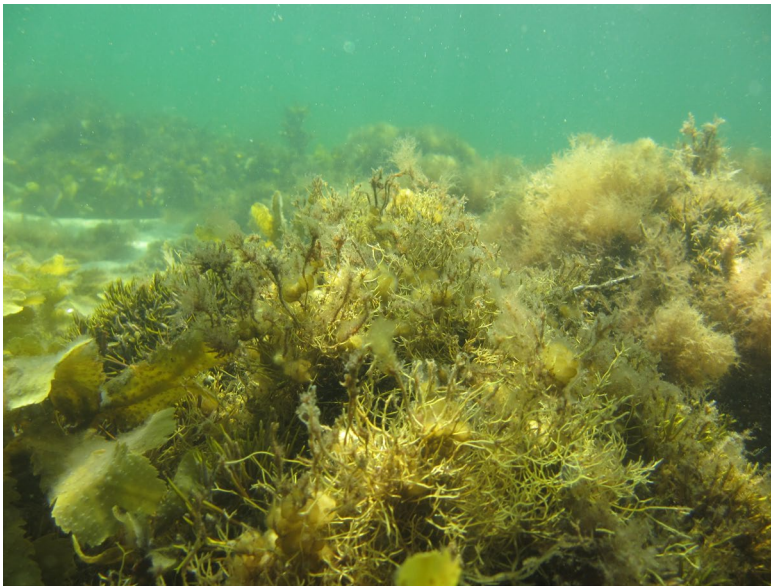


Analysis of historical macroalgae data

Scientific briefing from DCE – Danish Centre for Environment and Energy

Date: 27. April 2022 | **32**



AARHUS
UNIVERSITY

DCE – DANISH CENTRE FOR ENVIRONMENT AND ENERGY

Tel.: +45 8715 0000
E-mail: dce@au.dk
<http://dce.au.dk>

Data sheet

Scientific briefing from DCE – Danish Centre for Environment and Energy

Category: Scientific briefing

Title: Analysis of historical macroalgae data

Authors: Jacob Carstensen & Signe Høgslund
Institution: Department of Ecoscience

Referee: Karsten Dahl, Department of Ecoscience
Quality assurance, DCE: Anja Skjoldborg Hansen
Linguistic QA: Anne Mette Poulsen

External comment: The Danish Environmental Protection Agency. The comments can be found here:
http://dce2.au.dk/pub/komm/N2022_32_komm.pdf

Claimant: The Danish Environmental Protection Agency

Please cite as: Carstensen, J. & Høgslund, S. 2022. Analysis of historical macroalgae data. Aarhus University, DCE - Danish Centre for Environment and Energy, 35 pp. – Scientific briefing no. 2022|329
https://dce.au.dk/fileadmin/dce.au.dk/Udgivelser/Notater_2022/N2022_32.pdf

Reproduction permitted provided the source is explicitly acknowledged.

Front page photo: Jens Deding

Number of pages: 35

Content

1. Introduction	4
1.1 Background	4
2. Historical and NOVANA monitoring macroalgae data	6
2.1 Distribution of data across depth	7
2.2 Harmonisation of taxonomy	7
3. Descriptive analyses of macroalgae communities	9
3.1 Investigation effort in historical data	9
3.2 Comparing macroalgae communities	9
3.3 Comparing species richness	13
3.4 Comparing depth limits	15
3.5 Potential of historical data for describing reference conditions	18
4. Conclusion	19
5. References	20
Annex A: Species richness versus depth	21
Annex B: Species-specific depth distributions	28

1. Introduction

This scientific briefing examines a data set of historical macroalgae observations from Danish coastal and open waters. As the majority of the observations represent a period with relatively low human disturbance, the aim is to investigate the potential for utilising these observations to set reference conditions for macroalgae communities and diversity. The data set has been compiled in a previous project (Høgslund et al. 2018), and this study comprises an elaborate in-depth analysis of the data, including an assessment of the quality and validity of the historical data. Specifically, the following aspects are considered:

- Q1: Is it possible to assess the investigation effort connected with the historical observations, acknowledging that species richness increases with investigation effort?
- Q2: Is the historical macroalgae community similar to that found in the contemporary monitoring data?
- Q3: Is the historical macroalgae species richness similar to that found in the contemporary monitoring data? For the entire community and different taxonomic groups?
- Q4: Are depth limits for specific species similar between the historical and present-day data?
- Q5: Is it possible to describe reference conditions, both qualitatively and quantitatively, for the situation around 1900 using historical macroalgae data? If so, do these conditions correspond to the values obtained through modelling of K_{bio} (macroalgae attenuation coefficient) and Secchi depth?

This briefing is made on request from the Danish Environmental Protection Agency (Miljøstyrelsen) and is one of three concurrent scientific briefings concerning the development of macroalgae indicators applicable to environmental assessment of Danish coastal waters. The accompanying briefings are: "Literature review of general responses of macroalgae to light, nutrient, salinity and temperature variations relevant to Danish waters" and "Macroalgae indicators for assessing ecological status in the Baltic and North East Atlantic".

1.1 Background

Macroalgae and angiosperms comprise one of the biological quality elements (BQE) for assessing ecological status according to the European Water Framework Directive (WFD). In Denmark, the status of this benthic vegetation BQE is based on angiosperm depth limits only as an assessment protocol for macroalgae has not yet been accepted. Two macroalgae indicators, based on the changes in cumulative cover and the number of perennial species with depth, are suggested (Carstensen 2020), with reference conditions based on relationships between their attenuation with depth (K_{bio}) versus light attenuation (K_d) using reference conditions for K_d inferred from historical eelgrass depth limits. These proposed reference conditions for macroalgae indicators

rely on the assumption that the macroalgae community, in terms of diversity and functional composition, has not changed from around 1900 to the present.

Today, macroalgae communities are monitored by diver assessments along transects that typically extend from near-shore sites to the deepest locations with hard substrate. For depth-specific locations along the transect, the macroalgae community is quantitatively assessed by estimating the specific cover relative to the amount of suitable substrate for each species present. When the taxonomy of specimens cannot be determined by the diver, a sample is brought back to the laboratory and investigated under microscope; this has only been done for stone reefs, however. The macroalgae index cumulative cover represents the sum of all species-specific covers, except crust-forming algae that are not consistently assessed. The macroalgae index for the number of perennial species represents the species richness of erect macroalgae species. The attenuation of these two indices with depth constitutes the Kbio indicator. Monitoring of macroalgae communities has followed the same standard operating procedures since 2001.

Historically, macroalgae communities in Danish waters have been described by scientific pioneers such as Ostenfeldt and Rosenvinge (Høgslund et al. 2019). The earliest reports in the historical data set are from 1805, and the majority of the observations were made around 1900. The aim of these early investigations was primarily to identify and describe species and not necessarily to characterise the entire community. Hence, the historical data set can provide information about the presence of a species at a given location, but the absence of a species does not imply that it was not present. Supplementary information (e.g. depth, stratum) was occasionally gathered along with the macroalgae observations, although this practice did not become common until the second half of the 20th century. Details on the historical data can be found in Høgslund et al. (2019). In summary, the historical macroalgae data represent qualitative rather than quantitative observations.

2. Historical and NOVANA monitoring macroalgae data

The historical data set consists of 31,917 species-specific observations that are geographically scattered over Danish coastal waters and primarily obtained by dredging (Høgslund et al. 2019). For 599 of these observations, no exact time point (year or date) exists, but by combining information on the observers and taxonomists behind these observations with the time-referenced observations, an approximate time estimate can be made. However, as 47 observations did not have any connection with other observations allowing an estimation of the most likely period of the observation, they were removed from the analyses. For 28,733 observations, sampling depth was indicated, although consideration must be taken regarding their interpretation. For instance, sampling depth may have been assessed using a sandbag on a line from the ship or maybe even the length of the dredging line itself, which could result in a deeper depth registration than the vertical distance to the sampling location. The bottom topography can be quite variable, which makes it difficult to assess the depth from a dredging haul. Moreover, macroalgae material is typically deposited at deeper depths near the hard substrate that they are attached to, and the dredging approach cannot distinguish between attached and fresh floating macroalgae, implying that the depth registrations of the historical data should be interpreted with caution (Nielsen and Dahl 1992). Hence, occurrence of macroalgae at sampling depths exceeding 50 m was considered unrealistic, and a few other observations had sampling depths exceeding the depths of the location. This resulted in 28,304 observations with sampling depths. All species-specific observations were associated with a locality and an approximate location, acknowledging the fact that for many observations the locality description and the coarse spatial resolution (degree and minute) did not allow for an exact evaluation of the sampling point. For 625 species observations, the macroalgae were not attached to substrate (identified as floating macroalgae), but these observations were nevertheless included as they were assumed to represent specimens from the local macroalgae community (see modifications below). Similarly, 1874 species were recorded as epiphytes; these were, however, included in the analyses since they represent the local macroalgae community and are generally not obligate epiphytes. Moreover, epiphytic macroalgae are also recorded in the monitoring data. For assessing the macroalgae community under conditions representing reference conditions, only data from before 1930 were used (19,151 observations) as increasing nutrient inputs thereafter may have disturbed the macroalgae community. Almost 14,000 of these observations (~73%) were made by L.K. Rosenvinge between 1890 and 1927.

Macroalgae monitoring data (1989-2020) were extracted from the Danish national monitoring database (ODA), resulting in almost 650,000 taxonomical registrations. These data contained a considerable amount of inappropriate taxonomical classifications such as 'Bangiophyceae indet.' or 'brown brush', which are incomparable with the historical data. After deleting observations that did not specify taxonomy to at least genus level, 612,359 observations were used for comparison with the historical data.

The macroalgae recordings in both data sets were associated with their WFD water bodies and assigned to areas that corresponded approximately with the Danish open water typology (OWs), representing a salinity gradient spanning

six areas from Bornholm, the western Baltic Sea (Hjelm Bugt, Fakse Bugt, Køge Bugt and Øresund), the southern Belt Sea (Lillebælt, Sydfynske Øhav, southern Storebælt, Smålandsfarvandet), the Kattegat (including northern Storebælt, Samsø Belt, Aarhus Bugt and Sejerøbugten), the Skagerrak and the North Sea (including the Wadden Sea).

2.1 Distribution of data across depth

Almost half of the historical observations were from the surface (0-2 m) and with an almost even distribution of observations between 2 and 10 m, followed by a decreasing number of observations at deeper depths (Figure 2.1). Observations deeper than 30 m (n=117) were made around Bornholm (30-40 m), Lillebælt (Snævringen, 35-40 m), the North Sea (near Thyborøn, 31 m), the Skagerrak (near Hanstholm) and Storebælt (Langelandsbælt). The observations at Thyborøn include 40 observations of different macroalgae species that were all recorded at the same depth (31 m). Similarly, all observations at Hanstholm (n=10) were from the same depth (34 m). Most likely, these observations are from stone reefs, and the depth represents the deepest location and not the actual depth of the macroalgae. Moreover, it is also questionable if the macroalgae in Snævringen were attached to substrate or detached macroalgae accumulating at the bottom. The historical studies did not aim at establishing depth limits for macroalgae species, and caution is therefore needed when interpreting the depths associated with macroalgae observations.

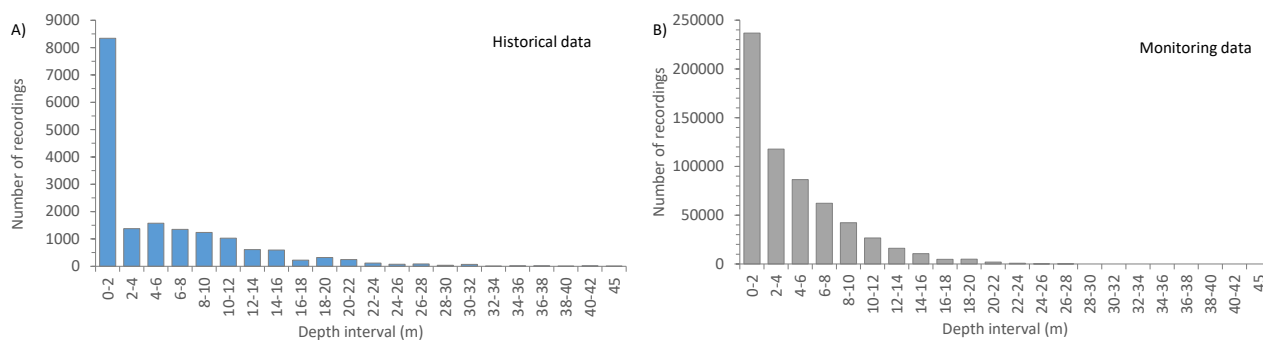


Figure 2.1. Distribution of macroalgae recordings across different depth strata for historical data (A) and recent monitoring data (B).

In the monitoring data, the macroalgae observations were also primarily from shallow depths, with the number of recordings decreasing almost exponentially with depth (Figure 2.1). The deepest recording of macroalgae was at 26.5 m from stone reefs in the Kattegat.

2.2 Harmonisation of taxonomy

Høgslund et al. (2019) matched the original data sources against the Global Biodiversity Information Facility (GBIF). In their study, all species names in the historical data set and from the monitoring data (ODA, 1989-2020) were matched against the World Register of Marine Species (WoRMS; <https://www.marinespecies.org/>) in January 2022 for harmonisation. In total, 11 species names came out as ambiguous, and the suggested name by WoRMS was chosen (*Callophyllis cristata* as *Euthora cristata*, *Ceramium fruticosum* as *Ceramium penicillatum*, *Ceramium nodulosum* as *Ceramium virgatum*, *Ceramium robustum* as *Ceramium rosenvingii*, *Ceramium secundatum* as *Acrochaetium secundatum*, *Ceramium strictum* as *Ceramium tenuicorne*, *Ectocarpus confervoides* as *Ectocarpus siliculosus*, *Ectocarpus fasciculatus* as *Ectocarpus flagelliformis*, *Eudorina elegans* as *Eudorina unicocca*, *Laurencia pinnatifida* as *Osmundea pinnatifida*, *Ralfsia*

clavata as *Stragularia clavata*). Furthermore, three species names were uncertain according to WoRMS without a suggestion for an alternative (and more correct) species name (*Ceramium virgatum*, *Coccomyxa ophiuræ*, *Ralfsia verrucosa*).

All identifications at higher taxonomical levels than genus (only monitoring data) were discarded. This resulted in 489 distinct taxonomical classifications at either species or genus level for the historical and monitoring data combined. Out of these, 31 taxonomical classifications were at genus level with only one species recorded in all data combined, and it was therefore assumed that recordings at genus level represented the (one and only) species (*Acrothrix* sp. as *Acrothrix gracilis*, *Audouinella* sp. as *Audouinella hermannii*, *Blastophyssa* sp. as *Blastophyssa rhizopus*, *Bolbocoleon* sp. as *Bolbocoleon piliferu*, *Capsosiphon* sp. as *Capsosiphon fulvescens*, *Chordaria* sp. as *Chordaria flagelliformis*, *Chroodactylon* sp. as *Chroodactylon ornatum*, *Chylocladia* sp. as *Chylocladia verticillata*, *Cruoria* sp. as *Cruoria pellita*, *Cutleria* sp. as *Cutleria multifida*, *Cystoclonium* sp. as *Cystoclonium purpureum*, *Delesseria* sp. as *Delesseria sanguinea*, *Derbesia* sp. as *Derbesia marina*, *Erythrocladia* sp. as *Erythrocladia irregularis*, *Gracilaria* sp. as *Gracilaria gracilis*, *Harveyella* sp. as *Harveyella mirabilis*, *Hecatonema* sp. as *Hecatonema terminale*, *Heterosiphonia* sp. as *Heterosiphonia plumosa*, *Lemanea* sp. as *Lemanea fluviatilis*, *Melobesia* sp. as *Melobesia membranacea*, *Mesogloia* sp. as *Mesogloia vermiculata*, *Monostroma* sp. as *Monostroma grevillea*, *Myriocladia* sp. as *Myriocladia lovenii*, *Myriotrichia* sp. as *Myriotrichia clavaeformis*, *Petalonia* sp. as *Petalonia fascia*, *Peyssonnelia* sp. as *Peyssonnelia dubyi*, *Pilinia* sp. as *Pilinia rimosa*, *Pylaiella* sp. as *Pylaiella littoralis*, *Pyropia* sp. as *Pyropia peggicovensis*, *Rhizoclonium* sp. as *Rhizoclonium riparium*, *Rhodochorton* sp. as *Rhodochorton purpureum*, *Spermothamnion* sp. as *Spermothamnion repens*, *Spongomorpha* sp. as *Spongomorpha aeruginosa*).

