different intersections. The goal is to improve the infrastructure to avoid severely injured or even killed traffic participants. The MicroTraffic software detects the vehicle speed and number and severity level of near accidents. A risk profile for all types of traffic participants is generated. The aim of this study is to make the four intersections comparable regarding to traffic safety based on vehicle speed and time to collision (Post-Encroachment Time PET). According to the results municipalities may improve the safety situation. That does not mean a rebuilding of the intersection or the roundabout. Traffic safety could also be improved by smaller measures as new markings or adjustments of signals (e.g. green times at traffic lights). As part of a study that was conducted in 2021 on behalf of the Institute for Road Safety Research SWOV (Netherlands), infrastructure measures based on a camera detection campaign of near accidents has contributed to increase traffic safety at two intersections in the city of The Hague. It can be proven by a before-after analysis that several improvements contributed to a decrease of critical situations because more cyclists may pass the intersection in the designated way. Passing behaviour becomes less complex when green spaces prevent crossing the road for vehicles and bicycles at certain spots. The study confirmed that the separation of green phases for (not) motorized traffic is an important factor for traffic safety [WND21].

Chapter two (methodology) gives an overview about the data collection and the near accident definition. In chapter three we will present the four intersections in the Netherlands and its specific circumstances. The most important Key Performance Indicators (KPIs) will be displayed to make the intersections comparable in terms of traffic safety. These numbers are the foundation for the discussion which intersections have the most critical profile regarding the total number and distribution of (high risk) near accidents. In the conclusion and outlook chapter the results are summarized.

2. Methodology

The detection of a near accident in this study is based on the Post Encroachment Time (PET) which is the time between one traffic participant is leaving a conflict area and another conflicting participant arriving at it. Risk levels are based on the relation between PET and the speed level of the traffic participant. The majority of the conflict points

belonging to a low risk level. A smaller number of situations can be perceived in the medium risk area. High risk situations with vehicle speed levels between 35 and 50 km/h are very rare. Critical risk situations were not monitored at the four intersections.

The vehicle speed levels are classified according to the Wramborg biomechanical survivability thresholds for collisions between vehicles and vulnerable road users (cyclists and pedestrians). Ten percent of the cyclists die in a collision with a vehicle with a speed level of 30 km/h. About 80 percent mortality is reached at 50 km/h [Wr05]. The Vulnerable Road Users Safe Systems indicator (VRUss) is a risk indicator for the categorization of the risk level that is based on vehicle speed and the position to a pedestrian or a cyclist. Vehicle speed is used to calculate the time that a vehicle needs to reach a conflict point. The minimum time to a conflict point is extracted from that and connected with the speed level of the vehicle to determine the risk level of a situation according to Wramborg's risk class definition [Wr05]. Tab. 1 shows the classification.

Risk Level	Critical Risk (C)	High Risk (H)	Medium Risk (M)	Low Risk (L)
VRU Conflicts	Gap < 2 seconds Speed > 50 km/h	Gap < 2 seconds Speed > 35 km/h	Gap < 3 seconds Speed > 15 km/h	Gap < 5 seconds Speed < 15 km/h

Tab. 1: VRUss Indicator Risk Thresholds. Source: MicroTraffic.

3. Intersections

3.1 Haadwei and Thiedamawei (Province of Friesland)

The first intersection is located in Damwâld (5.540 inhabitants in 2020 [CBS21]). The administration decided to rebuild two roundabouts at the Haadwei mainroad as intersections with shared space concept. A camera detection was conducted at the rebuilt Haadwei-Thiedamawei intersection at four days in June 2019 between 7 am and 11 pm.

Many critical situations between cyclists and vehicles are happening when cars driving through and leaving the intersection on the north-south axis. The coloured arrows in Fig. 1 indicate the driving direction of the vehicles. A lot of situations are happening when the vehicles are approaching from western direction. At the northern and the southern axis more than 300 critical situations at each side were detected. At the eastern approach more than 200 near accidents were registered during the measurement.

Fig.1: Near Accidents at Haadwei and Thiedamawei. Source: ArcGIS, copyright by Esri.

Vehicle speed levels were divided by the time to collision. The most conflicting situations are happening when vehicles approaching from the northern, southern or western direction. Turning to the left or to the right side is less critical. The risks are quite equal spatially divided to all four corners of the intersection.

According to a report by *BonoTraffic by* driving behaviour at this location is often unpredictable. Because of the barrier in the center, many people interpret the traffic situation still as roundabout what would mean that vehicles in the intersection are prioritized (and not the vehicles approaching from the right side). Citizens complained about high speed levels and dangerous overtaking situations. When traffic volume is low, some left turning vehicles crossing the barrier from the left side. In peak hours cyclists are preferring to use the pedestrian path, especially in the afternoon [LW19].

The Key Performance Indicators (KPIs) are the key to make traffic safety at the intersections comparable. The speed levels are mainly constant on a comparatively higher level with an increase above 20 km/h in the peak hours in the early morning and in the late afternoon/evening. It is the intersection with the highest number of conflicts (1137) what is also reflected by a very high frequency 00:05:03 of near accidents (there is one near accident in about five minutes, see Tab. 2). Although speed levels are high, two third of all near accidents (757) belonging to low risk situations.

Risk Level	Average vehicle speed per period	
Average speed total	19.2 km	
Average count per day	284.25	
Average count per hour	11.84	
	Count & Average speed (according to risk class)	Frequency (according to risk class)
Low	757 (17.79 km/h)	00:07:36
Medium	380 (22.03 km/h)	00:15:09
Total	1137	00:05:03

Tab. 2: KPIs Haadwei & Thiedamawei (Friesland). Source: https://bicycle-data.de/city-analysis/

3.2 Haadwei and Foarwei (Province of Friesland)

Haadwei-Foarwei which is nearby was also rebuilt as a shared space concept. If there are approaching vehicles from more than two sides this can lead to confusion what is especially the case in the afternoon hours. As Fig. 5 confirms, there are more conflict situations in the eastern compared to the western approach. One reason could be that many cyclists who are cycling from north to south are using the eastern side of the intersection. As car drivers do not expect cyclists moving on this side, some near accidents may be registered here [LW19]. There are still traffic safety problems at Haadwei-Foarwei as priorities are not clear to the traffic participants what can also be confirmed by local media reports about a car accident in August 2020 [RTV20].

The detection of near accidents took place at four days in June 2019 from 7 am until 11 pm. There is a conflict hotspot at the southern axis with more than 300 situations especially when cars driving through or turning to the left side (Fig. 2). Vehicles that driving through and leaving to the northern axis are also highly relevant. More than 250 situations were detected in the northern part what is a little bit less compared to the first intersection. Many situations (>250) are happening at the eastern approach.

Fig. 2: Near Accidents at Haadwei and Foarwei. Source: ArcGIS, copyright by Esri.

Conflict levels tend to be lower compared to Haadwei-Thiedamwei. Especially the western and eastern axis show very low risk levels. The northern axis seems to be the most conflicting spot. The southern axis tends to be more conflicting compared to the west and the east. According to speed and PET the whole traffic safety situation seems to be more critical at Haadwei-Thiedamawei compared to Haadwei-Foarwei.

Speed levels are very similar at both intersections. At Haadwei-Foarwei we perceive quite constant speeds on a higher level with average above 20 km/h in the peak hours (morning, late afternoon and also at noon 1-2 pm). The total number of situations (1050) and the frequency (00:05:29) are slightly lower (Tab. 3). However, three high risk situations were detected at the second intersection with an average speed level of more than 38 km/h. Two

of these happened at the northern, another one at the eastern axis. Both Friesland intersections show similar results.

Risk Level	Average vehicle speed per period	
Average speed total	19.79 km	
Average count per day	262.5	
Average count per hour	10.94	
	Count & Average speed (according to risk class)	Frequency (according to risk class)
Low	728 (18.33 km/h)	00:07:54
Medium	319 (22.96 km/h)	00:18:03
High	3 (38.13 km/h)	31:59:40
Total	1050	00:05:29

Tab. 2: KPIs Haadwei & Foarwei (Friesland). Source: https://bicycle-data.de/city-analysis/

3.3 Assendorperstraat and Luttenbergstraat (City of Zwolle)

Zwolle is a medium sized city (128.840 inhabitants in 2020 [CBS21]). According to a 2021 ranking of the American interest group PeopleForBikes, Zwolle is the best place for cycling in the world [TK21]. Assendorperstraat-Luttenbergstraat is an inner city intersection with a typical Dutch road design with partially separated (protection islands) and coloured bike paths at both sides. Space is limited and cannot be broadened.

The detection of the near accidents was conducted at six days in March 2020 from 6 am to 10 pm. The overwhelming majority of the near accidents is happening when vehicles turning to the left from Assendorperstraat to Luttenbergstraat (see Fig. 3). More than 900 (>85 percent) of the 1050 detected situations are happening when cyclists crossing the bike path. Only a very small number of situations occurring at other spots.

Fig. 3: Near Accidents at Assendorperstraat/Luttenbergstraat. Source: ArcGIS, copyright by Esri.

A look to the relation of the speed levels and the time to collision shows a slightly different image. The most critical situation that may occur is still a vehicle that turns from the

Assendorperstraat to the left into the Luttenbergstraat. Although the total number of situations is not high at this location, an increased share of medium risk situations occured when cyclists crossing the Assendorperstraat straight to the Southwest into the Bartjenstraat. Many vehicles are crossing the Assendorperstraat with high speed.