After consultation with the macroalgae taxonomical expert Ruth Nielsen, Copenhagen University, some further modifications were made. Five species in the historical data (*Alaria esculenta*, *Callithamnion granulatum*, *Himanthalia elongata*, *Saccorhiza polyschides*, *Vertebrata lanosa*) were removed as these are not indigenous species in Danish waters and only observed as floating macroalgae originating from the macrotidal southern North Sea. *Aorainvillea riukuensis* was observed once in Præstø Fjord in 1897, but this is presumably a misidentification. *Devaleraea ramentacea* was observed once in Kalundborg Fjord in 1879, but this was probably also a misidentification. *Ulvella lens* was observed twice at stonereefs in the Kattegat – also a misidentification. The same applies to *Nitophyllum punctatum*, which was also registered twice at stonereefs in the Kattegat. *Codium tomentosum* is the old name for *Codium fragile*. *Laurencia* sp. is today registered as *Osmundea oederi*. *Desmotrichum* sp. and *Lithoderma* sp. are today registered as *Punctaria* sp. and *Pseudolithoderma* sp. *Myrionema reptans* is most likely microthallus of *Asperococcus fistulosus*. *Myriophyllum alterniflorum* and *Myriophyllum spicatum* are angiosperms and were removed. *Batrachospermum gelatinosum*, *Bulbochaete* sp., *Hildenbrandia rivularis*, *Lemanea fluviatilis* and *Oedogonium* sp. were removed as they are freshwater species. *Eudorina unicocca* is a phytoplankton species and was removed (in the monitoring data). *Monostroma balticum* is highly questionable and most likely *Monostroma grevillea*. Finally, the following species were pooled as there is only one species within the genus (*Bangia atropurpurea* is *Bangia fuscopurpurea*, *Compsothamnion gracillimum* is *Compsothamnion thuioides*, *Osmundea pinnatifida* and *Osmundea truncate* are *Osmundea oederi*, *Plocamium cartilagineum* is *Plocamium lyngbyanum*).

As a consequence of the merging, 427 taxonomical classifications remain; out of these 261 occur in both the historical and the recent monitoring data, 98 occur in the historical data only, and 68 occur in the monitoring data only.

3. Descriptive analyses of macroalgae communities

The research questions outlined in the introduction will be addressed point by point, bearing in mind that the two data sets, the historical and the recent monitoring data, are not strictly comparable.

3.1 Investigation effort in historical data

The historical dataset is a mosaic of different investigations performed by different persons with different research objectives and investigation methods. The investigation effort is therefore highly variable.

The historical data from the period before 1930 were frequently collected with the objective of answering scientific questions concerning morphological characterisation, taxonomy and life history traits. Thus, systematic mapping of species occurrence and abundance was often not the main target of the investigation, and exhaustive species lists from a given investigation can therefore not always be expected.

L. Kolderup Rosenvinge's research was a notable exception to his. He aimed at making "... as uniform an investigation of the Danish waters as possible" and "... note all species present at a dredging" (Kolderup Rosenvinge 1909).

The investigations by Rosenvinge were undertaken on board different ship types: fishery control steamers, life-saving steamers, fishing boats and research ships. Although Rosenvinge aimed at a uniform investigation of Danish waters, it is reasonable to assume that this influenced the sampling intensity in the different geographical areas.

The historical investigations were restricted to collection of algae by different type of dredges. Dredging as a sampling method is difficult to standardise compared to today's methods of diving and video recording, and this also adds to the uncertainty regarding the investigation effort in the historical data.

Q1: Is it possible to assess the investigation effort of the historical observations, acknowledging that species richness increases with investigation effort?

Answer: No, the historical data are composed of different investigations undertaken with different methodologies and research targets, and no thorough sampling protocol is available for the individual investigations. A standardised measure of investigation effort can therefore not be made.

3.2 Comparing macroalgae communities

The majority of the taxa found in the historical data only (n=98 species) were often sporadically observed (less than 5 times in total), a pattern that was more or less the same for all groups (Figure 2.2). It should also be noted that many of the species recorded only in the historical data are (according to Ruth Nielsen) tiny, mostly epiphytic and difficult to identify at species level (e.g. *Botrytella micromora* is probably recorded as brown bush in the monitoring data). Moreover, many of such species are probably not known or easily overlooked by divers and

therefore not registered in the monitoring data. For the historical data set as a whole, 95.8% of the recordings were species found also in the monitoring data, and only 4.2% of the recordings were species found only in the historical data. The proportion of recordings within the common taxonomy was lowest for Chlorophyta (93.0%) and highest for Phaeophyta (98.1%), whereas it was comparable to the overall average for Rhodophyta (94.8%).

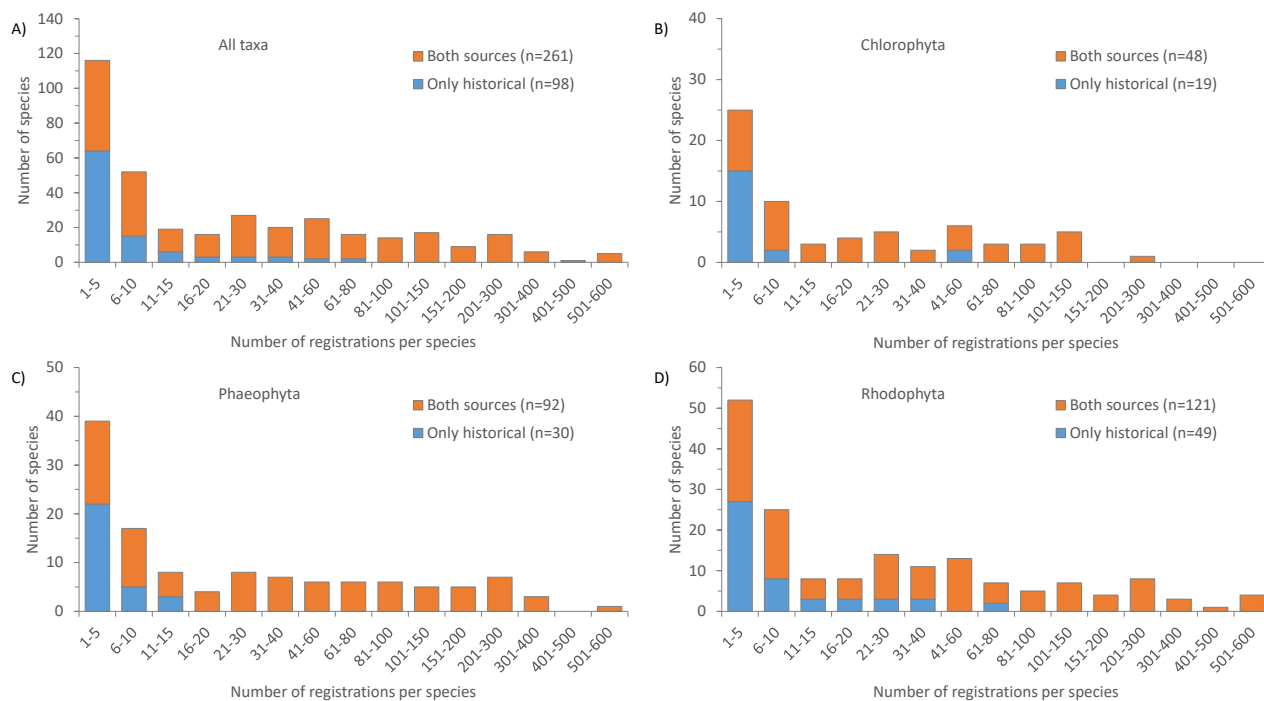


Figure 3.1. The number of species registered in both data sources and in only the historical data versus the number of times that the species were registered in the historical data for all macroalgae species combined (A) and for the three taxonomic groups (B-D).

Similarly, a number of taxonomical classifications were found in the monitoring data only (n=68), and nine of these were identifications at genus level, which may correspond to recordings in the historical data at species level. However, the majority of taxonomical classifications only observed in the monitoring data were recorded less than 10 times out of more than 600,000 recordings (Figure 3.2). An exception is the invasive species *Sargassum muticum* that was recorded 3671 times in the monitoring data. For the entire monitoring data set, 96.9% of the recordings are species also found in the historical data; only 3.1% of the recent taxonomical recordings do not appear in the historical data. The proportion of recordings within the common taxonomy was similar for all groups: 96.1% for Chlorophyta, 96.3% for Phaeophyta and 97.3% for Rhodophyta.

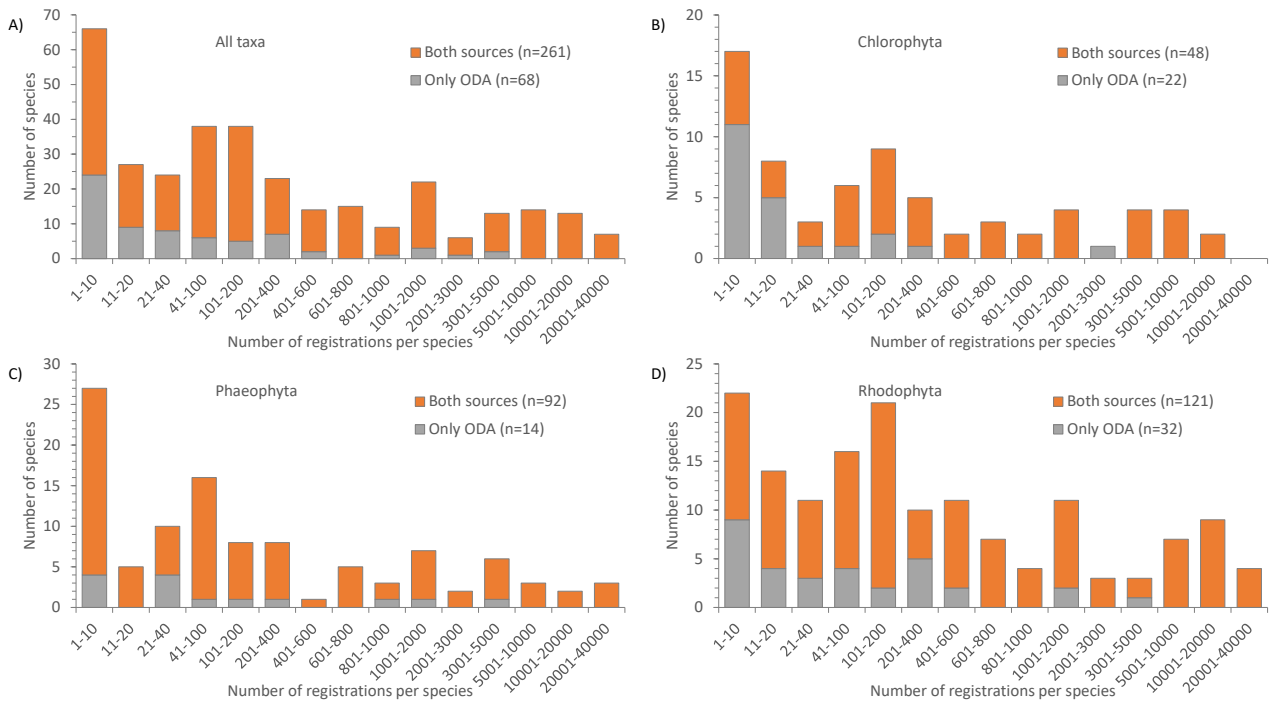


Figure 3.2. The number of species registered in both data sources and only in the monitoring data versus the number of times that the species were registered in ODA for all macroalgae species combined (A) and for the three taxonomic groups (B-D).

The number of taxonomical classifications was highest in the Kattegat, followed by the southern Belt Sea and the Skagerrak, whereas the lowest numbers were found around Bornholm and in the North Sea and the western Baltic Sea (Figure 3.3). This reflects the combination of the salinity gradient and the number of observations in the different areas. Most taxonomical classifications were made in the Kattegat, representing 68% of the historical observations and 50% of the monitoring data. The southern Belt Sea represented 15% of the historical observations and 47% of the monitoring data. For the western Baltic Sea, both data sets represented approximately 3% of the recordings. Importantly, the North Sea, the Skagerrak and Bornholm were relatively poorly represented in the monitoring data, each accounting for less than 0.5% of the monitoring records. For comparison, these three areas represented 2.2%, 8.7% and 3.2% of the historical data. Thus, data from the Kattegat and the southern Belt Sea are best suited for comparing the pool of taxonomical classifications between the two data sets.

In the Kattegat, 249 taxonomical classifications were common for both data sets, whereas 86 and 73 taxonomical classifications were found only in the historical and monitoring data, respectively (Figure 3.3). Excluding the most seldom observed taxonomical classifications (<5 recordings in the historical data and <10 recordings in the monitoring data) resulted in a common pool of 169, with 33 and 49 taxonomical classifications found only in the historical and monitoring data, respectively. There were generally more taxonomical classifications in common for Phaeophyta than for the two other groups. Phaeophyta are generally characterised by larger perennial species that are easier to identify.

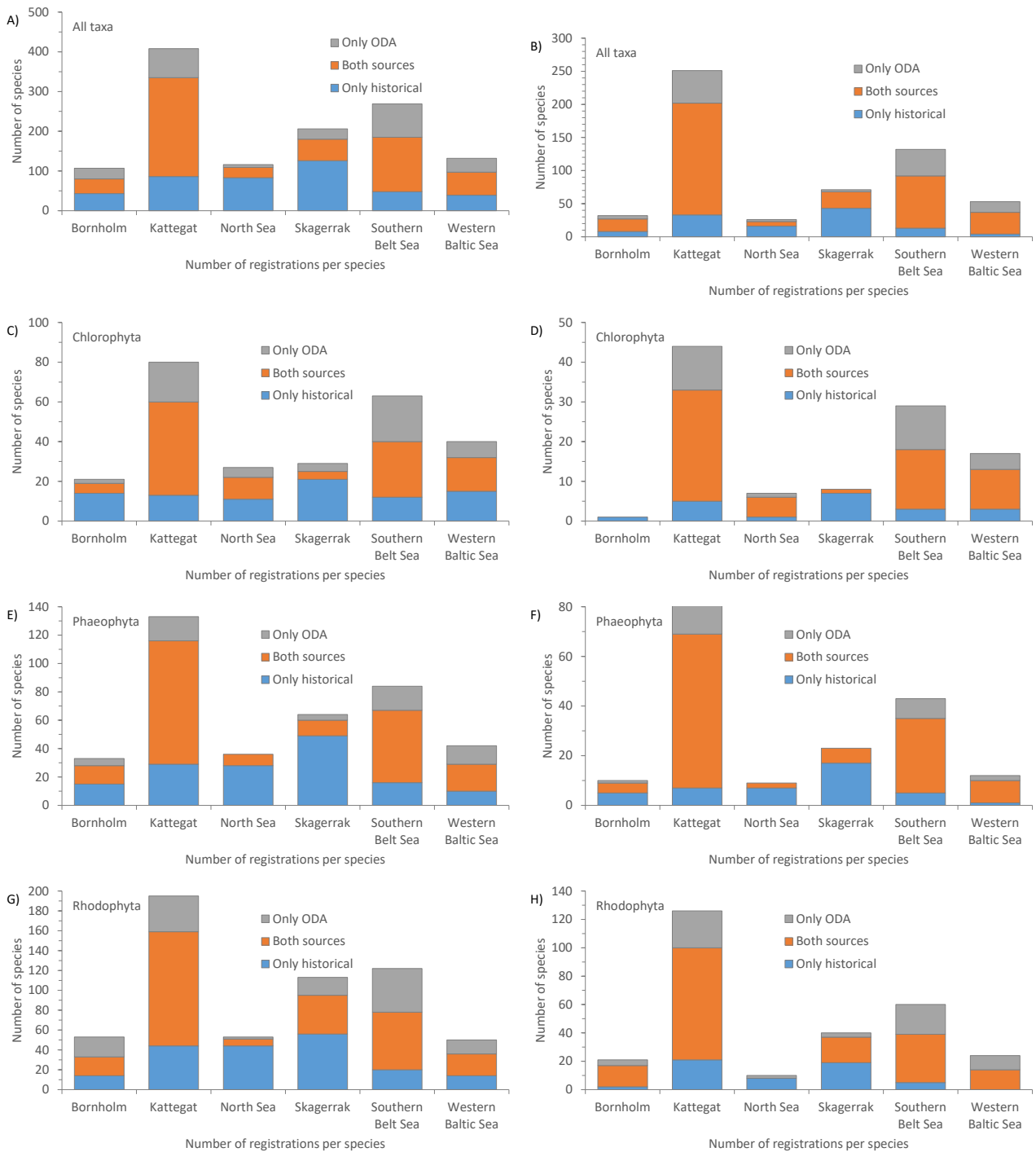


Figure 3.3. The number of taxonomical classifications registered in both the historical and the monitoring data or in one of the data sets alone for all species and divided into three taxonomic groups. Left panel: All taxonomical classifications included. Right panel: Excluding taxonomical classifications recorded less than 5 times in the historical data and less than 10 times in the monitoring data.

For the southern Belt Sea, 137 taxonomical classifications were common for both data sets, whereas 48 and 84 taxonomical classifications were found only in the historical and the monitoring data, respectively (Figure 3.2). Excluding the most seldom observed taxonomical classifications (<5 recordings in the historical data and <10 recordings in the monitoring data) resulted in a common pool of 79, with 13 and 40 taxonomical classifications found only in the historical and the monitoring data, respectively. As for the Kattegat, there were generally more taxonomical classifications in common for Phaeophyta than for the two other groups.

Q2: Is the historical macroalgae community similar to that found in the contemporary monitoring data?

Answer: It is not possible to compare the taxonomy of the two data sets directly. This is because the taxonomical resolution is different and because some species may be the same but identified with different names or to genus level only in one of the data sets. However, the majority (~95%) of all recordings in both data sets belong to the same common species pool. Furthermore, taxonomical classifications unique to one of the two data sets generally occur in low numbers, suggesting that these taxa could be sporadic occurrences or misclassifications. Thus, it appears most likely that the species pool did not change over the 20th century.

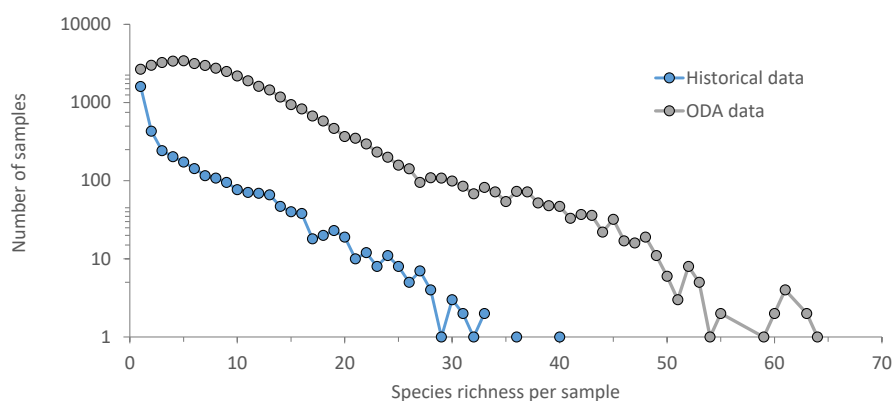
Shifts in the composition of the macroalgae community – with some species becoming more or less dominant over time – may have occurred. However, the historical data cannot support such an analysis due to their qualitative nature, i.e. species abundance is not recorded in the form of biomass or coverage.

3.3 Comparing species richness

The historical algae observations were not obtained according to a specific protocol or technical guidance, and the exact effort put into the investigation is therefore unknown. Historical investigations may have covered a small or large area and included data from longer periods (e.g. days to weeks to months). However, to obtain a relatively comparable expression of the species richness in the historical investigations, the number of distinct species within each investigation (defined as those having the same spatial location, year and depth) was calculated. Similarly, the species richness of the macroalgae community was calculated from the recent monitoring data for each sampling location, time and depth.

In the majority of the historical investigations, one species was registered, corresponding to 44% of all investigations, which is probably an indication of the fact that many historical investigations were focused on observing 'unusual' species (Figure 3.4). In relatively many historical investigations, just two species (12%) were observed, but the distribution of species richness followed a tendency similar to that in the monitoring data of a species richness above five. Therefore, it seems reasonable to assume that historical investigations including at least five species represented an evaluation of the entire macroalgae community.

Figure 3.4. The number of species registered in each investigation ("sample") in the historical and the recent monitoring data.



For comparing the species richness distribution between the historical and recent monitoring data, only investigations with at least five species were included, and the frequency of samples from the monitoring data was scaled to match the total number of historical investigations with at least five species. This normalisation balances, to some degree, the differences in sampling intensity between the two data sets. The species richness distributions for the historical and the monitoring data were relatively similar for the community as a whole (Figure 3.5A), particularly when taking into account the fact that there were only 10 historical investigations registering 30 or more species, leading to a slight deviation from the distribution curve based on the monitoring data. Half of these historical investigations were from stone reefs in the Kattegat. The majority of recent monitoring investigations with more than 30 species were also from stone reefs in the Kattegat.

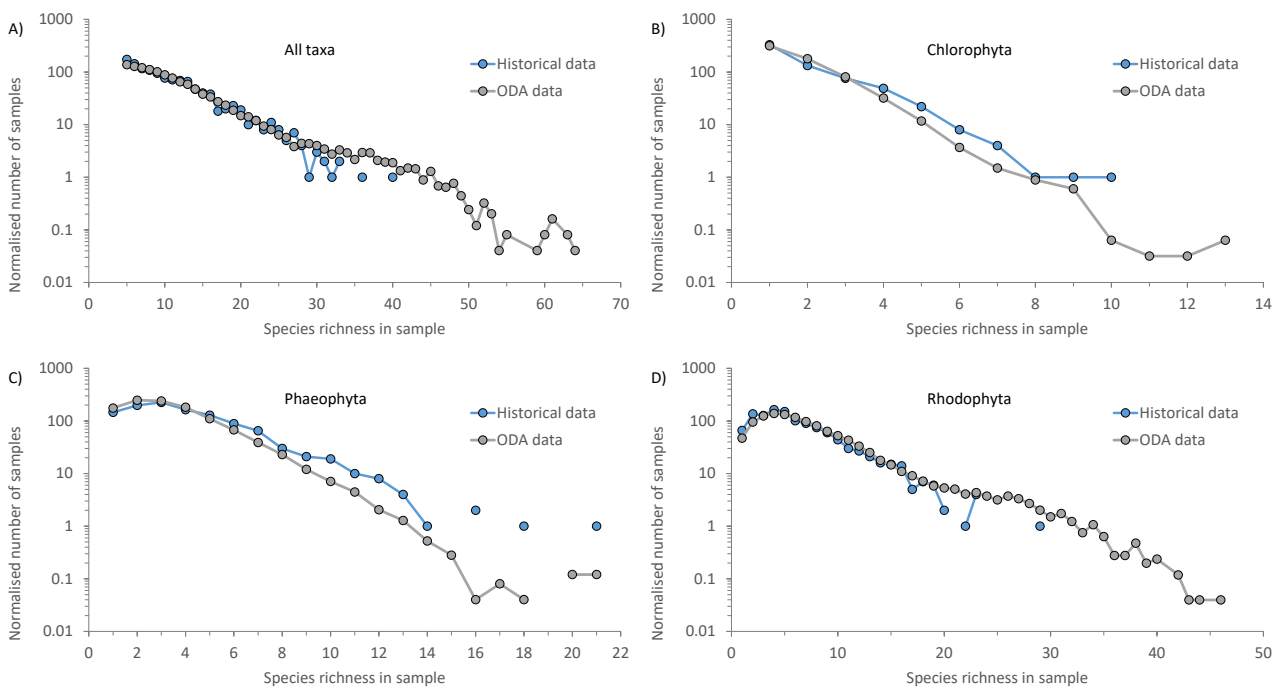


Figure 3.5. Comparison of species richness between the historical and the recent monitoring data for the entire macroalgae community (A) and the three taxonomic groups (B-D). The number of samples from the monitoring data were scaled to equal the total amount of historical investigations with at least five species. Note the logarithmic scale on the y-axis.

Although the species richness distributions in the historical and the monitoring data looked similar, there were some minor differences when considering the species richness of the different groups (Figure 3.5B-D). The historical data appeared to have a slightly higher species richness for Chlorophyta and Phaeophyta when many species were recorded. This could be due to the higher taxonomical resolution of the historical data, the monitoring data more frequently being pooled into genera. In fact, there were only 87 recordings (0.5%) at genus level in the historical data opposed to 26,602 (4.3%) at genus level in the monitoring data, i.e. a factor of 306 times more genus registrations in the monitoring data. For comparison, there were, in total, only 32 times more genus registrations in the monitoring data. Hence, the probability of identification only to genus level was almost 10 times higher for the monitoring data. In the monitoring data, Chlorophyta had the highest proportion of registrations at genus level (12.3%) followed by Phaeophyta (4.9%), Rhodophyta (2.1%) having the lowest proportion.

The distribution of species richness with depth was investigated in detail for 29 WFD water bodies, for which there were sufficient data to carry out such a descriptive analysis (Annex A). The majority of these water bodies (n=18) exhibited historical species richness distributions similar to those obtained from the monitoring data. For the other water bodies, the species richness in the historical data was apparently slightly higher (n=4) or lower (n=7) than the species richness found in the monitoring data, but for all these cases sampling conditions and taxonomical resolution explained the differences. For some water bodies there was a tendency for species richness to remain relatively higher at deeper depths in the historical data, which could be explained by better light conditions during this period. However, this effect is already incorporated in the macroalgae indicator Kbio that describes the species richness attenuation with depth.

Q3: Is the historical macroalgae species richness similar to that found in contemporary monitoring data? For the entire community and different taxonomic groups?

Answer: There is no evidence in the historical data that the species richness was higher at depths with no light limitation, but in some areas the decrease in species richness with depth is steeper in the monitoring data than in the historical data. Although the historical depth registrations should be interpreted with caution, this effect could possibly be due to changes in the light conditions over the past century, leading to stronger attenuation in species richness in the monitoring data. However, this effect was only observed in a few water bodies, including Limfjorden that has experienced large changes. Differences in species richness in the different groups over time were observed, but this is most likely due to differences in the taxonomical resolution where macroalgae are identified only to genus level more often in the monitoring data.

3.4 Comparing depth limits

Many species have been recorded in the historical and the monitoring data, but it is not possible to assess the depth limits for many of these species due to the low number of registrations. Therefore, this analysis focuses on 22 species registered at least 200 times in the historical data and at least 5000 times in the monitoring data, corresponding to approximately 1% of the registrations in the two data sets. These species include six Phaeophyta (*Ectocarpus siliculosus*, *Fucus serratus*, *Fucus vesiculosus*, *Laminaria digitata*, *Saccharina latisima*) and 16 Rhodophyta (*Ahnfeltia plicata*, *Callithamnion corymbosum*, *Carradoriella elongata*, *Ceramium virgatum*, *Chondrus crispus*, *Coccotylus truncatus*, *Cystoclonium purpureum*, *Delesseria sanguinea*, *Furcellaria lumbricalis*, *Leptosiphonia fibrillosa*, *Phycodrys rubens*, *Phyllophora pseudoceranoïdes*, *Polysiphonia stricta*, *Rhodomela confervoides*, *Spermothamnion repens*, *Vertebrata fucoïdes*), but no Chlorophyta although the criteria would apply to the genus *Cladophora* sp.

The monitoring data can be interpreted as presence/absence for the common easily recognisable species because the entire community is assessed. This is not the case for the historical data that only represent presence; thus, lack of a species cannot implicitly be interpreted as absence. Therefore, it is not possible to compare the two data sets by means of presence/absence analyses. Instead, the depths where the species are present can be compared for the two data sets, bearing in mind that the depth registrations in the historical data are not always representative. For this analysis, the deepest occurrence for each of the selected species for each investigation was determined. For the

historical data, this means that the deepest occurrence of the selected species was found for each specific location and year. Since the historical data also included investigations from shallow depths, potentially representing macroalgae washed ashore, all observations with a depth of 2 m or shallower were discarded. As for the monitoring data, this means that the deepest occurrence of the selected species was found for each transect investigation. Importantly, these observations do not represent depth limits but express a distribution of deep occurrences for the selected species.

The distribution of deep occurrences depends primarily on light conditions, implying that the depth distributions of deep occurrences in Danish waters vary broadly, particularly along the gradient from estuaries to coastal and open waters. Removal of hard substrate for building harbours and due to trawling has also impacted the depth distributions. Moreover, macroalgae species have adapted themselves to certain salinity ranges, and the salinity of an investigation area will therefore also affect the depth distribution of deep occurrences. Finally, competition for light among species will also impact the depth distribution of deep occurrences. All macroalgae registrations have been associated with a water body based on their geo-reference, and this analysis is restricted to the 29 water bodies with sufficient data (see above). In fact, none of the 22 species were found in all 29 water bodies, and the analysis of the depth distributions of deep occurrences was carried out on the subset of water bodies where the specific species were found in both the historical and the monitoring data. For the purpose of analysing the depth distribution of deep occurrences, it was assumed that there were no changes in salinity over time (Ringkøbing Fjord and Nissum Fjord are not included in the 29 water bodies) and that there were no changes in competition within the community (albeit this may not be the case for Limfjorden due to the invasive species *Sargassum muticum*; Stæhr et al. 2000), most likely leaving the species pool unaltered (see Q2). Therefore, if there are changes in the depth distributions of the deep occurrences over time, these are most likely caused by changes in the light conditions. To investigate this potential change, the following model was formulated for the depth (D) of species-specific deep occurrences:

$$D = \text{period}_i + \text{wb}_j + \text{period}_i \times \text{wb}_j + \text{LOCATION}_k(\text{WB}_j) + e_{ijkl}$$

where period_i describes the overall change from the historical data to the monitoring data (old vs. new data), wb_j describes the variability among water bodies due to general differences in light conditions and salinity, $\text{period}_i \times \text{wb}_j$ describes the changes between the historical and the monitoring data for the water bodies, $\text{LOCATION}_k(\text{WB}_j)$ describes the variability among locations (used for investigations/monitoring) within water bodies, and e_{ijkl} is the residual error.

Since all depth registrations did not represent the deep populations of the species, and since some extremely deep depth registrations in the historical data could be wrong or not associated with hard substrate, a robust regression approach was employed to identify obvious outlier observations. First, observations below the 1% prediction interval from the model above (relatively low depth observations) were removed, and the model was re-estimated. Second, observations above the 99.9% prediction interval from the model above (unusually high depth observations) were removed, and the model was re-estimated. This procedure was repeated until all depth observations were within the prediction interval of the model defined by the lower 1% and the upper 99.9% of the distributions. Hence, the depth distributions of the deep occurrences were defined by period- and water body-specific means (Annex B) and

by random variation among locations within water bodies and repeated visits to the same transect (residual variation). For 21 out of the 22 species, the random variation (given as standard deviation) among locations ranged between ± 2.2 m and ± 4.7 m, and the residual variation ranged between ± 0.2 m and ± 0.7 m. For *Fucus vesiculosus*, the variation among locations was only ± 0.2 m, whereas the residual variation was higher at ± 0.8 m. These estimates of the 'natural variability' suggest relatively wide depth distributions for the macroalgae species. The variation among water bodies was highly significant for all 22 species ($P(\text{wb}) < 0.001$).

Five out of the 22 species exhibited a significant overall difference between the two periods ($P(\text{period}) < 0.05$), with three species showing an increase in depth distribution (*Ceramium virgatum*, *Phyllophora pseudoceranoïdes* and *Spermothamnion repens*) and two species showing a decrease in depth distribution (*Fucus serratus* and *Fucus vesiculosus*) (Annex B). For the remaining 17 species, there was no significant consistent increase or decrease in depth distribution across all water bodies. Almost half of the species (10) demonstrated variable changes in depth distribution among the water bodies ($P(\text{period} \times \text{wb}) < 0.05$), but the changes often contrasted each other, with some showing significant increases and others showing significant declines. Moreover, the majority of water bodies exhibited both increasing and decreasing depth distributions over time for the 22 species. Thus, four water bodies (Det Sydfynske Øhav, Kalundborg Fjord, Limfjorden, Århus Bugt syd) had a general tendency of decreasing depth distribution, and four water bodies (Kalø Vig, Kattégat Læsø, Kattégat Nordsjælland, Nordlige Kattégat Ålbæk Bugt) had a general tendency of increasing depth distributions.

In Det Sydfynske Øhav, the majority of the species had deeper occurrences around 8-12 m in the historical data and around 5-8 m in the monitoring data. In Kalundborg Fjord, there were only two historical investigations at 7 and 11 m depth, most species being registered at 11 m. In comparison, the depth distributions for most species in the monitoring data were 7-8 m, and many transects do not go deeper due to lack of suitable substrate. For Limfjorden, the depth distributions in the monitoring data were typically 0.5-2.5 m lower than in the historical data. These differences were based on a more solid data set for both the historical and the monitoring data and were most likely caused by poorer light conditions and removal of hard substrate over the last century as a result of intense trawling activities. In Århus Bugt syd, the depth distributions in the historical data were typically 6-11 m, but these have decreased to 4-9 m, corresponding to an approximate 2 m reduction.

The increasing tendency observed for the species depth distributions in Kalø Vig was most likely due to absence of deeper observations in the historical data. The deepest distribution was recorded at 3 m, whereas the depth distributions in the monitoring data typically were around 4.5-6.5 m. For Kattégat Læsø, most species were distributed around 10-13 m in the monitoring data, but considerably lower depth distributions were found in the historical data, probably due to lack of deeper investigations. This was partially also the case for Kattégat Nordsjælland and Nordlige Kattégat Ålbæk Bugt.

Our analyses clearly demonstrate that determining the depth distribution of the deeper occurrences of specific species depends more on the depth distribution of the investigations than on the presence of macroalgae species at deeper depths. Even in the more numerous monitoring data, for many water bodies the number of deep observations are insufficient to determine whether the macroalgae community is light limited (Carstensen et al. 2020), which

highlights the difficulty of determining 'true' depth distributions or, particularly, the depth limits for macroalgae species (relative to light limitation). Finally, as mentioned above, the depth registrations in the historical data should be interpreted with caution.

Q4: Are depth limits for specific species similar between the historical and the present-day data?