The KPIs in Tab. 4 show that average speed levels are lower (15 km/h) compared to the Friesland cases. Before 2 pm the speed level is below 15 km/h. Later it is slightly increasing up 16.5 km/h. The total number of near accidents (1090, around 181 per day) and the frequency (seven minutes and 55 seconds for one situation) is also lower. The share of medium risk near accidents is higher: Nearly 50 percent of all near accidents (525) are belonging to the medium risk class. The relating frequency of 16 minutes and 27 seconds is close to Haadwei-Thiedamawei (15 minutes and nine seconds, see Tab. 2) and higher compared to Haadwei-Foarwei (18 minutes and three seconds, see Tab. 3). Three high risk situations were registered in the measurement in Zwolle.

Risk Level	Average vehicle speed per period	
Average speed total	15 km	
Average count per day	181.67	
Average count per hour	7.57	
	Count & Average speed (according to risk class)	Frequency (according to risk class)
Low	562 (11.25 km/h)	00:15:22
Medium	525 (18.9 km/h)	00:16:27
High	3 (36.55 km/h)	47:59:40
Total	1090	00:07:55

Tab. 4: KPIs Assendorperstraat & Luttenbergstraat (Zwolle). Source: https://bicycle-data.de/city-analysis/

3.4 Oldeneelallee (City of Zwolle)

The fourth use case of this study is the Oldeneelallee in Zwolle. The location in the Southern part of the city is a multilane road which is interrupted by railway tracks on the eastern side. In the Northern (connection to the Burgermeester de Vos van Steenwijklaan), the Western and the Southern part (connection to the Gouverneurlaan) the intersection can be crossed by cyclists. Many media and police reports about traffic accidents can be found [Al20, vdZ20, RTV18]. The camera detection took place for seven days in late October / early November 2020 from 7 am to 5.30 pm. At the 30st of October the system was only active from 7 am to 2 pm.

More than 40 percent (187) of all critical situations are happening at the northern axis. Here are some problems on the bike path when vehicles turning to the left into the Burgermeester de Vos van Steenwijklaan. The southern axis also shows a lot of conflicts (160). Here is also a problem with vehicles turning to the left leaving the Oldeneelallee into the Gouverneurlaan direction (70). In addition, many situations also occur when the vehicles driving straight from the northern axis and leaving the intersection at the Southern axis (79). At the western side of Oldeneelallee quite a lot of near accidents can be

perceived when vehicles approaching from the northern axis and turning to the right (71). Three high risk situations at the western axis were registered what what makes it to a conflicting spot. The site view seems to be not optimal at this part of the intersection when cyclists want to cross the Oldeneelallee to the Southern direction.

Average vehicle speed levels at Oldeneelallee (17 km/h) are slightly higher compared to Assendorperstraat-Luttenbergstraat but lower compared to the Friesland use case (Tab. 5). The speed level remains constant over the whole day around 16 and 17 km/h. As an outlier, in the late afternoon the speed level is increasing exorbitantly. Compared to the other intersections that were discussed before, the total number (447), the average count per day/hour (63/2-3) and the frequency (22 minutes and 32 seconds) of the near accidents are relatively low. The number (169) and share of low risk situations (37 percent) are the lowest of all four intersections. The distribution of high risk situations is remarkable: More than 60 percent of all near accidents are belonging to medium or high risk situations. Although the total number of conflicts is low, Oldeneelallee is the only intersection that counts five high risk situations including a hotspot. The average speed level (39.45 km/h) is higher compared to the Haadwei-Foarwei case. Oldeneelallee is the only intersection with a share of high risk situations (>one percent). The probability for a critical situation is much higher although near accidents are happening less frequently. It can be concluded that this intersection may have some severe traffic safety issues.

Risk Level	Average vehicle speed per period	
Average speed total	17.09 km	
Average count per day	63.86	
Average count per hour	2.66	
	Count & Average speed (according to risk class)	Frequency (according to risk class)
Low	169 (12.09 km/h)	00:59:38
Medium	273 (19.77 km/h)	00:36:55
High	5 (39.45 km/h)	33:35:48
Total	447	00:22:32

Tab. 5: KPIs Oldeneelalle (Zwolle). Source: https://bicycle-data.de/city-analysis/

4. Conclusion

As part of this study we tried to compare the cycling safety situations at four different intersections in the Netherlands based on a 3D camera data collection based on vehicle speeds and PET. We learned that two intersections in the Province of Friesland with a shared space concept show a high number of situations and high frequencies what means that more near accidents are occurring in the same time. An increased average speed of the vehicles can be perceived. The number of situations is slightly higher at Haadwei-Thiedamawei compared to Haadwei-Foarwei, but the speed level is slightly higher at the second one. The measurement at Haadwei-Foarwei also include some high risk situations

what is surprisingly not the case at the first intersection although this is the place with the highest total number of situations. Both data sets show increased average speed levels in the peak hours in the morning and in the afternoon.

Inner city intersections may show other characteristics regarding traffic safety compared to rural areas. The near accidents at the Assendorperstraat-Luttenbergstraat intersection that was designed in a typical Dutch style are happening with a comparatively low average vehicle speed of 15 km/h. A trend to increasing speed levels in the afternoon peak hours can be perceived. The high majority of the situations is detected when the cars turning to left. Another relevant risk potential was identified when the moving straight forward vehicles approaching the intersection with increased speed levels although these situations do not happen very often.

At the Oldeneelallee by far the lowest number of situations was detected. Only one near accident occurs in 22 minutes. Compared to Haadwei-Thiedamawei it is the other way around: The measurement at the intersection with the lowest number of detected situations include the highest number of high risk situations. The probability for a medium or high risk situation is highly increased at this location. Especially the crossing in the northwestern part of Oldeneelallee where several high risk situations were detected seems to have problems with the road design what is confirmed by local media reports. Similar as at Assendorperstraat-Luttenbergstraat, vehicles are speeding up when going straight through. The site view for cyclists seems to be limited and may be improved to increase traffic safety. After improvements has made, a before and after analysis on the effects of the measures can be realized as the The Hague study by the Institute for Road Safety Research SWOV has tried.

Bibliography

[Al88] Alrutz, D.: Gestaltung von Knotenpunkten für den Radverkehr. In (Walprecht, D. ed.): Fahrradverkehr in Städten und Gemeinden: Planung, Ausbau, Förderung. Heymanns, Köln, pp. 71-84, 1988.

[Ba13] Bakaba, J.E., Unfallforschung kompakt: Sichere Knotenpunkte für schwächere Verkehrsteilnehmer, https://udv.de/de/publikationen/unfallforschung-kompakt/sichere-knotenpunkte-fuer-schwaechere-verkehrsteilnehmer, accessed: 5/12/2021.

[CBS21] CBS, Kerncijfers wijken en buurten 2020, https://www.cbs.nl/nl-nl/maatwerk/2020/29/kerncijfers-wijken-en-buurten-2020, accessed: 5/12/2021.

[GHS21] Gildea, K.; Hall, D; Simms, C., Configurations of underreported cyclist-motorised vehicle and single cyclist collisions: Analysis of a self-reported survey. Accident Analysis & Prevention, Vol. 159, September 2021, 106264, ISSN 0001-4575, https://doi.org/10.1016/j.aap.2021.106264, https://www.sciencedirect.com/science/article/pii/S0001457521002955.

[KLP13] Kolrep-Rometsch, H.; Leitner, R.; Platho, C.; Richter, T.; Schreiber, A.; Schreiber, M.; Buterwegge, P.: Abbiegeunfälle Pkw/Lkw und Fahrrad,

https://digital.zlb.de/viewer/api/v1/records/16305330/files/images/fb_21_ab_pkw_rf.pdf/full.pdf, accessed: 4/12/2021.

[LW19] Lieftink, R.; Wietses, E.: Evaluatie herinrichting Haadwei Damwâld - Gemeente Dantumadiel. Report by BonoTraffic bv. https://www.dantumadiel.frl/sites/default/files/2019-12/191115%20Evaluatie%20herinrichting%20Haadwei%20Damwald-gecomprimeerd.pdf, accessed: 4/12/2021.

[Po21] Polizeiinspektion Oldenburg-Stadt/Ammerland, Die Polizeiinspektion Oldenburg-Stadt/Ammerland gibt die Verkehrsunfallstatistik für das Jahr 2020 bekannt, https://www.presseportal.de/blaulicht/pm/68440/4889307, accessed: 4/12/2021.

[RTV18] RTV Focus Zwolle, Zevende ongeluk in anderhalf jaar tijd op kruising Gouverneurlaan – Oldeneelallee. https://www.rtvfocuszwolle.nl/zevende-ongeluk-in-anderhalf-jaar-tijd-op-kruising-gouverneurlaan-oldeneelallee/amp/, accessed: 6/12/2021.

[RTV20] RTV NOF, Botsing op kruising Haadwei – Foarwei in Damwâld, https://www.rtvnof.nl/botsing-op-kruising-haadwei-foarwei-in-damwald/, accessed: 6/12/2021.

[RTV15] RTV Oost, Twee gewonden bij twee ongelukken op de Assendorperstraat in Zwolle, https://www.rtvoost.nl/nieuws/219433/Twee-gewonden-bij-twee-ongelukken-op-de-Assendorperstraat-in-Zwolle, accessed: 6/12/2021.

[SBL80] Stock, R.; Berr, W.; Linde, R.; Heuber, U.; Klass, W.: Sicherheit für den Radfahrer: Ergebnisse und Schlussfolgerungen aus dem Städtewettbewerb 1980. Bonn, German Federal Ministry of Traffic, 1980.