Answer: The historical and monitoring data were not collected with the aim of determining macroalgae depth limits. Therefore, it is not possible to determine absolute species-specific depth limits for macroalgae species from either the historical or the monitoring data. However, it is possible to estimate the depth distribution of deeper occurrences for more common macroalgae species in water bodies with a reasonable data coverage at deeper depths. In most cases, these distributions appear to be governed rather by the depth distribution of the observations (to a large extent determined by the availability of suitable substrate) than the light-dependent depth distribution of the macroalgae species. There is no general answer to the question, but weak indications exist that macroalgae species grew at deeper depths in Sydfynske Øhav, Limfjorden and Århus Bugt syd in historical times than today.

3.5 Potential of historical data for describing reference conditions

The historical macroalgae data (before 1930) represent a period believed to have relatively low anthropogenic disturbance, at least for coastal and open waters, and data from this period have been used to infer reference conditions for other biological and environmental variables such as eelgrass depth limit, Secchi depth and oxygen conditions. For these variables, large historical data sets of quantitative nature exist that have been assembled using a relatively homogeneous sampling methodology. Macroalgae historical data, on the other hand, have been collected without any specifications regarding sampling or methodology, which means that they cannot be quantitatively compared with today's monitoring data. Furthermore, depth registrations in the historical data must be interpreted with caution.

The qualitative nature of the historical macroalgae data means that these can be used to assess the species pool and partly also species richness, but this qualitative information cannot be translated into relevant characterising metrics for the reference conditions in all Danish waterbodies.

Q5: Is it possible to describe reference conditions, both qualitatively and quantitatively, for the situation around 1900 using historical macroalgae data? If so, do these conditions correspond to values obtained through modelling of Kbio and Secchi depth?

Answer: No, the historical macroalgae data are only qualitative and cannot be converted into quantitative information of relevance for assessing ecological quality through the calculation of EQR based on reference condition values. The qualitative aspects of the historical data do not suggest any changes in the macroalgae species pool (see Q2) and species richness (see Q3).

4. Conclusion

The historical macroalgae data represent an impressive collection of species registrations in Danish waters. The historical data are georeferenced and depths are associated with the majority of the registrations, albeit there are methodological issues implying that depth recordings should be interpreted with caution. The registrations can be interpreted as presence of a given species at a given time and location, but the data do not contain any absence data. Moreover, there is no information about investigation effort, so the more exact conditions for the registrations are unknown. As a consequence of this, the historical macroalgae data have a purely qualitative nature, which is incompatible with the present monitoring data, and they can therefore only be used for assessing changes in the species pool and, partially, changes in species richness.

The majority of the historical registrations (95%) are also found in the recent monitoring data, but, in addition to these, a considerable number of species are found in the historical data only. These (mostly sporadic) registrations are dominated by tiny microscopic macroalgae or belong to a genus for which it can be difficult to identify the actual species, and these species are most likely pooled to genus level in the monitoring data. It appears most likely that the species pool has not changed over the 20th century with the exception of *Sargassum muticum*.

Species richness also appears to be similar when comparing the historical data with the monitoring data at depths where there is no light limitation. For some water bodies, there are indications that the decrease in species richness with depth is steeper in the recent monitoring data, a mechanism that is most likely driven by changes in light conditions and thus incorporated into the existing Kbio indicator.

Depth limits for macroalgae species cannot be assessed either from historical or monitoring data, but it is in principle possible to estimate the species-specific depth distributions for the most common macroalgae species in water bodies with sufficient data coverage. However, the historical data have insufficient depth coverage to allow an accurate comparison of historical and recent depth distributions, but data from a few well-investigated water bodies indicate that some macroalgae species grew deeper in the past.

Despite the apparently overwhelming amount of historical registrations (>30,000), these are spread over many different areas, depths and species, which dilutes the overall qualitative information content. It is not possible to convert this relatively sparse information into reference conditions useful for assessing ecological status *sensu* the Water Framework Directive.

5. References

Carstensen, J. 2020. Macroalgae indicators for assessing ecological status in Danish WFD water bodies. Aarhus University, DCE – Danish Centre for Environment and Energy, 74 pp. Technical Report No. 170. <https://dce2.au.dk/pub/TR170.pdf>

Høgslund, S., Carstensen, J., Krause-Jensen, D. 2018. Historiske observationer af makroalger og andre blomsterplanter end ålegræs i danske kystområder https://dce.au.dk/fileadmin/dce.au.dk/Udgivelser/Notater_2018/Historiske_observationer_af_makroalger.pdf

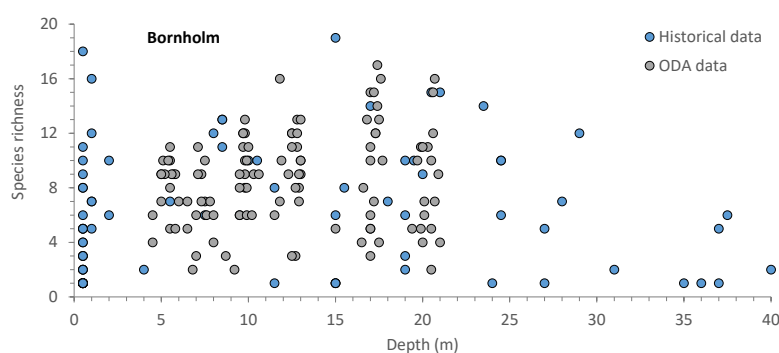
Kolderup Rosenvinge, L. 1909. The Marine Algae of Denmark. Contributions to their natural history. Part I. Det Kongelige Danske Videnskabers Selskab Skrifter. 7. Række, Naturvidenskab og Matematisk afd. VII. 1.

Nielsen, R., Dahl, K. 1992. Macroalgae at briseis Flak, Schultzs Grund and Store Middelgrund, stone reefs in southern and eastern part of Kattegat, Denmark. In Bjørnstad, E., Hagerman, L., Jensen, K. (eds): Proceedings of the 12th Baltic Marine Biologists Symposium. Olsen & Olsen, Fredensborg, Denmark.

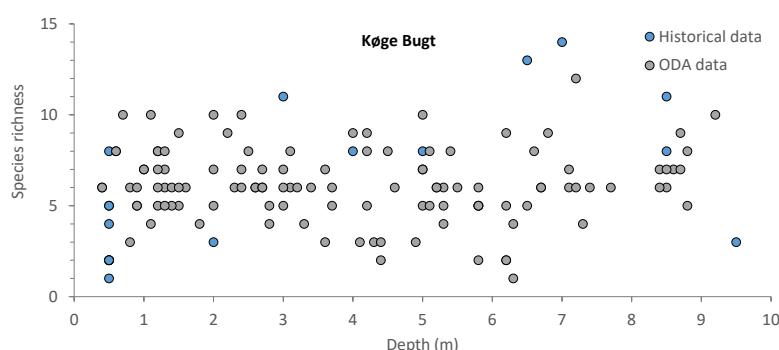
Stæhr, P., Pedersen, M.F., Thomsen, M.S., Wernberg, T., Krause-Jensen, D. 2000. Invasion of *Sargassum muticum* in Limfjorden (Denmark) and its possible impact on the indigenous macroalgal community. Marine Ecology Progress Series 207: 79-88. <https://doi.org/10.3354/meps207079>

Annex A: Species richness versus depth

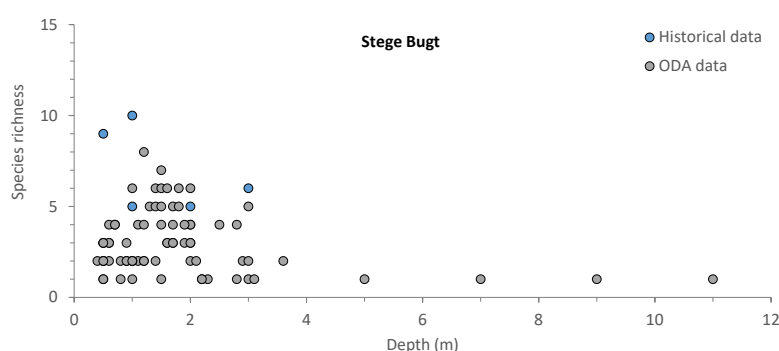
The distribution of species richness with depth was investigated for WFD water bodies with sufficient data for comparing historical data with recent monitoring data. The water bodies were pooled into a single unit around Bornholm, Kattegat (Nordsjælland), Skagerrak, Storebælt nord and Storebælt syd. Monitoring data between 1989 and 2001 were not included as the investigations in this period were carried out in depth intervals and not at specific depths. It should be stressed that the historical data registered at shallow depths (0-1 m), most likely represent macroalgae species washed ashore and thus an integration of depths rather than a single depth. Hence, the historical samples can have a higher species richness as they represent a multitude of macroalgae communities. The historical species richness observations are examined to see if they can support the hypothesis that the species richness was higher 100 years ago.



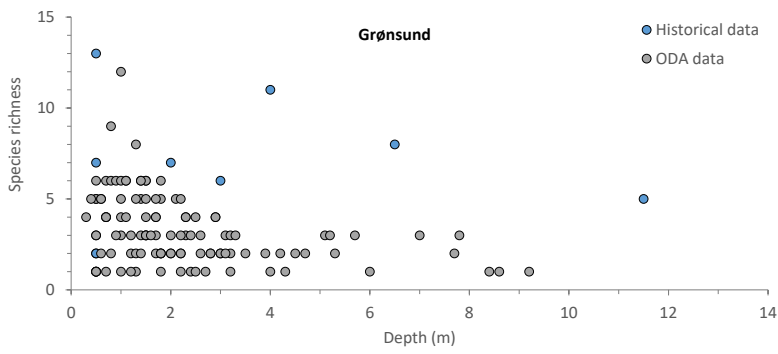
Similar species richness. Exceptions are two historical samples at shallow depths and one at 15 m from 1923. The two shallow samples probably reflect macroalgae species washed ashore from several depths, and they are therefore not comparable with monitoring data. The historical sample at 15 m with 19 species includes several *Ceramium* species, which are difficult to separate in the monitoring data. This can explain the slightly higher species richness.



Similar species richness. Two historical investigations at 6.5 and 7 m have 13 and 14 species compared to 12 as the highest number in the monitoring data. All historical species are also found in the recent monitoring data. However, these two investigations include 15 and 17 registrations where several species are recorded twice. This indicates that the investigations might have covered a broader area or period. It is noteworthy that the species richness is only slightly higher for these two investigations.

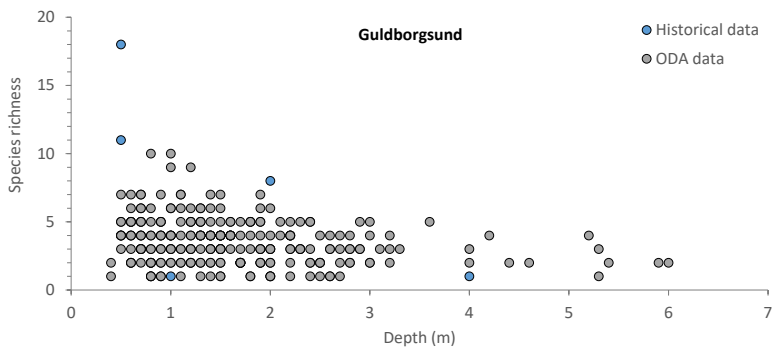


Similar species richness. There are five historical investigations of which three are consistent with the monitoring data. Two historical investigations have nine and 10 species compared to a maximum of eight in the monitoring data. The historical data include two species that are not found in the monitoring data and two *Ceramium* species that are most likely identified to genus level in the monitoring data. The higher species richness in the historical data is most likely due to their higher taxonomical resolution.

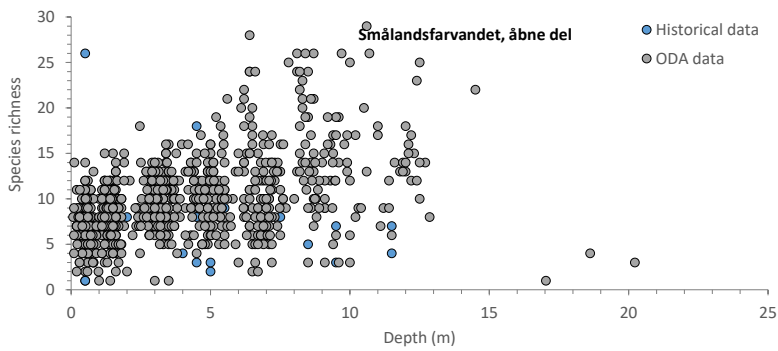


Higher species richness in historical data.

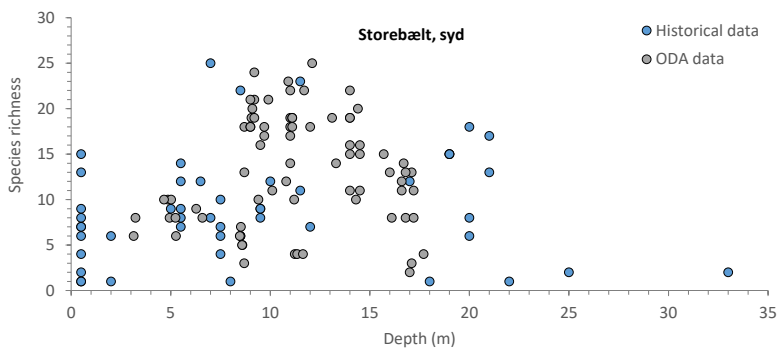
Maximum species richness is 13 for the historical data and 12 for the monitoring data. All the historical species are also found in the recent monitoring data. Species richness is generally higher at deeper depths in the historical data. It has been difficult to find suitable locations for monitoring in Grønsund, which might explain the relatively low species richness despite good light conditions (Carstensen 2020). Monitoring was discontinued in Grønsund for this reason.



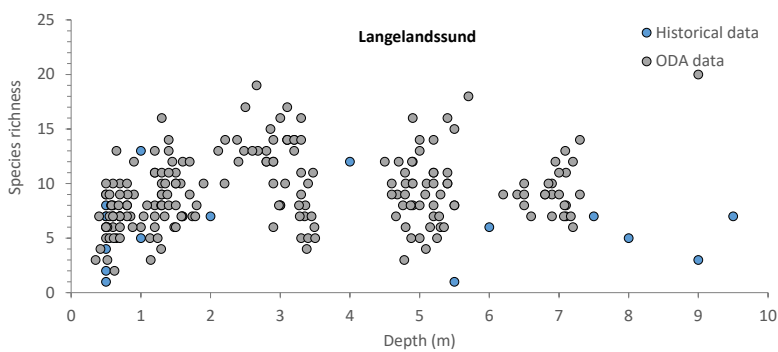
Similar species richness. There are two investigations from shallow depths, and these might represent specimens washed ashore, thus signifying a depth-integrated community. Otherwise, the historical species richness appears similar to that revealed by the monitoring data.



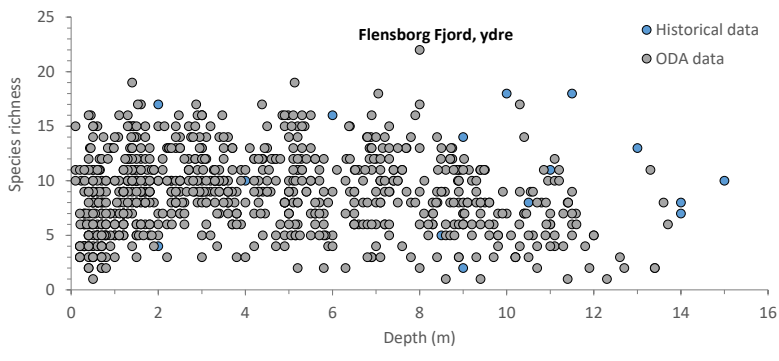
Similar species richness. There is one investigation from 0.5 m, potentially representing specimens washed ashore, thus covering a depth-integrated community. However, the historical richness observed is similar to that in the monitoring data at 10 m depth. Otherwise, the historical species richness appears to be somewhat lower than in the monitoring data, possibly due to the fact that a large proportion of the monitoring data are derived from stone reefs.



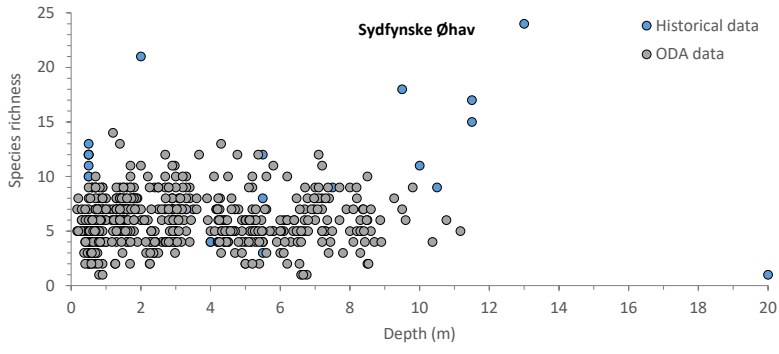
Similar species richness. A number of historical investigations were conducted at 0.5 m depth and therefore likely represent macroalgae specimens washed ashore. However, the observations are within the expected range. Otherwise, there is no major difference in species richness at the optimal depth. At deeper depths, the historical species richness is higher than the present one, which might be attributed to better light conditions then, i.e. a mechanism covered by the proposed indicator.



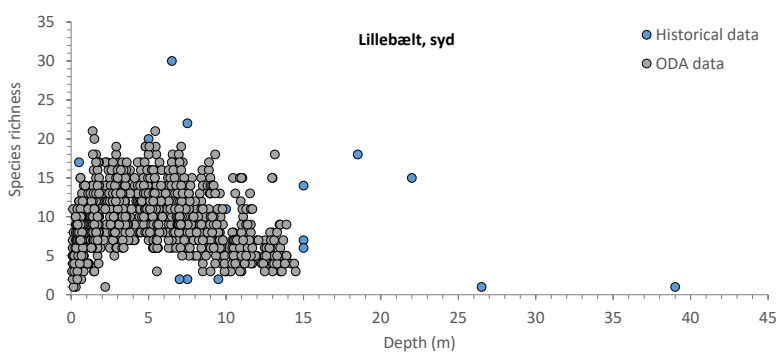
Similar species richness. The historical species richness is consistent with recent observations.



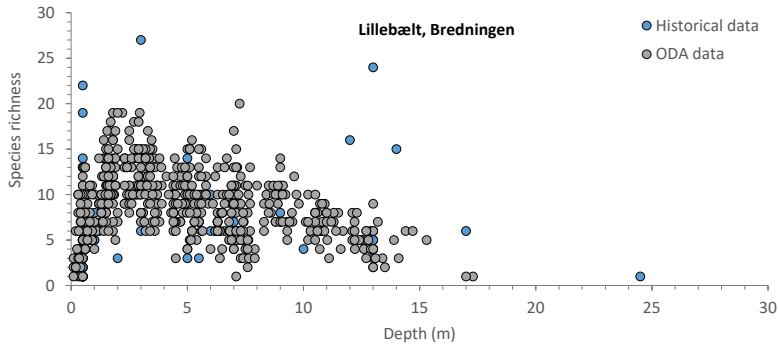
Similar species richness. The historical species richness is consistent with recent observations at intermediate depths. At deeper depths, the historical species richness is higher, which may be attributed to better light conditions, i.e. a mechanism covered by the proposed indicator.



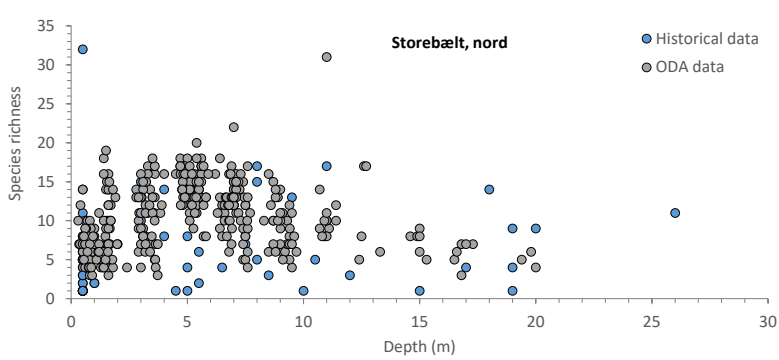
Few historical investigations with higher species richness. The historical species richness (32 out of 37) is overall consistent with the present-day monitoring data, but five investigations show a considerably higher historical species richness. In these five investigations, three rare species are only present in the historical data, and there are several registrations of species that would typically be registered at genus level in the monitoring data (e.g. *Ceramium*, *Hildenbrandia*). Finally, multiple registrations of the same species indicate a broader spatial survey.



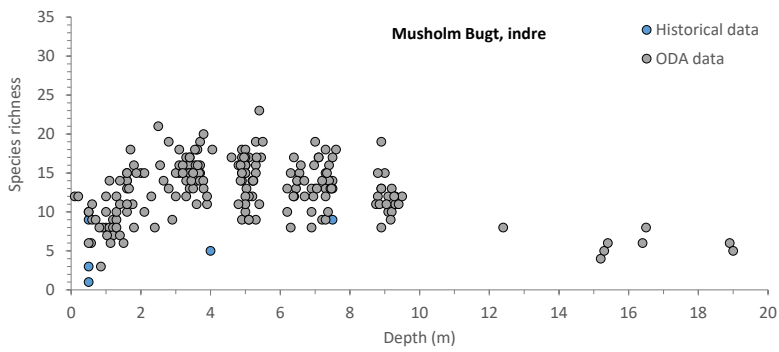
Few historical investigations with higher species richness. The historical species richness (32 out of 34) is generally consistent with the monitoring data, and two investigations (~7 m) have a higher species richness.



Few historical investigations with higher species richness. The historical species richness (67 out of 70) is overall consistent with the monitoring data, and three investigations (1, 3, 13 m) have a higher species richness. In these three investigations, there are three rare species occurring only in the historical data, and there are some registrations of species that would typically be registered at genus level in the monitoring data (e.g. *Ceramium*, *Hildenbrandia*). Finally, multiple registrations of the same species indicate a broader spatial survey.

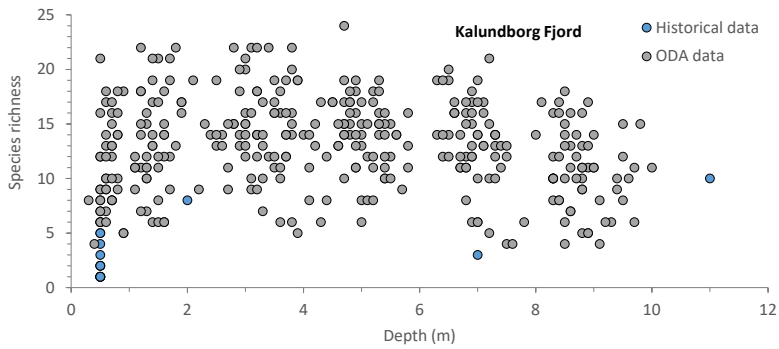


Similar species richness. Species richness appears similar in the historical and the monitoring data, except for two unusually high species richness observations – one in the historical data and one in the monitoring data. The historical observation is made at shallow depth, indicating that it consists of macroalgae specimens washed ashore. The monitoring registration includes three high species richness observations (three depths), of which one is excessively high. This sample appears suspect.



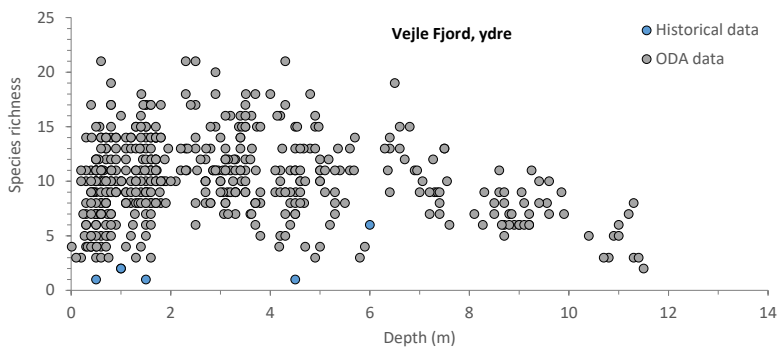
Historical data have low species richness.

The species richness in the historical investigations comes from the lower part of the depth distribution, indicating that some investigations did not examine the entire community.



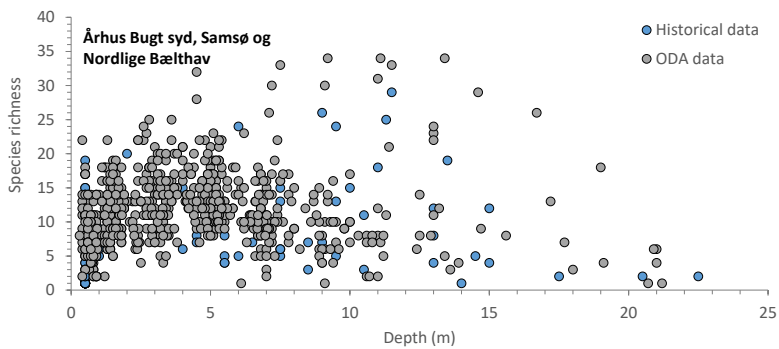
Historical data have low species richness.

The species richness in the historical investigations comes from the lower part of the depth distribution, indicating that some investigations did not examine the entire community.



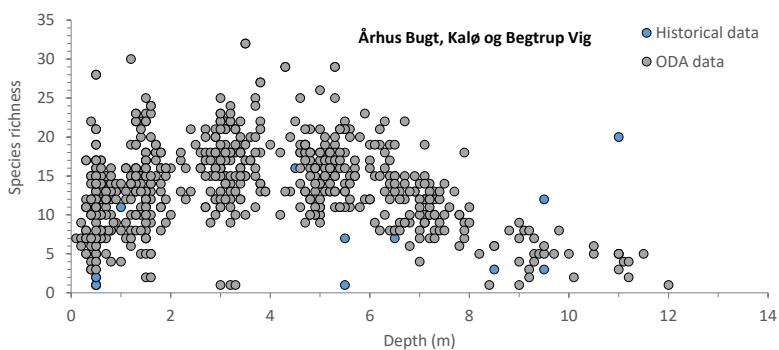
Historical data have low species richness.

The species richness in the historical investigations comes from the lower part of the depth distribution, indicating that some investigations did not examine the entire community.



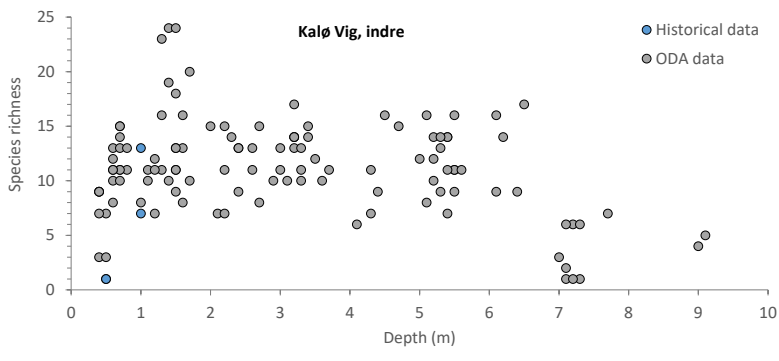
Similar species richness.

The species richness in the historical data and the monitoring data generally shows the same distribution with depth.

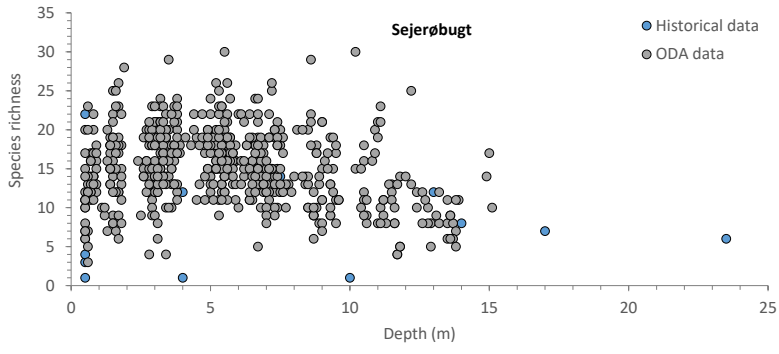


Similar species richness.

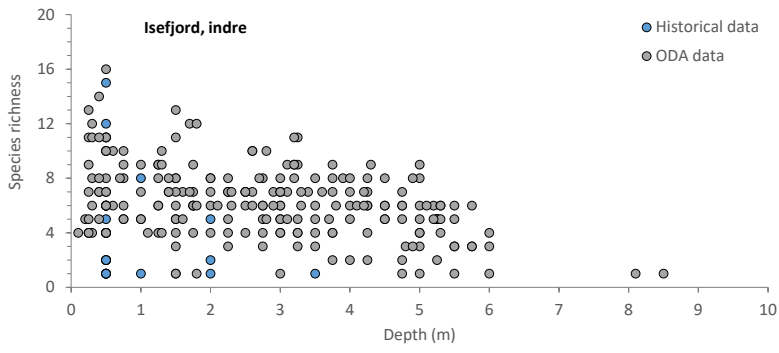
Species richness in the historical and the monitoring data generally shows the same distribution with depth, although two historical investigations – one with a relatively low and one with a relatively high species richness – deviate from the overall pattern. The historical observation with relatively high species richness has multiple registrations of the same species, indicating that a broader spatial survey was performed.



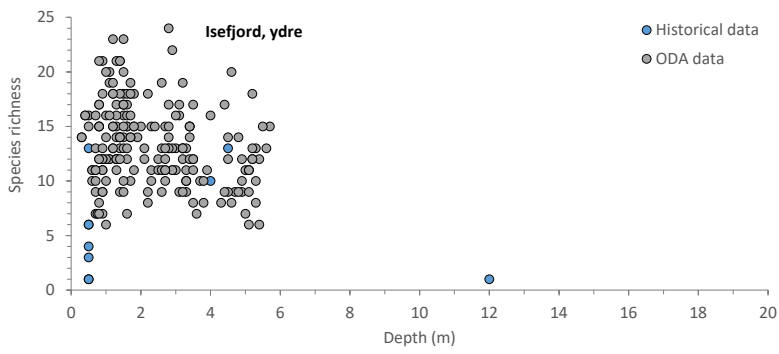
Similar species richness. The species richness in the historical and monitoring data generally shows the same distribution with depth.



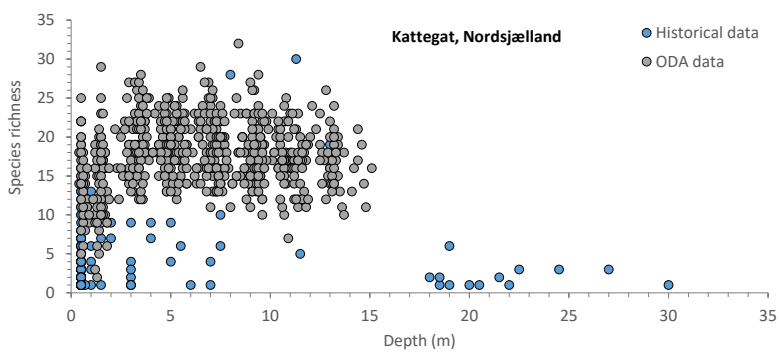
Similar species richness. The species richness in the historical and monitoring data generally shows the same distribution with depth.



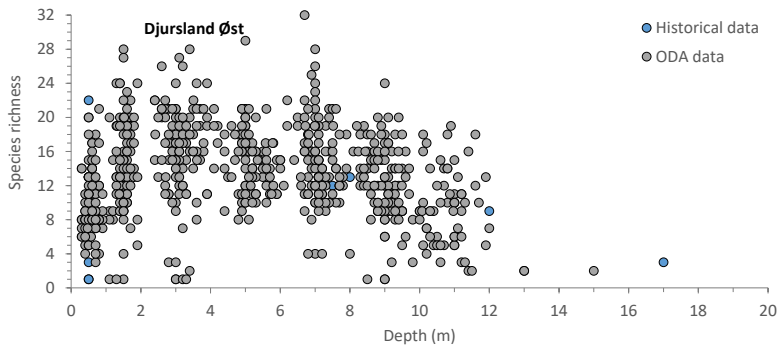
Similar species richness. The species richness in the historical and monitoring data generally shows the same distribution with depth.



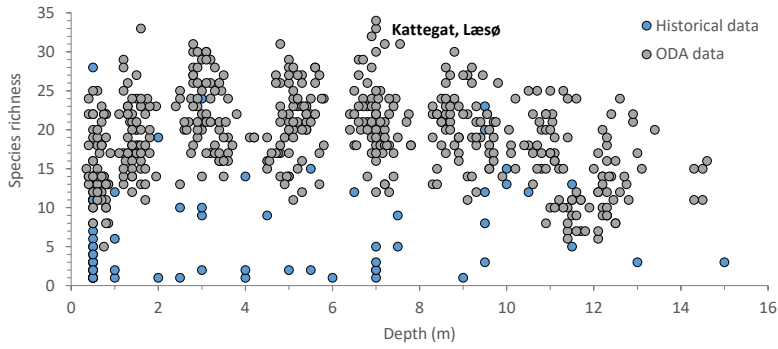
Similar species richness. The species richness in the historical and monitoring data generally shows the same distribution with depth.



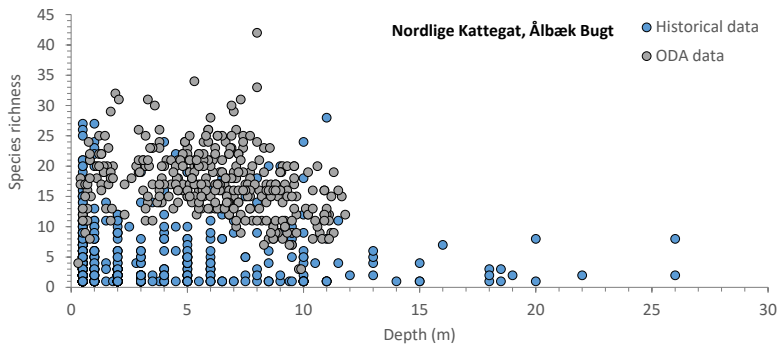
Historical data have low species richness. The species richness in the historical investigations comes from the lower part of the depth distribution, except for two investigations with a relatively high species richness. This suggests that most investigations did not examine the entire community.