[TK21] Ten Kleij, E.; Kuik, C., Zwolle door verbeterde rekenfout ‘beste fietsstad ter wereld’ ten koste van Utrecht, https://www.destentor.nl/zwolle/zwolle-door-verbeterde-rekenfout-beste-fietsstad-ter-wereld-ten-koste-van-utrecht~a9db0721/, accessed: 07/12/2021.

[UV20] Unfallforschung der Versicherer, Kreuzung, https://udv.de/de/strasse/kreuzung, accessed: 5/12/2021.

[vdZ20] van der Zee, P., Fietsster gewond na aanrijding met auto in Zwolle-Zuid, https://www.weblogzwolle.nl/nieuws/87406/fietsster-gewond-na-aanrijding-met-auto-in-zwolle-zuid.html, accessed: 07/12/2021.

[WND21] Wijlhuizen, G. J.; Nabavi Niaki, M.; Dijkstra, A.: Evaluatie fietsveiligheid bij herinrichting van kruispunten in Den Haag: Voor-nastudie met geautomatiseerde videoanalyse. Institute for Road Safety Research SWOV, The Hague, 2021.

[Wr05] Wramborg, P.: A New Approach to a Safe and Sustainable Road Structure and Street Design for Urban Areas. In Proceedings of the Road safety on four continents: Warsaw, Poland 5-7 October 2005. Statens väg- och transportforskningsinstitut, Linköping, 2005.

Digital Mobility Services for Communities: Flexible boarding points for campus ridesharing

Moritz Gieza[1], Bernhard Kölmel[2], Thomas Schuster [3] and Lukas Waidelich[4]

Abstract: Mobility is still characterized by individual transport. Despite changes in recent years, it still influences infrastructure development and results in car-friendly cities. As a result, traffic congestion reveals weaknesses in efficiency and sustainability of this model. This is exacerbated in metropolitan areas with high growth rates and in areas with below-average public transport services. Besides congestion, emission such as pollution and noise are a major problem. In this article, we give explain how this affects communities in general and transport from and to our university campus particularly. We will examine how digital mobility services can extend public and individual transport. We will explore how digital services can promote intermodal transport and lead to more sustainability in mobility. Within that discussion, we present a ridesharing platform and study its influence on directions to and from our campus.

Keywords: Sustainable mobility, intermodal mobility, public transport, ridesharing, digital mobility service.

Addresses Sustainable Development Goal 11: Sustainable cities and communities

1. Introduction

Increasing individual mobility and growing freight logistics are leading to congestion and pollution. Especially in larger urban and metropolitan areas, citizens suffer from these effects. Cars are responsible for a large proportion of these problems. The slight pandemic-related decline in individual mobility in the past two years is not permanent and has already been reversed in many cases. Especially in industrialized countries, cars still dominate infrastructure development. The resulting issues are evident by increased total travel time relative to a congestion-free situation (congestion level). Especially in evening rush hour traffic, commuting by car is typically congested. The willingness to change remains low so far. This is the case for citizens as well as for urban development. The focus is currently still on road planning and centralized traffic guidance systems, although

[1] Hochschule Pforzheim, Tiefenbronner Straße 65, 75175 Pforzheim, moritz.gieza@hs-pforzheim.de
[2] Hochschule Pforzheim, Tiefenbronner Straße 65, 75175 Pforzheim, bernhard.koelmel@hs-pforzheim.de
[3] Hochschule Pforzheim, Tiefenbronner Straße 65, 75175 Pforzheim, thomas.schuster@hs-pforzheim.de, https://orcid.org/0000-0002-9539-1627
[4] Hochschule Pforzheim, Tiefenbronner Straße 65, 75175 Pforzheim, lukas.waidelich@hs-pforzheim.de, https://orcid.org/0000-0003-2602-0942

these have contributed little to improvement until now. The promotion of electric mobility can contribute to the reduction of local emissions. Since only the powertrain is changed, this has no influence on the congestion situation. In this article, we will focus on individual mobility in general and its impact on traffic congestion in selected urban areas. We will take a closer look at resulting difficulties (emissions, environment, congestion, and citizens) on site. As one solution approach, we discuss digital mobility services and their potential to change mobility behavior. The latter is outlined by a concrete example of an information and ridesharing platform, which is being used at our university. In this paper we will present ideas on how our ridesharing app could be made even more efficient.

The article is structured as follows: Section two presents related work about mobility behavior. This includes both statistical analysis and digital mobility assistants. In the third section, we explain the challenges arising from these findings and relate them to the context of the situation on site. In Section four, we present the current state of our developed information and ridesharing platform. We will highlight its potential for changes in mobility behavior and present suggestions that we believe might increase the app's usage. In the final section, we summarize main findings and provide an outlook.

2. Mobility Behavior

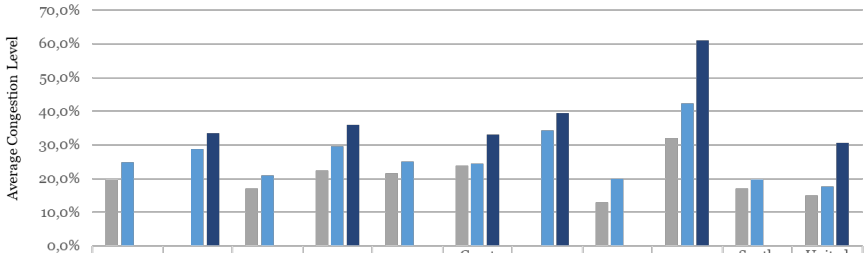
In this section, we examine global mobility developments based on current traffic data. Statistical data on traffic development and mobility behavior are used for this purpose. Our attention is on traffic congestion and intermodal mobility behavior. Known approaches to reduce congestion will be presented and challenges will be addressed.
Especially regions with high growth rates (both in terms of population and economy) suffer from traffic congestion. A tempering of congestion levels was observed in and at the beginning of the pandemic. In 2021, global congestion levels were 10% inferior to the previously observed peaks. However, compared to 2020, the situation has already worsened again. Istanbul (Turkey) recorded the highest congestion level at 62%. Bengaluru (India) peak in 2019 (with 72 %), on the other hand, was able to reduce the congestion level (currently still ranked tenth). In London, citizens spent an average of 148 hours per year in traffic jams, in New York 102 [TT22].

In Fig. 1 we display average congestion levels at country level. Congestion level in regions (cities, metropolitan areas) below 800.000 inhabitants is indicated in gray bars, above 8 million it is shown in dark blue and everything in-between in is shown in light blue. The congestion level is a percentage and describes the longer travel time caused by congestion. We derived the data from TomTom's Traffic Index 2021 [TT22] by aggregation of available city data. Compared to the 2019 survey, congestion levels in Europe and North America have decreased by 7 and 14 percent respectively due to the pandemic. The index does not show data for Chinese regions and cities anymore.

Digital transformation of workplaces and new transport services also mitigated the rise in congestion levels [Pi21]. A transfer to home office, the introduction of digital

conferencing and flexible working hours led to a global shift in peak traffic times in almost 40% of cities.

In Germany, pre-pandemic levels were already almost reached again by 2021. On average, German drivers were stuck in traffic jams for 40 hours nationwide. INRIX estimates the resulting overall economic costs at € 3.5 billion. That is €371 per driver. Congestion and the associated emissions are a challenge not only for the largest German cities (see Tab. 1). In terms of population, smaller municipalities such as Potsdam or Pforzheim show very high congestion statistics [Pi21].



	Australia	Brazil	Canada	France	Germany	Great Britain	Japan	Norway	Russia	South Africa	United States
<800.000	19,8%	n.a.	17,0%	22,4%	21,5%	23,8%	n.a.	13,0%	32,0%	17,0%	14,9%
>800.000	24,8%	28,8%	21,0%	29,5%	25,0%	24,5%	34,4%	20,0%	42,4%	19,7%	17,6%
>8.000.000	n.a.	33,5%	n.a.	36,0%	n.a.	33,0%	39,5%	n.a.	61,0%	n.a.	30,6%

Fig. 1: Congestion Statistics by Country [TT22]

Congestion and modal split statistics show that the car is still the dominant mode of transport [Ec21]. However, since a couple of years we slowly observe changes. For example, the use of alternative modes of transport (e.g., rail, bicycle, and e-scooters) is increasing. The pandemic has shifted the trend to a certain extent now. Public transport lost considerable attractiveness at times, as commuters began to prefer private means of transport again. New forms of mobility gained in popularity. The use of e-scooters and bicycles is on the rise, often also supported by newly constructed bicycle paths. Many changes can be attributed to the so-called sharing economy and Internet-based mobility services [SW19]. Young people are increasingly using multiple modes of transport. They show a multi- or even intermodal mobility behavior. Multimodal mobility behavior is evident when people switch their main mode of transportation for different trips (e.g., different trips within a week). Intermodal transportation refers to a combination of multiple modes of transportation during a single trip, e.g., a public transportation trip combined with a bicycle or car. Moreover, different modes of transport are used in relation to specific situations (context of the trip and personal preferences) [Ec21, We16]

The reduction in monomodal vehicle use is correlated with age and affinity for Internet services. The introduction of sharing concepts and Internet use are also correlated. Internet-based services show higher growth rates in countries that promote alternative modes of transport [Eu19]. However, it is still not clear today whether there is also a longer-term and causal relationship in this respect. It remains an open question whether

these people change their behavior again when they get older or whether they maintain their attitude. Other important influencing factors could be income level and other personal preferences. These typically change with age as well. Individuals make their choice of transport mode depending on factors such as purpose, route, type of trip, origin, and destination of their trip, as well as personal preferences.