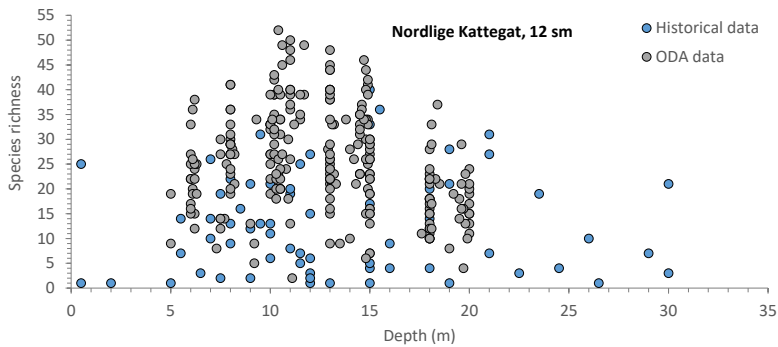
Similar species richness. The species richness in the historical and monitoring data generally shows the same distribution with depth.



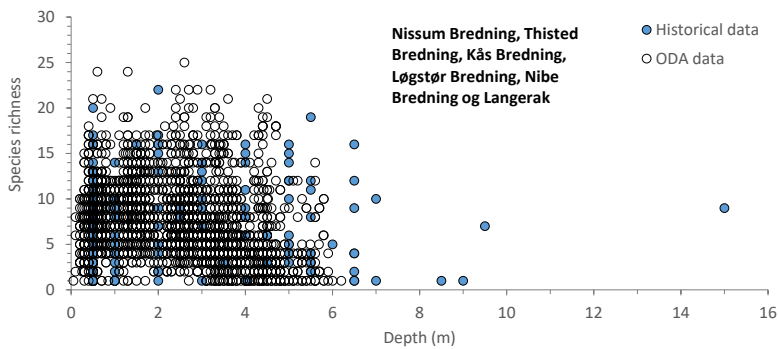
Historical data have low species richness. The species richness in the historical investigations represents the lower part of the depth distribution, except for seven investigations where the species richness is comparable to that observed in the monitoring data. This indicates that most investigations did not examine the entire community.



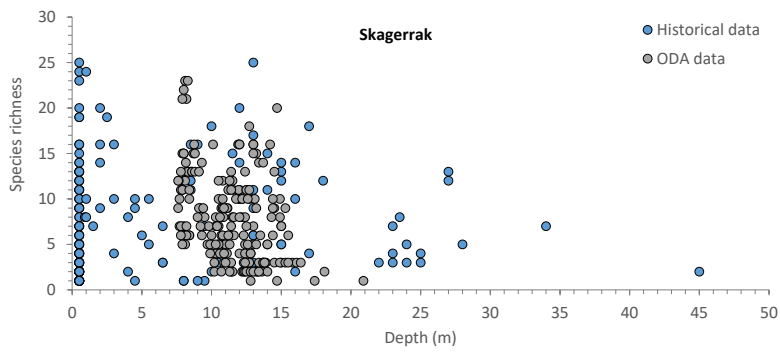
Historical data have low species richness. The species richness in the historical investigations comes from the lower part of the depth distribution, except for approximately 30 investigations where the species richness is comparable to that of the monitoring data. This indicates that most investigations did not examine the entire community.



Historical data have low species richness. The species richness in the historical investigations comes from the lower part of the depth distribution, except for approximately 20 investigations where the species richness is comparable to that of the monitoring data. This indicates that most investigations did not examine the entire community.



Similar species richness. The species richness distributions are similar (note that due to the many observations, many historical observations are hidden), albeit with a tendency to a faster decline with increasing depth in the monitoring data. This is most likely due to changes in light conditions and is thus incorporated into the indicator.



Similar species richness. Within the relatively narrow depth range of the monitoring data, the species richness distributions of the historical and monitoring data are similar. One species richness in the historical data is high (n=25 at 13 m depth), although not unusually considering the fact that the historical data are more numerous than the monitoring data.

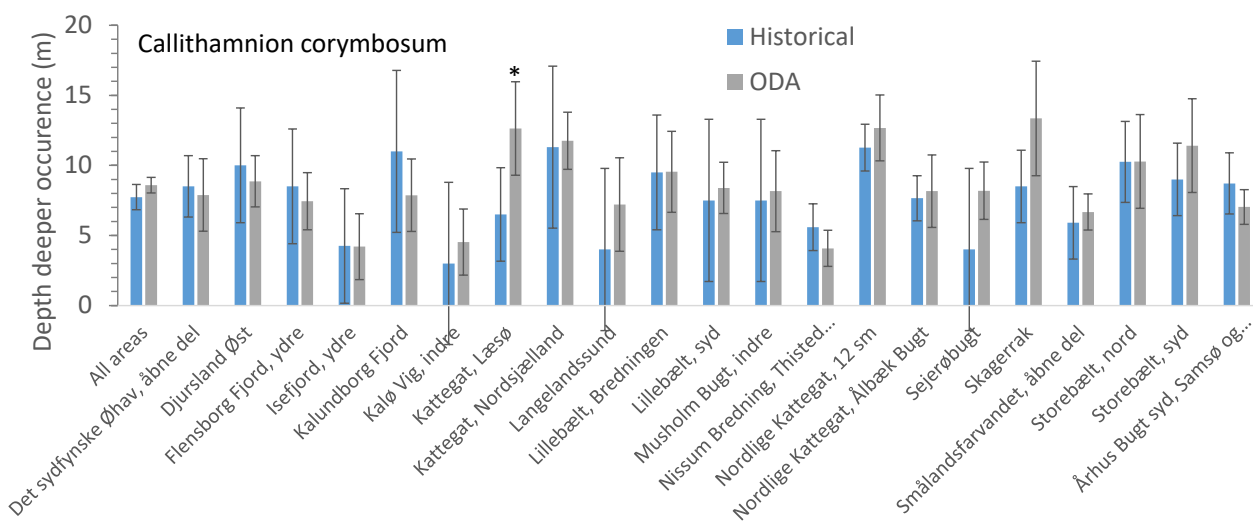
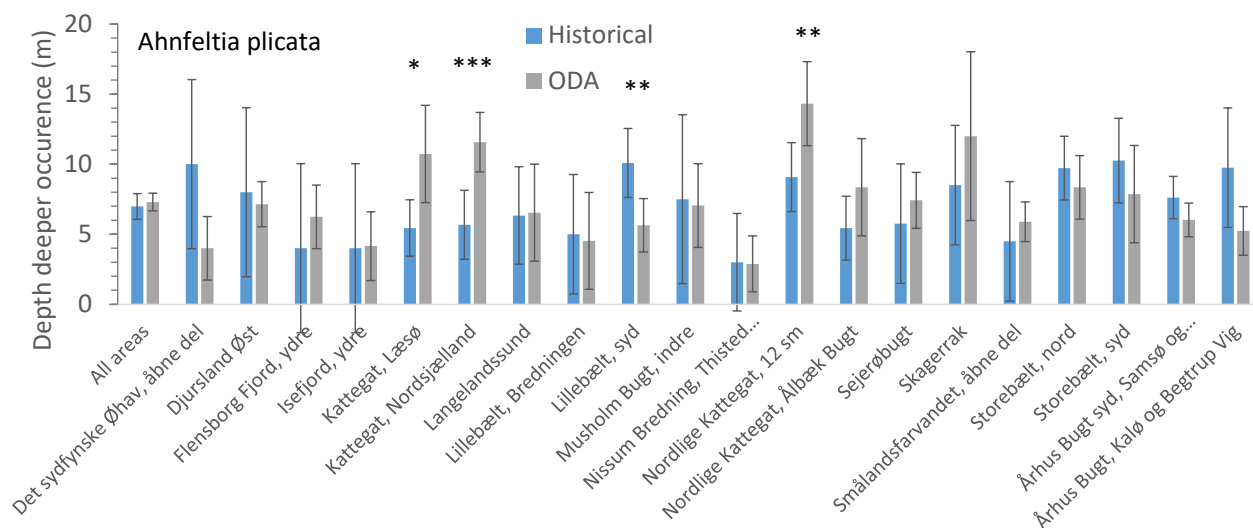
Figure A.1. Species richness versus depth for 29 selected water bodies with sufficient data for comparing historical data with recent monitoring data (2002-2020). The water bodies are organised along an approximate gradient from the Baltic Sea towards the Skagerrak.

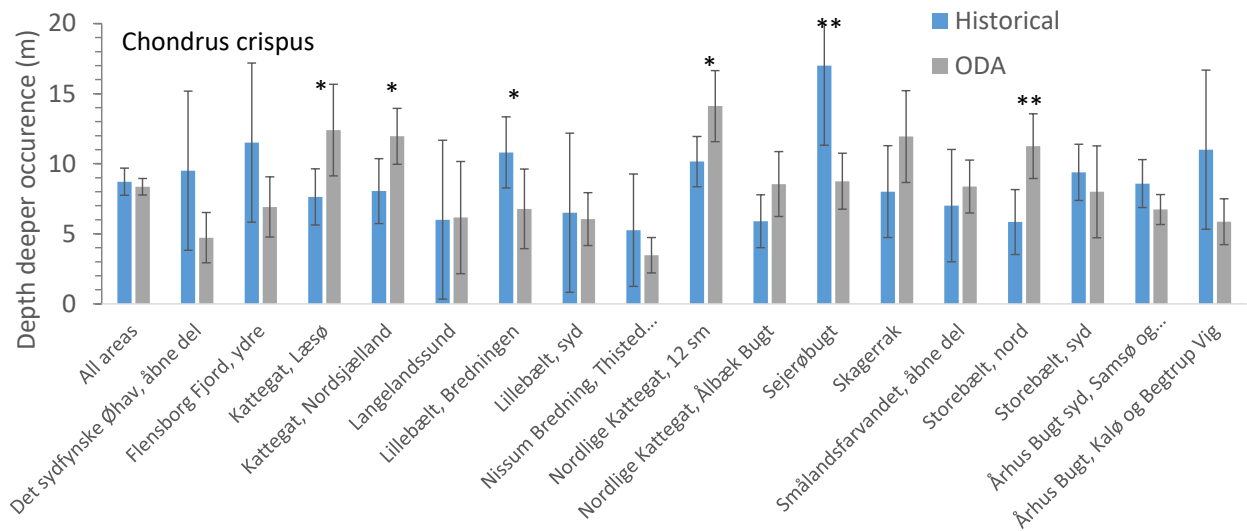
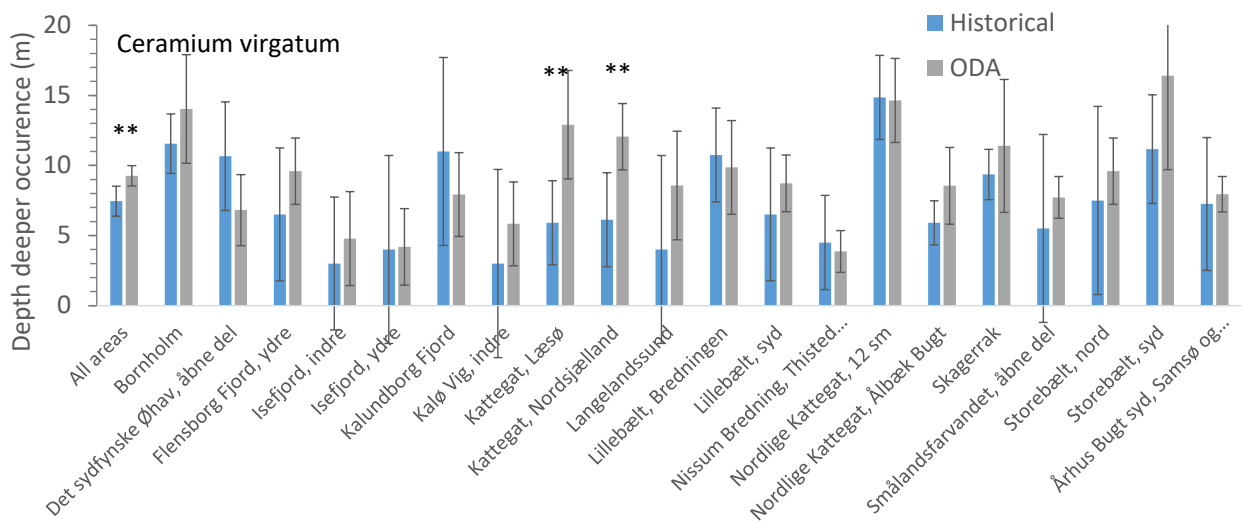
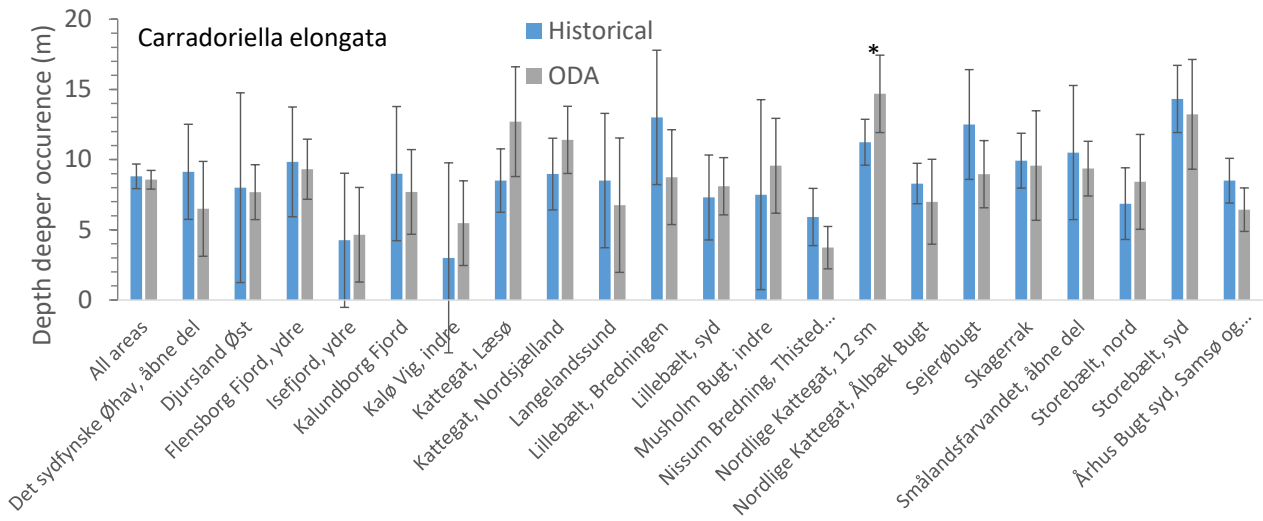
References