Rank	City	Time loss [in hours]	Residents [in million]	Cost per driver	Costs for the Municipality [in million]
1	Munich	79	~ 1,49	740 €	388 €
2	Berlin	65	~ 3,66	610 €	823 €
3	Hamburg	47	~ 1,85	443 €	291 €
4	Potsdam	46	~ 0,18	428 €	27 €
5	Pforzheim	44	~ 0,12	410 €	18 €
6	Dusseldorf	43	~ 0,62	400 €	88 €
7	Cologne	42	~ 1,08	398 €	152 €
8	Nuremberg	41	~ 0,52	382 €	71 €

Tab. 1: Traffic Congestion & Impacts, Germany [Pi21]

Much work also exists in this area for freight transportation. Optimal routing results in reduced fuel consumption and thus reduced emissions (e.g., carbon dioxide emissions). This includes the idea of rerouting based on the current traffic situation and pre-planning based on the traffic forecast. Routing based on real-time traffic information assumes that traffic congestion can be reduced through intelligent route planning [APL91]. Alternative means of transportation were also proposed [CRS09, GM13]. On the other hand, route optimization is subject to a saturation effect as soon as the information is used by a certain number of users [MC91]. Therefore, routing algorithms are additionally contextualized. For instance through parking models [AP04, FP13] or ad-hoc change of transport modes [CMJ13, GM13]. Besides intelligent routing and information systems, there is another psychological aspect. Users tend to accept suggestions once their preferences are met. These include user interface design, transportation mode (for ad hoc decisions), cost, approximate delay, driver comfort and flexible boarding points.

In terms of new mobility services, Mobility-as-a-Service (MaaS) is proving particularly popular. These are based on Internet and app offerings as well as on the principles of the sharing economy [SW19]. To establish MaaS, public and private providers usually rely on distributed information systems to organize rides. Users usually use the service via apps or web applications. Participation is thus linked to the use and availability of Internet services. The services are mostly subject to usage-based charges (pay-per-use). Costs are usually assessed per trip, according to duration of use and distance, or depending on the current traffic situation. There are currently many different MaaS. To increase attractiveness, many services take personal preferences into account to improve offers and planning. A lot of the services in the sharing space are designed to be unimodal (such as Uber, Didi, Lyft, Lime, Zeus, Bird). In the sharing sector, a basic distinction can be made

between ridesharing (private trips) and rideselling (commercial trips) [MS17]. Ridesharing is the provision of rides to private vehicles. The ride can but does not have to be remunerated.

In the case of rideselling, the rides are organized commercially. Rideselling can be subdivided into ridehailing (for individual transportation, e.g., cabs, rental cars, Lyft or Uber Pop) and ridepooling (collective transportation via trip bundling, e.g., Uber Pool, Clever Shuttle or MOIA).

3. Current Status and Challenges

Congestion and resulting delays as well as loads characterize people's everyday lives today. With high growth rates, this situation will continue to worsen. This requires new solutions to improve the overall mobility situation. To reduce congestion levels and the environmental pollution caused by traffic, the following fields of action emerge:

1. Service offerings to meet the demand for mobility
2. Transport infrastructure modernization
3. Reduction of environmental pollution (e.g., particulate matter, carbon dioxide emissions, noise)

The modernization of transportation infrastructure is part of urban or regional transportation planning and management. Different methods aim to improve transportation planning based on predictive analytics and optimization methods [CK19]. The reduction of environmental impacts should be a fundamental goal along all measures. This can be achieved particularly through the choice of propulsion technology, trip avoidance, and route optimization according to ecological criteria.

Consider more closely the opportunities and challenges of deploying MaaS in our city /region and in relation to our university. The city has a population of about 125,000, the surrounding county about 200,000, and the region just over 600,000. The university has about 6,200 students enrolled and 880 people (professors, adjunct professors, and staff) employed. The university has buildings at three different locations in the city and a main campus which about 88% of students primarily attend. It is located on a hillside on the outskirts of the city. A shopping center and leisure facilities (zoo, city forest and restaurants) are located nearby [Ri19]. The approach is mainly by private motorized transport (car, motorcycle) and public transport (bus). There is a direct connection to the main railway station of the city by the public bus. From there, it is possible to get to the surrounding towns and the nearest larger cities that are part of the university's catchment area. There are 524 (319 for students) free parking spaces directly at the university. In addition, 300 parking spaces of the adjoining zoo may also be used. The total capacity is 824 spaces. A highway connection close by favors access by private automobile additionally. At present, a major infrastructure project on the freeway is having a negative

impact on congestion. A nearby dorm encourages students to walk or bike to campus as well. However, the strong slope (and a difference of about 100 meters in altitude compared to the city center) and lack of bike lanes prevent many students and staff from using this option for longer commutes.

Preliminary studies of student mobility demands showed transportation peaks as expected. These are in the morning at 8 a.m., at noon at 1 p.m. and in the evening from 5 p.m. onward. During these times, there is insufficient transport capacity in the public transport system. The peak demand occurs on average on Wednesday mornings at 8 a.m. (over 1,800 students). This creates an acute shortage of parking spaces (even outside of peak demand) and results in parking traffic having to move to neighboring residential areas. A lack of parking spaces is a permanent source of dissatisfaction from the perspective of students and residents. The review reveals that about 1,400 students live in the city (daily mobility demand for travel to and from the city about 12,000 person-km (pkm), 1,100 students live outside the city up to 20 km (mobility demand 41,000 pkm/d), and another 750 students live up to 40 km (50,000 pkm/d) away. The remaining students live at distances greater than this. In total, this results in a daily student mobility demand of about 130,000 pkm. 50% of students use cars, 40% use public transport (bus and train), 10% use other means of transport. The measurement of the intensity of use of student cars showed a car occupancy of just under 1.2 persons per car. With the mentioned daily mobility demand and an assumed CO_2 emission of 0.12 kg CO_2/km, the daily CO_2 emission by the students traveling individually is arithmetically 6.5 t/d. The emissions from the mobility of employees and those arriving by public transport or other means of transport must also be added. In a direct comparison, the mobility requirements of the university thus exceed those of the municipal administration (approx. 2500 employees) and presumably also those of the larger companies in the city. The main challenges relate to the following points:

1. Inefficient utilization of vehicles of individual motorized transport (1.2 persons per car).
2. Increased congestion (the city is generally one of the most congested in Germany) and the university's contribution to this.
3. High emission levels due to daily directions (CO_2, particulate matter, and noise)
4. Conflicts caused by parking in the surrounding area.
5. Issues created by searching for parking (cruising).
6. Return to classroom teaching increases pressure on issues 1- 5.

4. Community Service Development

The university has defined sustainability goals. This includes a 50% reduction in total emissions by 2030. Regarding traffic and directions, the following measures are intended

to contribute to this objective:

1. Development of a new mobility culture including innovative mobility offers.
2. Reduction of individual transport via e.g., bundling/sharing approaches.
3. Reduction of emissions through zero-emission drives in individual and public transport, e.g., through electromobility.
4. Avoiding trips by offering a modern range of services.

We conducted three steps to prepare and develop these measures. First, an investigation of the current approach situation was carried out (see above). Then, alternative arrival concepts were investigated and designed. Finally, the design and implementation of MaaS was started by development of a ridesharing platform. The ridesharing platform was deployed in winter semester 2021 but was only used briefly (over a two-week period) due to the pandemic and the resulting online classes. An overview of the ridesharing platform is visualized inFig. 2.

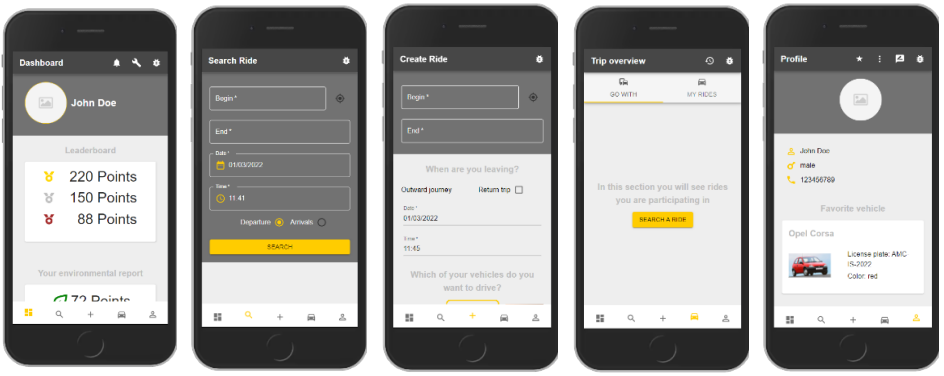


Fig. 2: Ridesharing platform user interface

The ridesharing platform allows users to search for and offer rides. To increase the possibility of user participation, the associated app is implemented as a progressive web app. This makes it possible to use it on any smartphone. No app store is required, users simply access the app via a browser. In addition, the use of the app is also completely unrestricted on any tablet or desktop PC. The app consists of five main areas: 1) Home; 2) Ride search; 3) Ride offer; 4) Ride overview; 5) User profile. The areas are designed according to common usability criteria and partially feature a tile layout. A dashboard serves as the home, presenting the user with currently relevant information (area 1). In addition to the information, the dashboard implements a gamification concept. Users can see how much carbon has been avoided and who has currently contributed the most. This allows students and staff to receive vouchers for the cafeteria. The second area (2) provides a search dialog to find offered rides and to request rides. Once a user has requested a ride,

the provider can confirm or decline it. In case of agreement on the rideshare, both sides receive ride-relevant information (e.g., the rider receives information about the vehicle). The third space (area 3) is used for the administration of the trip offers. Ride offers can be created as one-time or regular rides. An overview of pending and past rides is provided in the fourth area (4). Users see which rides they are offering, who is riding with them or where they are riding themselves, and the ride's status. The last area serves to manage the user profile (area 5). Users who want to offer rides themselves must also enter a vehicle. In addition, ratings received from other users for rides are viewed here. Feedback can be submitted after a ride has been completed and, as well as a points system, also allows a comment to be entered. A review can be submitted anonymously or personalized.

The ridesharing platform was well received after a promotional campaign (see Fig. 3). As a result, 764 users have registered in the system in a short period of time. Users can simply use their university user account for this purpose. In the two weeks leading up to the renewed pandemic lockdown, 303 trips were scheduled with an average trip duration of 28 min and 6 seconds (variance: 5 minutes and 17 seconds, standard deviation: 17 minutes and 50 seconds). In line with expectations, most of the trips were offered on Wednesdays. The exact distribution on the weekdays was: Monday: 53, Tuesday: 64, Wednesday: 87, Thursday: 51, Friday: 33, Saturday: 4, and Sunday: 11. Due to the planned trips and results in 0.12 kg CO_2/km there is already a saving of about 1.1 t CO_2.

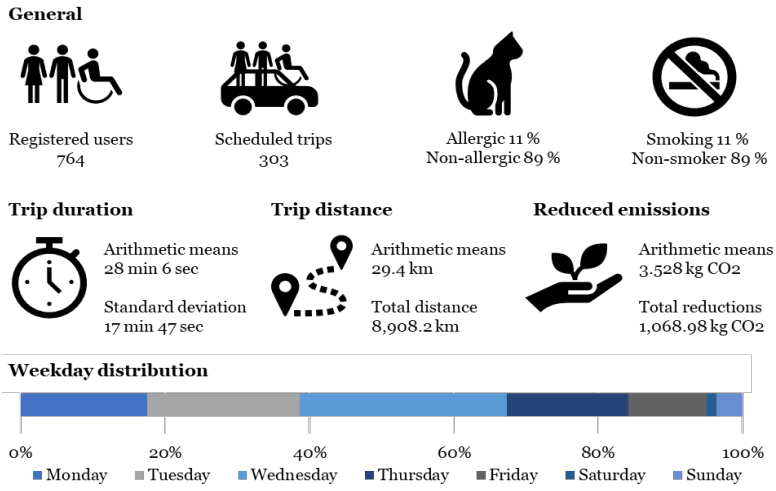


Fig. 3: Ridesharing platform usage patterns

Since the initial deployment of the ridesharing platform experienced a very positive user response, its planned for further and extended use upcoming terms. By returning to classroom teaching, we expect increased and continuous use. This will make a sustainable contribution to the university's emissions goals. The short evaluation period has already attracted attention from outside the university. City officials intend to offer the platform to city employees as well. A corresponding project has been launched.

We see a lot of potential to evolve the state of the ridesharing platform further. The ridesharing platform is aimed at bundling travel needs to the university in a flexible manner, enabling alternatives and improving efficiency (vehicle utilization, congestion situation). Dynamic boarding points are to be generated automatically and integrated into the routing of ridesharing trips. By using the generated entry points, the existing platform shall be enhanced by adaptive route selection. Thus, efficiency and attractiveness of routes and ridesharing shall be increased. To this end, the following points are to be addressed in the future: First, the improvement of the ridesharing application including validation through A/B testing with users. Second, the analysis of the access routes and identification of the main constraints. Third, the identification, generation and testing of dynamic boarding points. Fourth, to improve traffic planning and routing in the application.

5. Conclusion

Traffic development and congestion is still an enormous problem worldwide. This has a negative impact on the environment and on individuals. Problems also arise economically, e.g., due to time losses and stress. Numerous research projects and studies are addressing the problem and aim to reduce traffic congestion. The usage of ubiquitous MaaS can be a promising approach. MaaS provide an opportunity to explore and use alternatives, especially to motorized individual transport. Information when to use which service properly is an important factor in achieving behavioral changes. In many cases, awareness of alternative mobility services is almost as important as their availability.

Our ridesharing platform demonstrated that digital planning and collaboration amongst users can create a considerable amount of interest and behavioral changes. The evaluation over a longer period was not yet possible due to the pandemic. Hence, this represents a consistent next step in our research. It will include evaluating the success of the gamification aspects of the application. Key questions will be which factors will create long lasting changes and how to make ridesharing experience increasingly sustainable.

Furthermore, we want to supplement the application with additional services. We intend further development of intelligent route planning, so that both routes can be optimized according to ecological criteria and intermediate stops can be planned as optimal as possible. The latter involves individual travel routes of platform users as well as traffic of user people (who should not be impacted by stops and onboarding of ridesharing users). Finally, the application will in future draw greater attention to intermodal route chains.

Bibliography

[TT22] TomTom International BV, https://www.tomtom.com/en_gb/traffic-index/ranking/, accessed February 9, 2022.

[Pi21] Pishue, B.:2021 INRIX Global Traffic Scorecard. INRIX Research, Kirkland, 2021.

[Ec21] Ecke, L.; Chlond, B.; Magdolen, M.; Vallée, J.; Vortisch, P.: Deutsches Mobilitätspanel (MOP). Wissenschaftliche Begleitung und Auswertungen Bericht, 2020/2021, Alltagsmobilität und Fahrleistung, Karlsruhe, 2021.

[SW19] Schuster, T.; Waidelich, L.: Internetbasierte Mobilitätsdienste – Ein Lösungsansatz zur effizienten individuellen Mobilität? In Journal für Mobilität und Verkehr, 2, pp. 50–58, 2019.

[We16] Weiß, C.; Chlond, B.; Hilgert, T.; Vortisch, P.: Deutsches Mobilitätspanel (MOP). Wissenschaftliche Begleitung und Auswertungen, Bericht 2014/2015, Alltagsmobilität und Fahrleistung, Karlsruhe, 2016.

[Eu19] Eurostat Passenger Transport Statistics - Statistics Explained, http://ec.europa.eu/eurostat/statistics-explained/index.php/Passenger_transport_statistics, accessed May 1, 2022.

[APL91] Arnott, R.; de Palma, A.; and Lindsey, R.: Does Providing Information to Drivers Reduce Traffic Congestion? In Transportation Research Part A: General, 25:5, pp. 309–318, 1991.

[MC91] Mahmassani, H. S.; Chen, P. S.-T.: Comparative Assessment of Origin-Based and En-Route Real-Time in-Formation under Alternative User Behavior Rules. In TRANSPORTATION RESEARCH RECORD. 1306, pp. 69–81, 1991.

[MS17] Mehlert, C.; Schiefelbusch, M.:Mobility On-Demand: Disruption oder Hype - Entwicklung und Zukunft von Rufbus, Sharing Und Robotaxi. In Nahverkehr, 7+8:35, pp. 6–12, 2017.

[CK19] Iliopoulou, C., Kepaptsoglou, K.: Combining ITS and optimization in public transportation planning: state of the art and future research paths. European Transport Research Review. 11, 27 (2019). https://doi.org/10.1186/s12544-019-0365-5.

[Ri19] Richter, A.; Waidelich, L.; Kölmel, B.; Bulander, R.; Glaser, P.; Proske, M., Brügmann, S.: Digitalisation and Future Challenges in Rural Areas: An Open Innovation Based Research. In Proceedings of the 16th International Joint Conference on E-Business and Telecommunications, Prague, SCITEPRESS, pp. 147–153, 2019.

[CRS09] Crainic, T. G.; Ricciardi, N.; Storchi, G. 2009: Models for Evaluating and Planning City Logistics Systems. In Transportation Science, 43:4, pp. 432–454, 2009.

[GM13] Giannikas, V.; McFarlane, D.: Product Intelligence in Intermodal Transportation: The Dynamic Routing Problem. In Dynamics in Logistics, H.-J. Kreowski, B. Scholz-Reiter, and K.-D. Thoben (eds.), Berlin, Springer Berlin Heidelberg, pp. 59–69, 2013.

[AP04] Anderson, S. P.; de Palma, A.: The Economics of Pricing Parking. In Journal of Urban Economics, 55:1, pp. 1–20, 2004.

[CMJ13] Caris, A.; Macharis, C.; Janssens, G. K.: Decision Support in Intermodal Transport: A New Research Agenda. In Computers in Industry, 64:2, pp. 105–112, 2013.

[FP13] Fosgerau, M.; de Palma, A.: The Dynamics of Urban Traffic Congestion and the Price of Parking. In Journal of Public Economics, 105, pp. 106–115, 2013.

R3 – Resilient, Regional, Retail: Implementation of SusCRM Approach within a Local Retail Platform

Sarankumar Haridas[1] , Richard Schulte[2] , Mattes Leibenath[3] and Benjamin Wagner vom Berg[4]

Abstract: Current consumption patterns in offline and online retailing pose challenges in terms of sustainability aspects such as overcoming the last mile, and product-related and energy-related environmental impacts. In the last decades, customers have changed their mindset from money to sustainability which includes e-Commerce neglecting the greenhouse gas emissions which is linked to the last mile delivery. Alternative choices of delivery method and individual mobility are resulting in different emissions. To bridge the gap this work investigates how to apply a SusCRM (Sustainability Customer Relationship Management) to a local retail platform to meet both customer expectations and sustainable and conscious consumption. An overview of potential incentive systems for implementation within a local e-marketplace is presented.