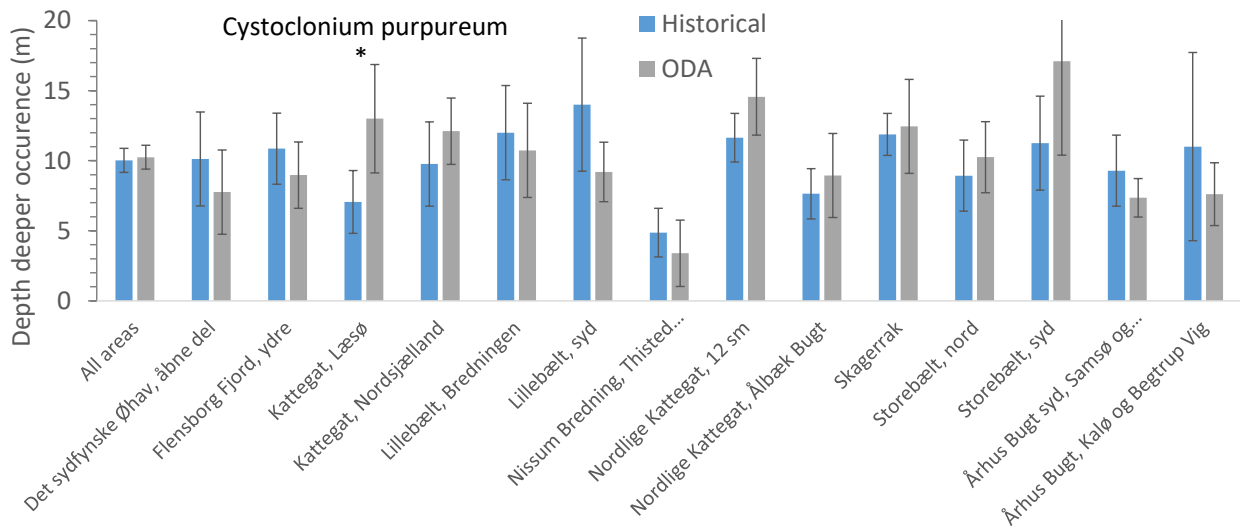
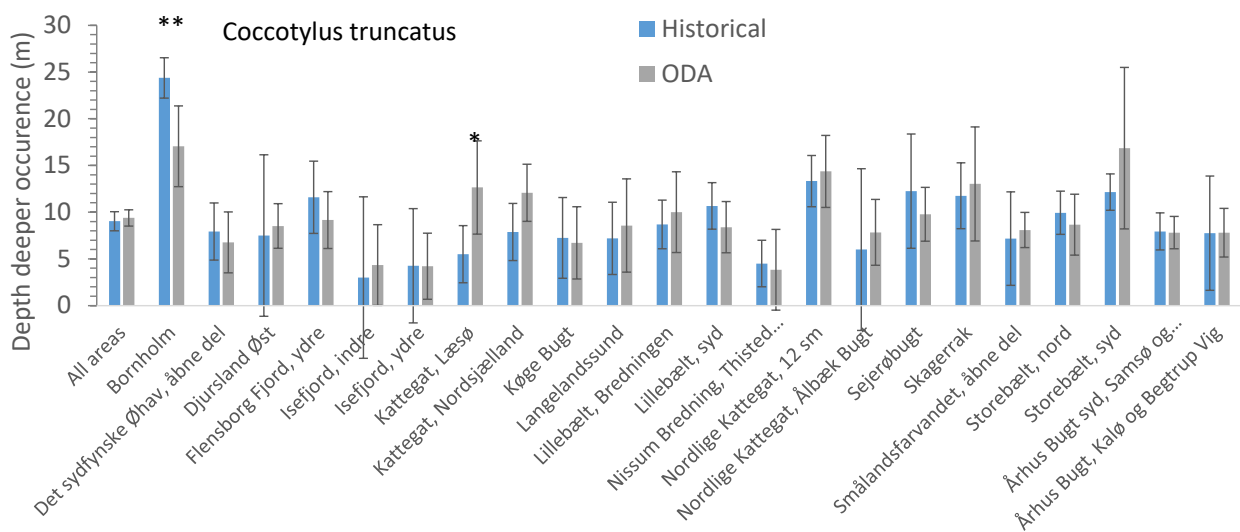
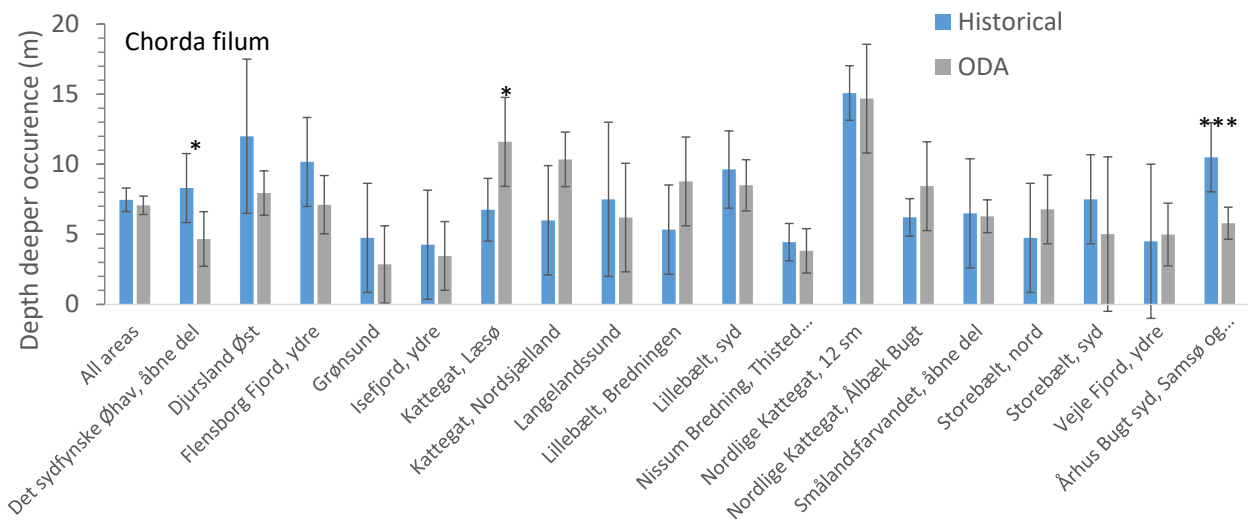
Carstensen J. (2020) Makroalgeindikatorer og deres anvendelse til VRD tilstandsvurdering. Notat fra DCE - Nationalt Center for Miljø og Energi, 28. september 2020, 25 pp.

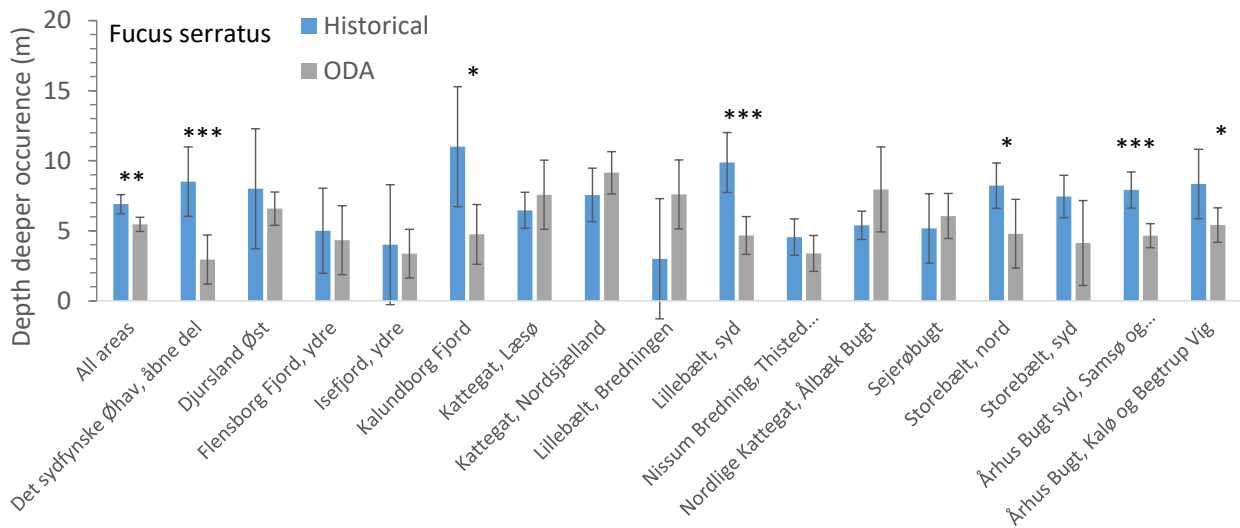
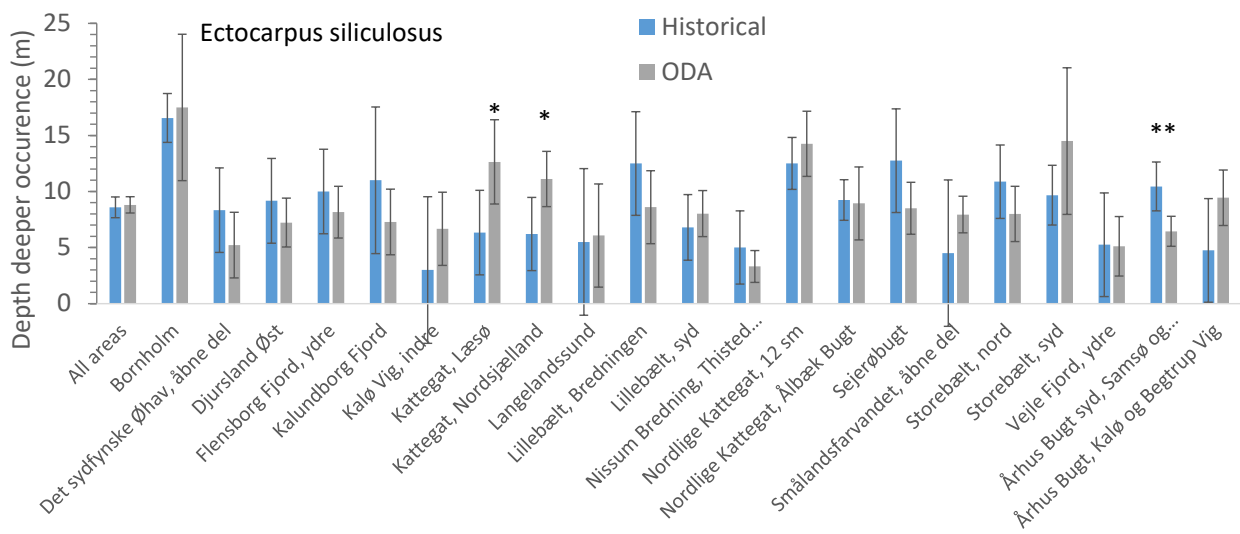
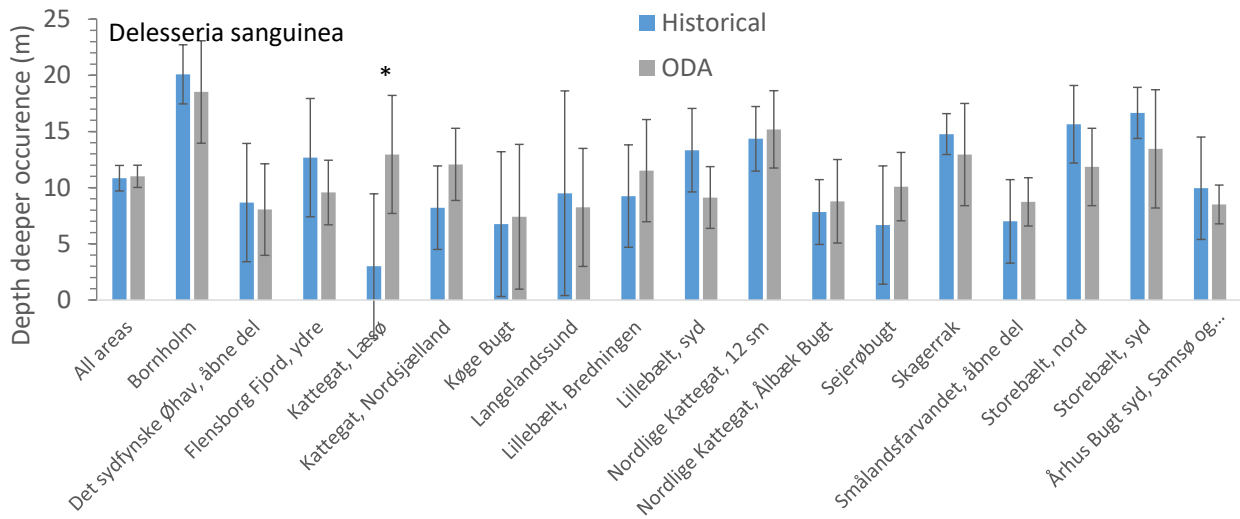
Annex B: Species-specific depth distributions

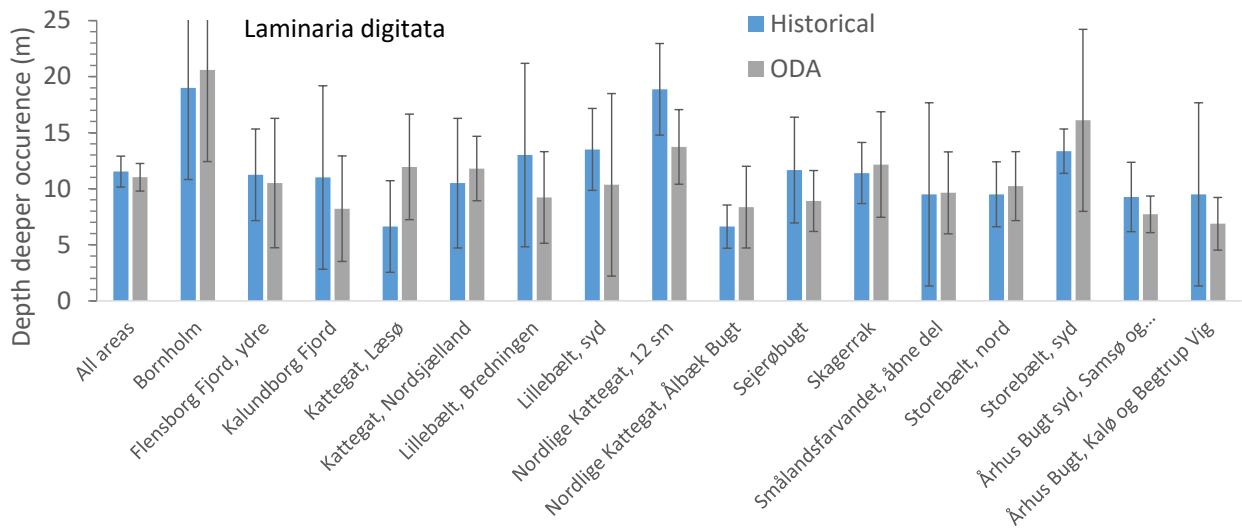
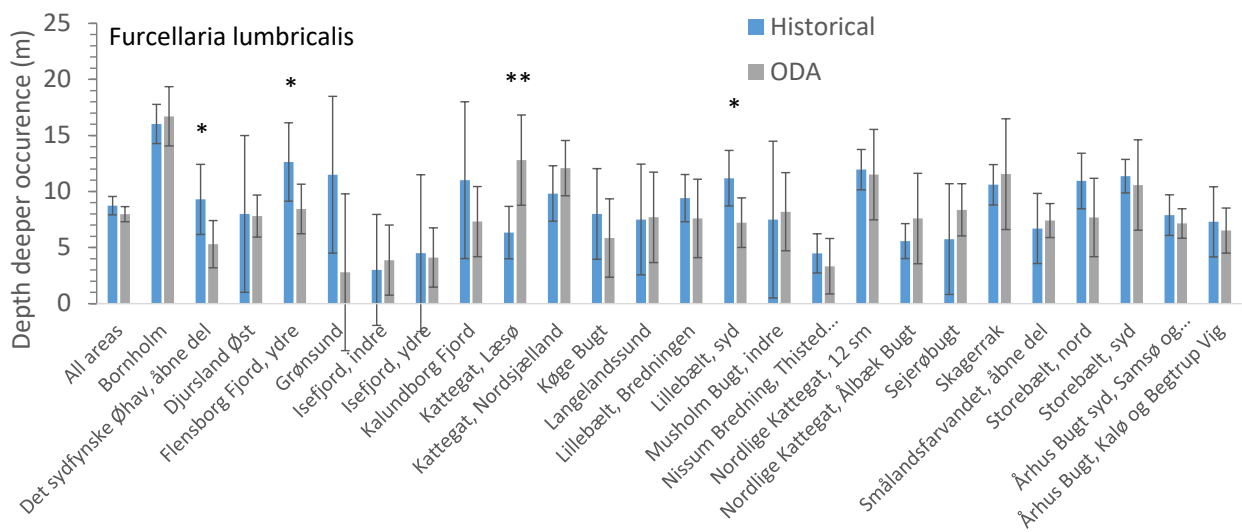
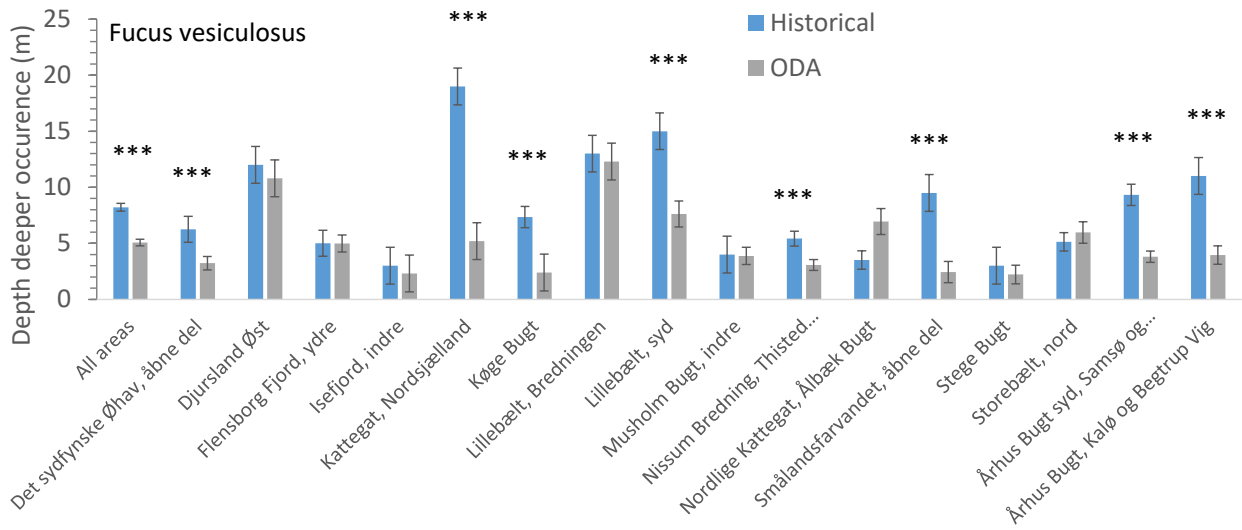
The depth distributions of the deeper occurrences of the 22 most common species were analysed by a statistical model, and the mean depths of deep occurrences were estimated for the two periods (before 1930 and after 2001) for all water bodies combined and for each water body separately. See details above.