Keywords: SusCRM, Incentive model, Last mile, Energy Management, Retail, Sustainability, Marketplace

Addresses Sustainable Development Goal 8: Decent Work and economic growth

1. Introduction

1.1 Motivation

The increment in innovation gives great openings to the dealer to reach the client in a much quicker, less demanding, and financial way. E-commerce is tremendously increasing in the last few years. In recent days the retail sector is booming with help of the internet. Thousands of people shifting towards online shopping. On the contrary, still people are going to the local retail shop and purchasing their things which is very important for the local retail shop to survive [Rh22]. Many consumers decide to buy things in a local shop

[1] Hochschule Bremerhaven, Smart Mobility Institute, An der Karlstadt 8, Bremerhaven, 27568, sarharidas@smail.hs-bremerhaven.de
[2] Hochschule Bremerhaven, Smart Mobility Institute, An der Karlstadt 8, Bremerhaven, 27568, rschulte@hs-bremerhaven.de
[3] Hochschule Bremerhaven, Smart Mobility Institute, An der Karlstadt 8, Bremerhaven, 27568, mleibenath@studenten.hs-bremerhaven.de
[4] Hochschule Bremerhaven, Smart Mobility Institute, An der Karlstadt 8, Bremerhaven, 27568, benjamin.wagnervomberg@hs-bremerhaven.de, https://orcid.org/0000-0001-6976-9625

to have a look at them and acquire ownership of the item only after paying for it. Due to digitalization, the Retail sector is constantly growing and this change is a well-known phenomenon across the retail industry [HK18]. People who are convenient choose the online shopping platform. In the report of PWC, around 56% of German people choose online shopping platform because of their convenience [PW18].

Contrastingly around 35% of german people said cost was the only reason for them to move towards online shopping. This effect was clearly shown in 2020 global e-Commerce sales. Several studies have already been conducted on the environmental impact of online retailing, but only a limited number of these studies address the environmental impact of FMCG products ordered online. These studies are often limited in scope, comparing an online retail system with traditional shopping and ignoring important impacts such as associated consumer travel. Failed deliveries, product returns, trip chaining (to what extent is the shopping trip embedded in another transport destination, e.g. commuting to and from work, going to school, etc.) and rummage trips (a trip to check out a product before buying it online) can be an important contributor to environmental impacts (estimated as a carbon footprint) [EJ09].

Businesses, politicians, and the public are increasingly realizing that our resources are limited and that numerous problems in the economic, social, and ecological dimensions of our society result from our economic activity and consumption. In its groundbreaking strategy paper on corporate social responsibility, the EU Commission sees it as a central task to support consumers in sustainable consumption behavior [BP21]. The biggest CO_2 polluters in the EU are road transport with a share of almost 27%, which is quite a lot [AE21].

Growing CO_2 emissions are seen as the main cause of global warming and related climate change. The Intergovernmental Panel on Climate Change has formulated the 2-degree target to reduce greenhouse gases in order to prevent climate change and its consequences. The mobility sector plays a central role in terms of energy consumption and CO_2 emissions and is one of the most relevant implementation areas for sustainable consumption. Mobility has strong negative impacts here in both the ecological and economic dimensions of sustainability. Canzler and Knie exaggerate the role of mobility in the CO_2 problem as follows: Transport is the big problem for a successful decarbonization strategy [CK16].

If we look at the transport mix, the so-called modal split, we see that a large part of our spatial mobility is still realized by motorized individual transport today, while at the same time this usually has the most negative effects in terms of sustainability. While the annual growth rates in the Federal Republic of Germany have even been declining in some cases since the beginning of the 21st century, this does not apply to the global perspective. Especially in emerging countries such as India, the number of cars is increasing due to further industrialization and the broader prosperity of the population, and the associated negative impacts are also increasing rapidly. This results in an acute need for action.

1.2 Background

Within the project "R3 - Resilient, Regional, Retail in the Metropolitan Region Northwest", a platform for regional retail and a sustainable and competitive supply and logistics structure is to be conceptually developed that strengthens regional retail against large online platforms. The project started on June 1st, 2021, and is planned for two years. The project is being carried out by the University of Applied Sciences Bremerhaven while the applicant for Funding by the Metropolregion Nordwest was the Erlebnis Bremerhaven GmbH (society for marketing and tourism for the strengthening of the city of Bremerhaven / who are also a marketing partner). The project is supported by a large number of partners, ranging from companies, associations, societies, cities, municipalities, districts, and the state of Bremen. In the project, a platform is first designed with the various stakeholders and then engineered in software terms. Competitive advantages such as sustainability and regionality, competent consulting as well as delivery and pick-up services are going to be integrated within the platform. In this platform, a SusCRM-approach is implemented, in which stakeholders, such as consumers, retailers, institutions, and logistics providers shall be incentivized to more sustainable acting.

The platform approach aims to lead to more orders from local retailers, who are participating in the local marketplace. Therefore, an increase in traffic resulting from retail last-mile logistics and individual mobility is expected. Sustainable aspects of the last mile have to be evaluated regarding the different delivery options, as well as the individual mobility to local retailers.

2. State of Art

In today's world, customers have so many options to choose what they want to buy. Because of digitalization, they have enough pathways to purchase what they desire to buy. For example, due to the introduction of mobile commerce and e-commerce, there is a lot of rivalry between different channels and because of this, we need to realize what makes customers buy from one channel to another. The main research direction identified explores the evaluation of product sustainability.

In recent years, consumers have had an unprecedented choice when buying goods and services. Research has examined the impact of consumer choices on carbon consumption, greenhouse gas emissions, and climate change. The research also shares the consensus, that consumer decisions are carbon-emitting. However, some consumer choices, decisions by logistics and supply chain providers, transport, and even technical solutions contribute to more greenhouse gas emissions [DF20]. Our aim is to understand the environmental footprint of consumer’s retail behavior. As the purchase and delivery of goods has expanded to multiple channels, we want to measure the role of retail, transport, and energy consumption. The aim is to help those involved in the delivery and receipt of goods to minimize their carbon footprint [MZ17].

A study by the German Federal Environment Agency compares stationary and online shopping. While a 5 km shopping trip by car causes 600-1,100 g of CO_2, the CO_2 emissions for going to a local shop by bicycle or by walk are zero. But when it comes to online shopping it emits around 200-400 g of CO_2. These low emitted methods are related to car driving where it has better space to allocate things, utilization of car to the fullest, proper transportation routes, and use of EV mode of transportation. The most important aspect of a retail shop is the usage of energy and mode of transport. The immense disadvantage in online shopping is package waste and last-mile delivery. [Ub22]

Over the past quarter-century, the retail sector has changed a lot driven by technological advances, supply chain innovations, and ever-changing consumer behavior. While consumers used to flock to suburban malls and brick-and-mortar shops in cities, e-commerce - combined with the possibility of next-day delivery - has completely changed the way we buy all kinds of goods. In parallel with this change, climate change has reached a point globally where it is undoubtedly impacting our environment, our economy, and our resilience as a society. The question we now need to ask ourselves is which of these forms of consumer behavior will lead to lower carbon emissions and be better for the world we live in [DF20]. The energy efficiency of commercial buildings is becoming more and more important due to the annually increasing number of commercial buildings [MZ17].

The strategic orientation of companies towards their customers and the associated maintenance of customer relationships is often viewed in a purely economic dimension. The main objective of "Sustainable Customer Relationship Management" (SusCRM) is to extend this traditional approach with an ecological and social dimension, which can provide the customer with alternative options for more sustainable consumption.

3. Methods for Sustainable approach in Retail store

SusCRM can be understood as any CRM approach that supports the continuous development of inter-company and customer relationships, considering the environmental, social and economic value of stakeholders and third parties involved. SusCRM strategies can identify sustainability values and goals as key differentiators for customer relationships and present a CRM process designed to increase loyalty as well as to motivate its target group to interact in a more sustainable business model. [BW15]

Within the approach an incentive model for the realization of the SusCRM strategy is implemented. This incentive model introduces measures, which can incentivize the consumer to adopt a more sustainable shopping and mobility behavior. Each incentive can be assigned to one of the four categories:

1. Information-based incentives
2. Game and competition-based incentives
3. Social incentives

4. Reward-based incentives

Based on Wagner vom Berg, information-based incentives include all those incentives that lead to a change in the customer's shopping behavior by passing on information to the customer. At this point, however, it must be stressed that the mere provision of information on sustainability is not sufficient for a change in behavior. It must be supplemented by the identification of attractive alternatives and further incentives for action.

In the case of social incentives, the motivation for changing behavior arises from the "expected reciprocity". This limits the individual's desire to elevate his own status in the community to satisfy his need for social attention.

The effectiveness of game- and competition-based incentives is based on a person competing with himself or with others. The performance of the others and/or the self-imposed performance limits are intended to motivate a person extrinsically or intrinsically to improve his own performance in a certain area.

In addition to the information-, game- and competition-based incentives already described, as well as the social incentives, reward-based incentives are also to be integrated within a local retail platform. Here, the focus is on material and non-material rewards in order to influence the purchasing behavior of end customers. Material rewards can be divided into monetary, such as credits and discounts, and non-monetary rewards (e.B. vouchers). Non-material rewards do not influence behavior through material value but relate more strongly to social incentives. Motifs to take part in a SusCRM approach can be intrinsic, extrinsic, and extrinsic-instrumental [BW15].

In this project, the following action areas can be identified to be suitable for the implementation of a SusCRM-approach within a local marketplace scenario: energy management in retail stores, use of sustainable products, and sustainable last-mile approaches.

4. Application of SusCRM Approach within a Local Marketplace

4.1 Energy Management in Retail Store

Savings in the environmental impact of a retail shop's energy consumption can be achieved by improving energy efficiency and using renewable energy sources. For example, choosing a renewable energy plan can reduce CO_2 emissions from energy consumption from 1.2 tonnes to 0.08 tonnes per year in a 2400 kWh consumption scenario, a reduction of over 90 percent. [Cw22]. Energy efficiency in buildings means the ability of buildings to reduce energy consumption by using efficient technologies or methods. Proper planning for the buildings is needed to achieve maximum efficiency. Retrofitting buildings is one of the practices of energy management in buildings (e.g. replacing window frames) [GE11]. Awareness of energy consumption and the behavior of users (e.g. owners,

managers, tenants, and customers) can also influence the energy efficiency of buildings. There are different technologies for energy saving, such as energy-saving lamps, BEMS (Building Energy Management System) or BAS (Building Automation System), and renewable energy systems (e.g. photovoltaics and energy storage). Sustainable solutions such as decentralized energy generation (e.g. wind turbines or PV solar panels) are installed in buildings for various reasons, e.g. power outages [HM15].

4.2 Use of Sustainable Products

Sustainability is increasingly becoming a standard of living for many people. Not only companies are changing their processes, but individuals are also changing their lifestyles. Given the state of our planet, it is imperative that manufacturers, businesses, and individuals work together to protect the earth's natural resources. A sustainable lifestyle reduces your carbon footprint and improves the entire supply chain by reducing energy consumption, protecting valuable ecosystems, and reducing pollution. We can work towards sustainability by making responsible choices for the products we select for ourselves and our families. The current economic model is still based on the "take-make-replace" principle. It depletes our resources, pollutes our environment, and damages biodiversity and the climate. It also makes Europe dependent on resources from other countries. To solve these problems, the EU wants to move to a more circular economic model based on more sustainable products. Some of the approaches, which can likely be used in the project's platform approach, are as follows [Cc22]:

- Avoid products with excessive packing
- Buy products with certified Eco-labels
- Use products made from recycled materials
- Buy environmentally friendly cleaners
- Use products that are biodegradable
- Buy products that use less water
- Buy products that reduce energy consumption

4.3 Sustainable Last Mile

The surge in e-commerce volumes has put enormous pressure on the last mile delivery system, the process by which products are transported from distribution centers to the end consumer. The COVID -19 pandemic has only accelerated this trend as more and more people demand contactless forms of shopping. Some of the sustainable last-mile approaches are as follows [TT22]:

- Market the green delivery and pick-up option

- Pack delivery more efficiently
- Explore sustainable packing options
- Joint delivery system, including networking sorting center

For the delivery option, we still have a sustainable option to deliver the product for example use of EVs and bicycles helps us to deliver in time as well as in a sustainable way. Nowadays packing stations have attracted a lot of consumers, which helps consumers to collect their products whenever they are free. Customers also have their part to play in this entire sustainable option. They have to be very careful in their approach to buying products. They have to purchase environmentally friendly products, reuse most of the online buying products, and particularly purchase with proper sustainable packing. We don't have to always buy a new product. In recent days people started to use second-hand products which are more sustainable and cost-effective.

4.4 Approaches for SusCRM Implementation in a Local Marketplace

The literature proposes different strategies and instruments that can be implemented by retailers to engage customers for more sustainable acting with their shopping behavior. A local marketplace approach needs to adapt from those methods to be able to promote sustainability in both online and offline sales. Studies in this field identify different approaches to retail marketing interventions [Ba21]. And also provides sustainable choices, staging shopping experiences that enable consumers to make sustainable choices in store

1. Editing choices in favor of sustainable consumption
2. Educating consumers on sustainable consumption
3. Informing consumers about sustainability-related aspects
4. Promoting sustainable shopping and consumption behavior
5. Getting third party certifications for sustainability.

Those approaches need to be shaped for use inside a local marketplace. The implementation of SusCRM with a marketplace provides the opportunity to influence both the suppliers and demand side regarding sustainable behavior. Local marketplaces especially allow to impact the consumers choice of individual mobility traveling to store by implementing incentive models to lead the consumer to a more sustainable choice. Retailers offering sustainable products and services can be promoted within the marketplace to give them a competitive advantage and incentivize sustainable acting. In the following Tab. 1 a selection of incentives is proposed to adopt the SusCRM strategies to a local marketplace:

Incentive system	Example
Information-based incentive	Comparison of the environmental impact of different delivery and pick-up options Certificates and labels for sustainable products, services and energy use
Game and competition-based incentive	Ranking of consumers regarding sustainable shopping behaviour Ranking of retailers regarding sustainable offers and acting
Reward-based incentive	Credit for no return of orders; credits for choosing sustainable alternatives Preference for sustainable retailers in product and store searches
Social incentive	Individual sustainability report that compares to other shoppers and retailers

Tab. 1: Examples of potential incentive systems suitable for marketplace integration

The majority of retailers say they currently participate in flagship retail coalitions, and despite the fact that 3% of them say coalitions have had a really hard impact on their Supportability efforts, there is still a solid belief that collaboration is essential. Around 90% of the supportability groups surveyed accept that engagement and collaboration with peers and others in the community will be critical supporters in moving forward on supportability [SH22]. There are clear opportunities for retailers to work together to avoid exposing problems. Collaborative activities can be particularly useful in creating a standardized approach to vendor metrics, definitions, and information levels across the retail industry. Retailers can then use this information to set baselines, make progress, engage suppliers and educate their customers on climate-friendly products. Other industries have already shown positive results based on this approach. A local retailing marketplace holds the opportunity to create an ecosystem for sustainable acting within retailing by through reciprocal effect on the stakeholders involved.

5. Future Outlook

The proposed project has the objective of implementing SusCRM approaches to contribute to sustainable consumption. Since the platform includes both offline and online sales channels, instruments must be found that, in addition to possible delivery options, also influence personal mobility. Furthermore, the other two big contributors to environmental issues resulting from energy and product-level have to be considered.

The implementation of incentive models needs to be executed within the design and conception of the marketplace. Suitable options for the integration of those instruments have to be evaluated in the sense of acceptance of the different stakeholders (customers,

retailers, logistics service providers). When talking about information-based incentives it has to be ensured that provided information is scientifically approved and communicated in a comprehensible way. This means that different delivery and pick-up alternatives need to be evaluated regarding their actual impacts as well as product-related information such as certificates and labels. In the sense of more sustainable energy management executed by retailers, instruments to inform customers about the application need to be found and evaluated. Further research also provides a valuation framework that can be used to assess the values of exact emission of CO_2 for the overcoming of the last mile.

Bibliography

[AE21] Annual European Union greenhouse gas inventory 1990–2019 and inventory report 2021, pg 69

[Ba21] Balan, C.: How Does Retail Engage Consumers in Sustainable Consumption? A Systematic Literature Review. Sustainability 2021,

[BP21] BP Statistical Review of World Energy 2021

[BW15] vom Berg, B.W.: Konzeption eines Sustainability Customer Relationship Management (SusCRM) für Anbieter nachhaltiger Mobilität. Shaker Verlag, Aachen (2015)

[Cc22] Corpcofe., https://www.corpcofe.com/the-impact-of-sustainable-products/ : accessed 13.05.2022

[CK16] Weert Canzler & Andreas Knie (2016) Mobility in the age of digital modernity: why the private car is losing its significance, intermodal transport is winning and why digitalisation is the key, Applied Mobilities, 1:1, 56-67.

[Cw22] Calculated with CO2 emission calculator of the German Federal Environmental Agency: https://uba.co2-rechner.de/de_DE/: accessed on 20.05.2022

[DF20] Diego, FB., Chegut A., Retail Carbon Footprint Measuring impacts from real estate and technology 2020

[EJ09] Edwards, J.B., McKinnon, A.C., 200 A shopping trip or home delivery: which has the smaller carbon footprint. Logistics. Transp. Focus 11, 20-24

[FJ06] Freyssinier, J. P., Frering, D., Taylor, J., Narendran, N., and Rizzo, P. 2006. “Reducing Lighting Energy Use in Retail Display Windows.” pp. 63371L-11.

[GE11] Gordo, E., Campos, A., and Coelho, D. 2011. “Energy Efficiency in a Hospital Building Case Study: Hospitais da universidade de Coimbra.” In *Proceedings of the 2011 3rd International Youth Conference on Energetics (IYCE)*, pp. 1-6.

[HK18] Henning Kagermann and Winter, J. (2018). The Second Wave of Digitalisation. [online] ResearchGate. Available at: https://www.researchgate.net/publication/344106566_The_Second_Wave_of_Digitalisation [Accessed 14 May. 2022].

[HM15] Hijjo, M., Bauer, P., Felgner, F., and Frey, G. 2015. “Energy Management Systems for Hospitals in Gaza-Strip.” In *2015 IEEE Global Humanitarian Technology Conference*

(GHTC), pp. 18-25.

[MZ17] Ma, Z., Billanes, J. D., Kjærgaard, M. B., and Jørgensen, B. N. 2017. "Energy Flexibility in Retail Buildings: From a Business Ecosystem Perspective." In *2017 14th International Conference on the European Energy Market (EEM)*, Dresden, Germany, p. 6.

[PW18] PWC (2018). They say they want a revolution [online] PWC. Available at: https://www.pwc.com/gx/en/retail-consumer/publications/assets/total-retailglobal-report.pdf [Accessed 12 May. 2022].

[Rh22] Repos.,https://repos.hcu-hamburg.de/bitstream/hcu/131/1/Ebert_Kirsten.pdf, p.53: accessed 20.05.2022

[SH22] Shalini Unnikrishnan et al.,"Sustainability in retail is possible but there's work to be done" In April 22 World retail congress article.