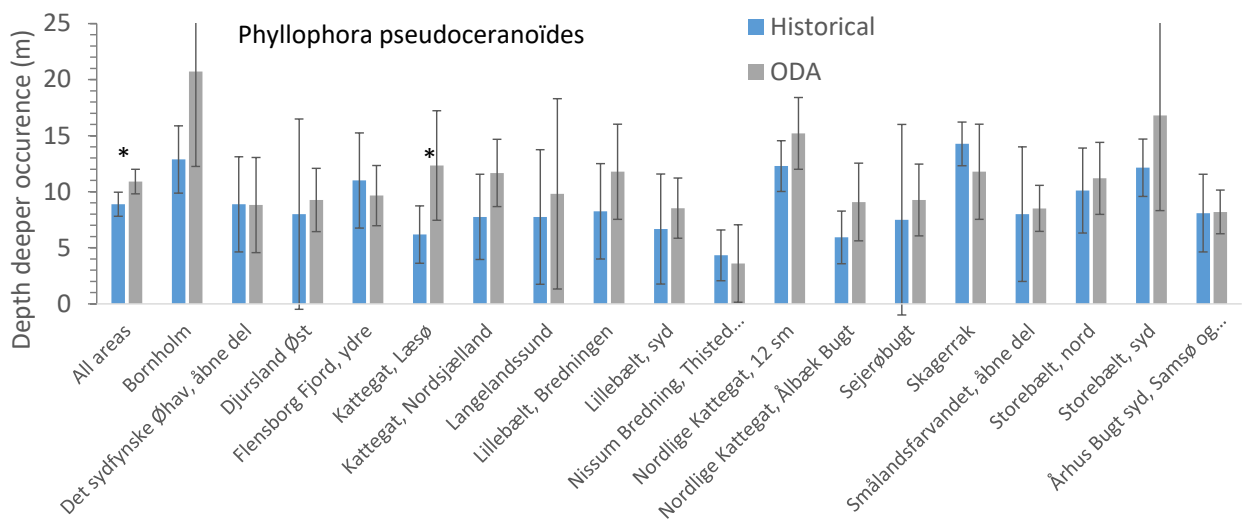
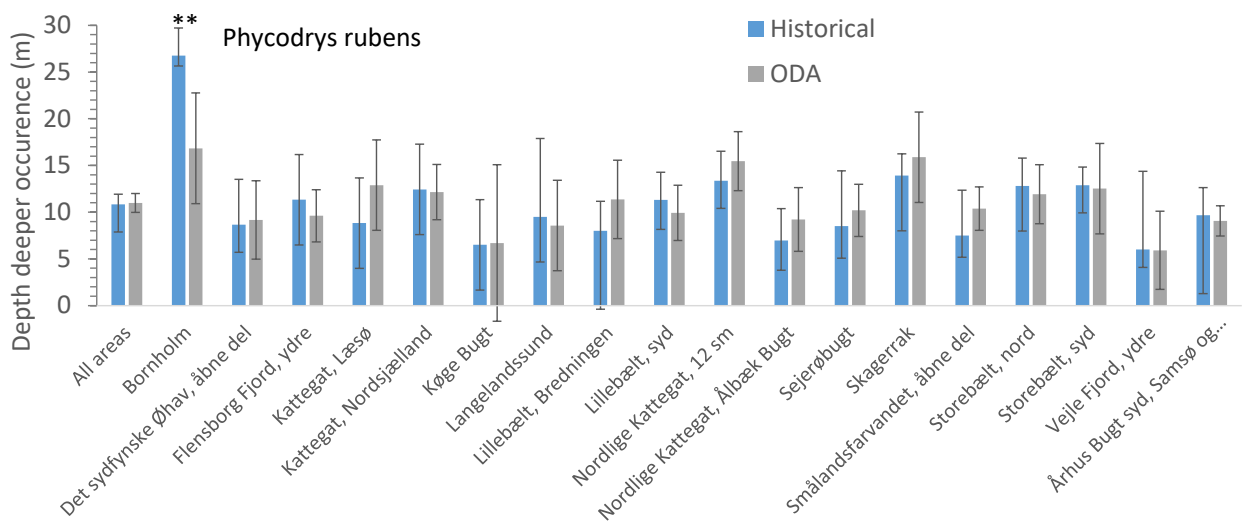
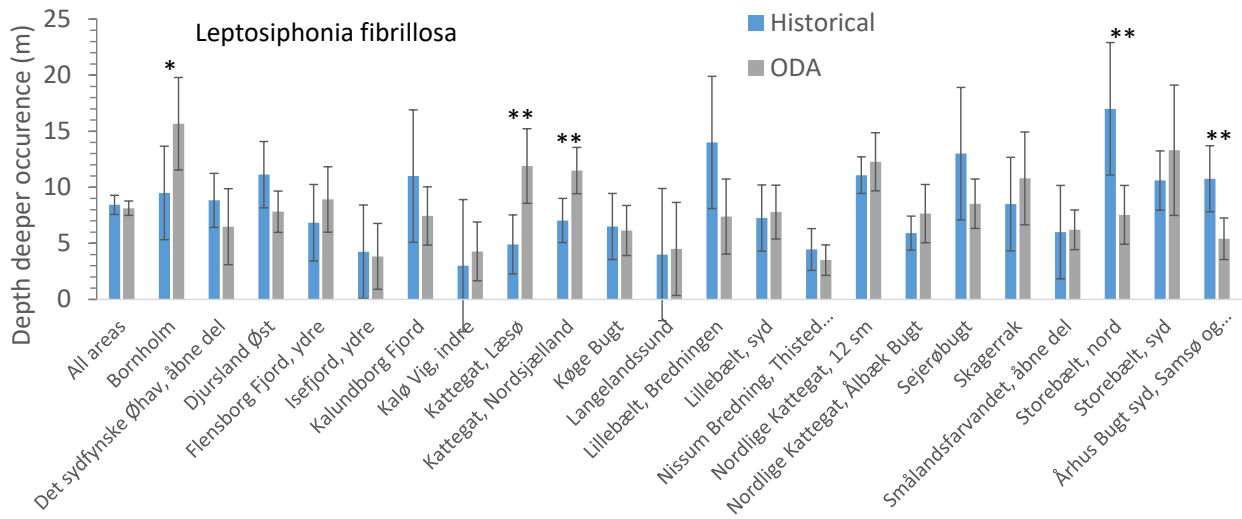


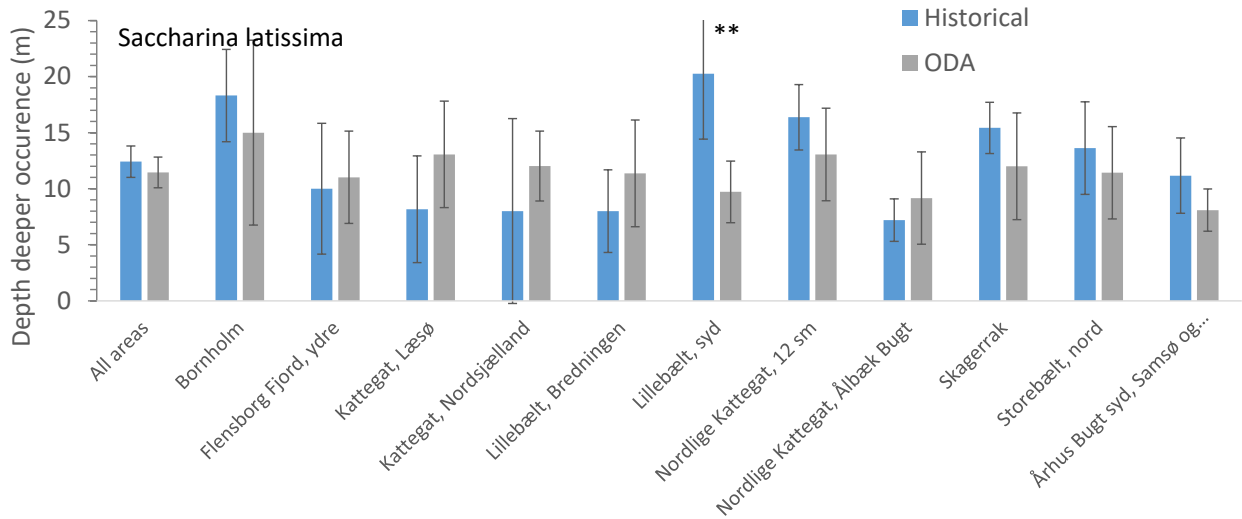
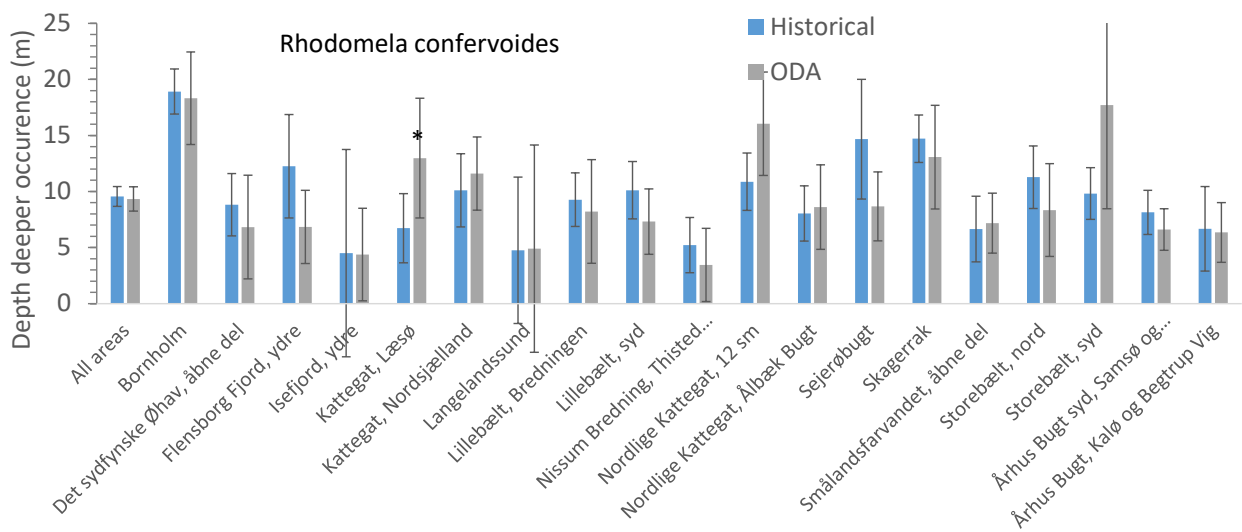
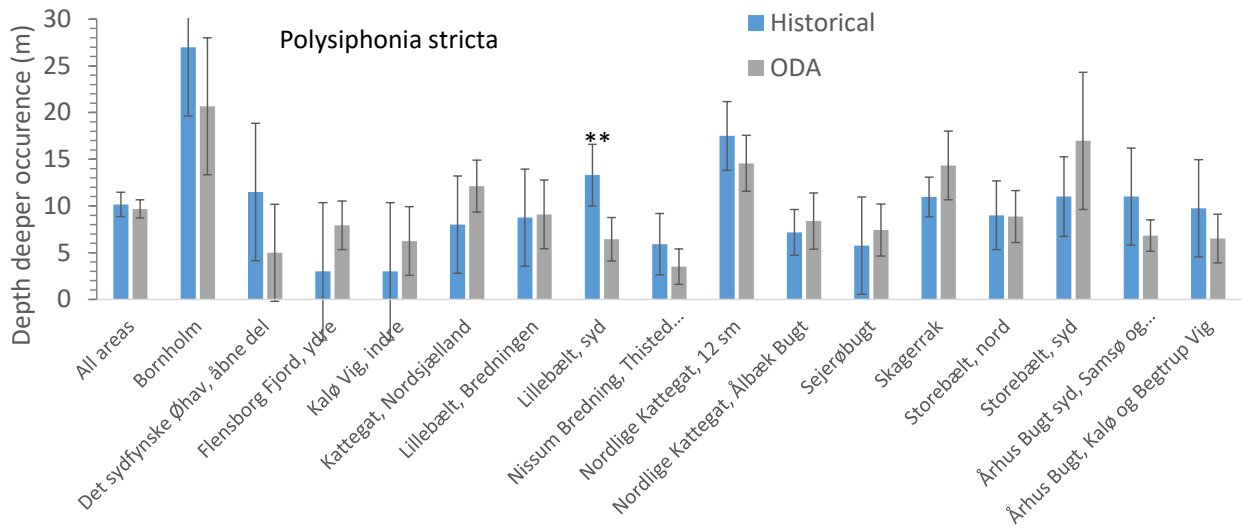












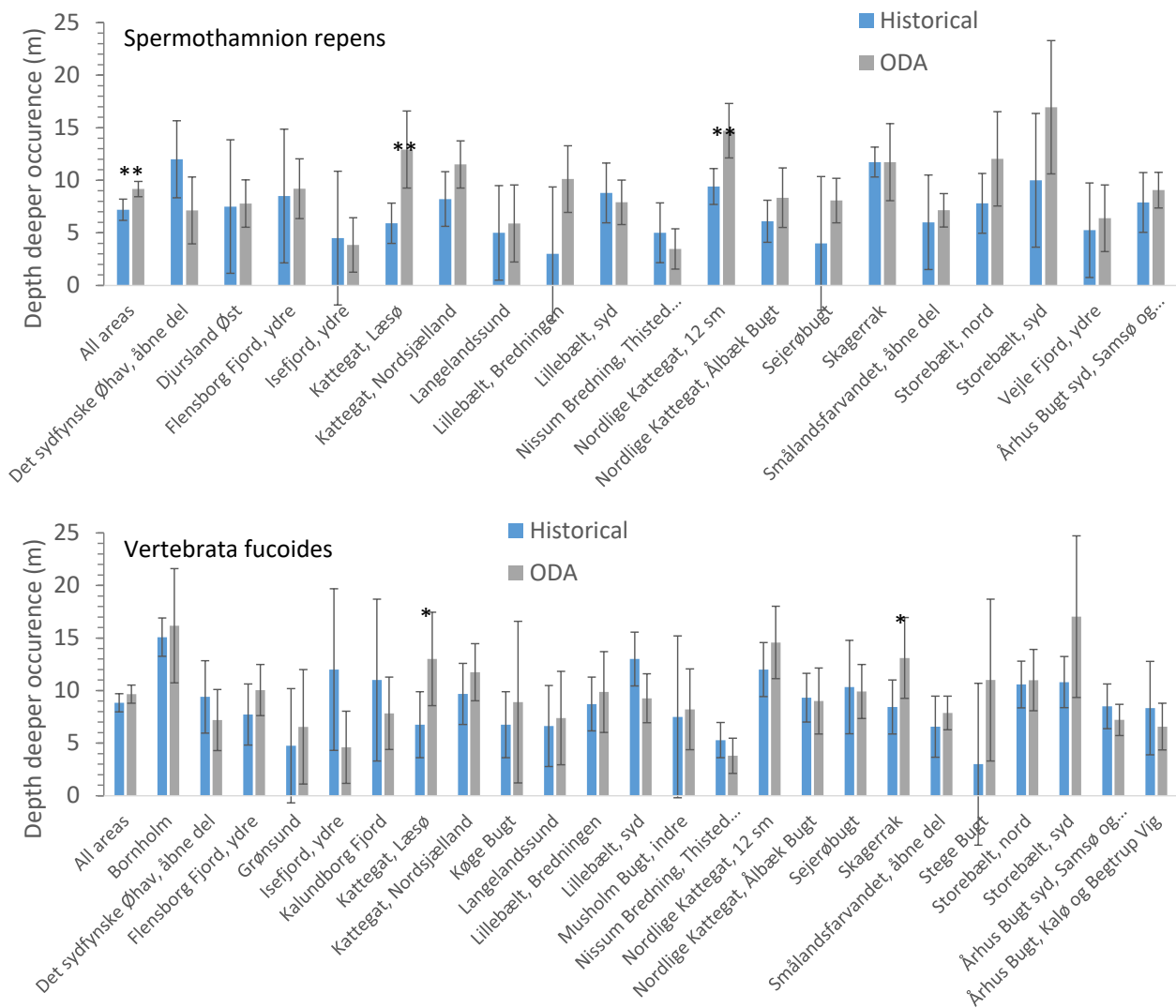


Figure B.1. Estimated mean depths for the deeper occurrence of specific macroalgae species for all areas combined and for each specific water body where both the historical and the monitoring data contained species-specific depth registrations. Error bars show 95% confidence intervals for the mean depths. Significance of testing differences between periods is indicated with * ($P < 0.05$), ** ($P < 0.01$) and *** ($P < 0.001$). Note that the number of water bodies vary between species.