[TT22] Tech Target., https://www.techtarget.com/searcherp/feature/10-strategies-for-more-sustainable-last-mile-delivery: accessed 16.05.2022

[Ub22] Umweltbundesamt., https://www.umweltbundesamt.de/en/press/pressinformation/climate-impact-of-online-retail-stationary-trade: accessed 17.05.2022

GI-Edition Lecture Notes in Informatics

P-299 M. Gandorfer, A. Meyer-Aurich, H. Bernhardt, F. X. Maidl, G. Fröhlich, H. Floto (Hrsg.)
Informatik in der Land-, Forst- und Ernährungswirtschaft
Fokus: Digitalisierung für Mensch, Umwelt und Tier
Referate der 40. GIL-Jahrestagung
17.–18. Februar 2020,
Campus Weihenstephan

P-300 Michael Felderer, Wilhelm Hasselbring, Rick Rabiser, Reiner Jung (Hrsg.)
Software Engineering 2020
24.–28. Februar 2020
Innsbruck, Austria

P-301 Delphine Reinhardt, Hanno Langweg, Bernhard C. Witt, Mathias Fischer (Hrsg.)
Sicherheit 2020
Sicherheit, Schutz und Zuverlässigkeit
17.–20. März 2020, Göttingen

P-302 Dominik Bork, Dimitris Karagiannis, Heinrich C. Mayr (Hrsg.)
Modellierung 2020
19.–21. Februar 2020, Wien

P-303 Peter Heisig, Ronald Orth, Jakob Michael Schönborn, Stefan Thalmann (Hrsg.)
Wissensmanagement in digitalen Arbeitswelten: Aktuelle Ansätze und Perspektiven
18.–20.03.2019, Potsdam

P-304 Heinrich C. Mayr, Stefanie Rinderle-Ma, Stefan Strecker (Hrsg.)
40 Years EMISA
Digital Ecosystems of the Future: Methodology, Techniques and Applications
May 15.–17. 2019
Tutzing am Starnberger See

P-305 Heiko Roßnagel, Christian H. Schunck, Sebastian Mödersheim, Detlef Hühnlein (Hrsg.)
Open Identity Summit 2020
26.–27. May 2020, Copenhagen

P-306 Arslan Brömme, Christoph Busch, Antitza Dantcheva, Kiran Raja, Christian Rathgeb, Andreas Uhl (Eds.)
BIOSIG 2020
Proceedings of the 19th International Conference of the Biometrics Special Interest Group
16.–18. September 2020
International Digital Conference

P-307 Ralf H. Reussner, Anne Koziolek, Robert Heinrich (Hrsg.)
INFORMATIK 2020
Back to the Future
28. September – 2. Oktober 2020, Karlsruhe

P-308 Raphael Zender, Dirk Ifenthaler, Thiemo Leonhardt, Clara Schumacher (Hrsg.)
DELFI 2020 –
Die 18. Fachtagung Bildungstechnologien der Gesellschaft für Informatik e.V.
14.–18. September 2020
Online

P-309 A. Meyer-Aurich, M. Gandorfer, C. Hoffmann, C. Weltzien, S. Bellingrath-Kimura, H. Floto (Hrsg.)
Informatik in der Land-, Forst- und Ernährungswirtschaft
Referate der 41. GIL-Jahrestagung
08.–09. März 2021, Leibniz-Institut für Agrartechnik und Bioökonomie e.V., Potsdam

P-310 Anne Koziolek, Ina Schaefer, Christoph Seidl (Hrsg.)
Software Engineering 2021
22.–26. Februar 2021,
Braunschweig/Virtuell

P-311 Kai-Uwe Sattler, Melanie Herschel, Wolfgang Lehner (Hrsg.)
Datenbanksysteme für Business, Technologie und Web (BTW 2021)
Tagungsband
13.–17. September 2021,
Dresden

P-312 Heiko Roßnagel, Christian H. Schunck, Sebastian Mödersheim (Hrsg.)
Open Identity Summit 2021
01.–02. Juni 2021, Copenhagen

P-313 Ludger Humbert (Hrsg.)
Informatik – Bildung von Lehrkräften in allen Phasen
19. GI-Fachtagung Informatik und Schule
8.–10. September 2021 Wuppertal

P-314 Gesellschaft für Informatik e.V. (GI) (Hrsg.)
INFORMATIK 2021Computer Science & Sustainability
27. September– 01. Oktober 2021, Berlin

P-315 Arslan Brömme, Christoph Busch, Naser Damer, Antitza Dantcheva, Marta Gomez-Barrero, Kiran Raja, Christian Rathgeb, Ana F. Sequeira, Andreas Uhl (Eds.)
BIOSIG 2021
Proceedings of the 20th International Conference of the Biometrics Special Interest Group
15.–17. September 2021
International Digital Conference

P-316 Andrea Kienle, Andreas Harrer, Jörg M. Haake, Andreas Lingnau (Hrsg.)
DELFI 2021
Die 19. Fachtagung Bildungstechnologien der Gesellschaft für Informatik e.V.
13.–15. September 2021
Online 8.–10. September 2021

P-317 M. Gandorfer, C. Hoffmann, N. El Benni, M. Cockburn, T. Anken, H. Floto (Hrsg.)
Informatik in der Land-, Forst- und Ernährungswirtschaft
Fokus: Künstliche Intelligenz in der Agrar- und Ernährungswirtschaft
Referate der 42. GIL-Jahrestagung
21. - 22. Februar 2022 Agroscope, Tänikon, Ettenhausen, Schweiz

P-318 Andreas Helferich, Robert Henzel, Georg Herzwurm, Martin Mikusz (Hrsg.)
FACHTAGUNG SOFTWARE MANAGEMENT 2021
Fachtagung des GI-Fachausschusses Management der Anwendungsentwicklung und -wartung im Fachbereich Wirtschaftsinformatik (WI-MAW), Stuttgart, 2021

P-319 Zeynep Tuncer, Rüdiger Breitschwerdt, Helge Nuhn, Michael Fuchs, Vera Meister, Martin Wolf, Doris Weßels, Birte Malzahn (Hrsg.)
3. Wissenschaftsforum:
Digitale Transformation (WiFo21)
5. November 2021 Darmstadt, Germany

P-320 Lars Grunske, Janet Siegmund, Andreas Vogelsang (Hrsg.))
Software Engineering 2022
21.–25. Februar 2022, Berlin/Virtuell

P-321 Veronika Thurner, Barne Kleinen, Juliane Siegeris, Debora Weber-Wulff (Hrsg.)
Software Engineering im Unterricht der Hochschulen SEUH 2022
24.–25. Februar 2022, Berlin

P-322 Peter A. Henning, Michael Striewe, Matthias Wölfel (Hrsg.))
DELFI 2022 Die 20. Fachtagung Bildungstechnologien der Gesellschaft für Informatik e.V.
12.–14. September 2022, Karlsruhe

P-323 Christian Wressnegger, Delphine Reinhardt, Thomas Barber, Bernhard C. Witt, Daniel Arp, Zoltan Mann (Hrsg.)
Sicherheit 2022
Sicherheit, Schutz und Zuverlässigkeit
Beiträge der 11. Jahrestagung des Fachbereichs Sicherheit der Gesellschaft für Informatik e.V. (GI)
5.–8. April 2022, Karlsruhe

P-325 Heiko Roßnagel, Christian H. Schunck, Sebastian Mödersheim (Hrsg.)
Open Identity Summit 2022
Fachtagung vom 07. - 08. July 2022, Copenhagen

P-327 Masud Fazal-Baqaie, Oliver Linssen, Alexander Volland, Enes Yigitbas, Martin Engstler, Martin Bertram, Axel Kalenborn (Hrsg.)
Projektmanagement und Vorgehensmodelle 2022
Trier 2022

P-328 Volker Wohlgemuth, Stefan Naumann, Hans-Knud Arndt, Grit Behrens, Maximilian Höb (Editors)
Environmental Informatics 2022
26. - 28. September 2022, Hamburg, Germany

All volumes of Lecture Notes in Informatics can be found at https://dl.gi.de/handle/20.500.12116/21.

The titles can be purchased at:

Köllen Druck + Verlag GmbH
Ernst-Robert-Curtius-Str. 14 · D-53117 Bonn
Fax: +49 (0)228/9898222
E-Mail: druckverlag@koellen.de

Gesellschaft für Informatik e.V. (GI)

publishes this series in order to make available to a broad public recent findings in informatics (i.e. computer science and information systems), to document conferences that are organized in cooperation with GI and to publish the annual GI Award dissertation.

Broken down into
• seminars
• proceedings
• dissertations
• thematics
current topics are dealt with from the vantage point of research and development, teaching and further training in theory and practice. The Editorial Committee uses an intensive review process in order to ensure high quality contributions.

The volumes are published in German or English.

Information: http://www.gi.de/service/publikationen/lni/

ISSN 1617-5468
ISBN 978-3-88579-722-7

The 36th annual conference of the Environmental Informatics took place from September 26th. to September 28, 2022 in Hamburg, with a strong focus on sustainability. Combining and shaping national and international activities in the field of applied informatics and Environmental Informatics, the EnviroInfo conference series aims at presenting and discussing the latest state-of-the-art development on information and communication technology (ICT) and environmental-related fields. A special focus of the conference was on the question how Environmental Informatics can help to reach the Sustainable Development Goals of the UN and which goals are explicitly addressed by the Environmental Informatics Community.