

Multi-dimensional characterization of pelagic habitats and the potential to detect ecological niches in a highly dynamic ecosystem

Dissertation

With the aim to achieving a doctoral degree at:

Faculty of Mathematics, Informatics and Natural Science

Department of Biology, University of Hamburg

by

Rene-Marcel Plonus

Hamburg, Mai 2023

Reviewer:

Prof. Dr. Christian Möllmann

Prof. Dr. Stefanie Vogl

Disputation: 10.07.2023

Table of Contents

CONTENTS	2
ABSTRACT	6
ZUSAMMENFASSUNG	8
GENERAL INTRODUCTION	10
REFERENCES	17
CHAPTER I	24
ABSTRACT	25
INTRODUCTION	25
MATERIAL AND PROCEDURES	29
DESCRIPTION OF INSTRUMENT.....	29
DESCRIPTION OF HARD- AND SOFTWARE.....	29
FIELD SAMPLING.....	30
IMAGE PRE-PROCESSING	34
AUTOMATED IMAGE CLASSIFICATION	35
RESULTS	43
MODEL TRAINING.....	43
TEST SET PREDICTIONS.....	44
IMAGE FILTERING.....	46
FIELD SET PREDICTIONS	48
TOP-K PREDICTIONS	51
SPATIAL DISTRIBUTIONS.....	54
DISCUSSION	55
FILTERING	57
SPATIAL DISTRIBUTIONS.....	58
TOP-K PREDICTIONS	58
COMMENTS AND RECOMMENDATIONS	59
DATA AVAILABILITY STATEMENT	60
ACKNOWLEDGMENTS	61
REFERENCES	62
SUPPLEMENTARY MATERIAL	67
CHAPTER II	74

ABSTRACT	75
INTRODUCTION	75
MATERIAL AND PROCEDURES	77
DATA ACQUISITION AND PREPARATION.....	77
MODEL DESCRIPTION	80
HABITAT SEGREGATION	81
ANALYSES	81
RESULTS.....	82
MODEL TRAINING.....	82
CLUSTERING.....	83
PROJECTIONS	84
SILHOUETTE METHOD	85
HABITAT MAPS	86
SPECIES ABUNDANCE PLOTS	87
LLOYD.....	88
TEST DATASET HE534	89
DISCUSSION	90
SELECTION OF PARAMETERS.....	90
RECONSTRUCTION LOSS.....	91
AGGREGATION	91
PELAGIC HABITATS	91
COMMENTS AND RECOMMENDATIONS.....	93
ACKNOWLEDGMENTS	94
REFERENCES	95
SUPPLEMENTARY MATERIAL	102
TABLE.....	102
FIGURES	103
CHAPTER III	106
ABSTRACT	106
INTRODUCTION	106
MATERIAL AND METHODS.....	109
DATA ACQUISITION AND PREPARATION.....	109
MODEL DESCRIPTION	110

MICRO-HABITAT SEGREGATION	111
MACRO-HABITAT SEGREGATION	111
IDENTIFICATION OF KEY PARAMETERS	112
SENSITIVITY ANALYSIS	113
RESULTS.....	113
MODEL TRAINING.....	113
IDENTIFICATION OF KEY PARAMETERS	114
SENSITIVITY	115
CHARACTERISATION OF MACRO-HABITATS	116
DISCUSSION	118
MODEL INPUTS	118
GLOBAL MACRO-HABITATS	119
KEY FACTORS BEHIND HABITAT SEGREGATIONS.....	119
BIO-PHYSICAL CHARACTERISTICS OF MACRO-HABITATS.....	120
PLANKTON NICHES.....	122
ACKNOWLEDGMENTS	124
REFERENCES	125
SUPPLEMENTARY MATERIAL	132
GENERAL DISCUSSION.....	142
THE ROLE OF DATA SET SHIFT	142
HOW TO HANDLE RARE SPECIES?.....	142
IS AUTOMATIC CLASSIFICATION SUFFICIENT?.....	143
WHAT DRIVES PLANKTON SPATIAL DISTRIBUTIONS?	144
CONCLUSION AND OUTLOOK	147
REFERENCES	149
OUTLINE OF PUBLICATIONS.....	155
CHAPTER I: PLONUS, R.-M., CONRADT, J., HARMER, A., JANSEN, S. AND FLOETER, J. 2021 ‘AUTOMATIC PLANKTON IMAGE CLASSIFICATION—CAN CAPSULES AND FILTERS HELP COPE WITH DATA SET SHIFT?’, LIMNOLOGY AND OCEANOGRAPHY: METHODS, 19(3), PP. 176–195. DOI: HTTPS://DOI.ORG/10.1002/LOM3.10413	155
CHAPTER II: PLONUS, R.-M., VOGL, S. AND FLOETER, J. 2021 ‘AUTOMATIC SEGREGATION OF PELAGIC HABITATS’, FRONTIERS IN MARINE SCIENCE, 8. DOI: HTTPS://DOI.ORG/10.3389/FMARS.2021.754375	155
CHAPTER III: PLONUS, R.-M., RIETHMOELLER, R. AND FLOETER, J. UNPUBL. ‘IDENTIFICATION OF PLANKTON HABITATS IN THE NORTH SEA’	155
DANKSAGUNG.....	156

Abstract

This thesis deals with multi-dimensional data sets sampled in the pelagic zone of the North Sea. Unlike terrestrial or benthic ecosystems, there are no readily observable differences in the pelagic zone. It is therefore difficult to define habitats or ecological niches and to predict the response of an ecosystem or its community to anthropogenic or environmental pressures. However, this is essential for sustainable management, especially in the face of global warming and increasing anthropogenic impacts on coastal ecosystems, such as offshore wind farms (OWFs). The results of this thesis highlight the potential of machine learning to improve our knowledge of the processes that shape plankton communities, but also its limitations. Fully automated methods are presented to (a) classify *in-situ* plankton images and (b) detect differences in the pelagic zone based on physical and biological measurements. The potential for detecting ecological niches and anthropogenic impacts on a highly dynamic ecosystem such as the North Sea is discussed.

In **chapter I** ('Automatic plankton image classification—can capsules and filters help cope with data set shift?') the potential for automatic classification of *in-situ* plankton images using a Capsule Neural Network (CapsNet) was investigated. Data Set Shift (DSS) in this case describes the problem of shifting plankton communities in both spatial and temporal dimensions. The CapsNet was less affected by DSS than a standard convolutional neural network (CNN), but it also had a lower overall recall, especially for rare classes. The CNN classifications were more affected by DSS, but were still sufficient to reflect the spatial distributions observed in the field, at least in the case of the more abundant groups. For rare classes, an alternative method called 'top-3 accuracy' is proposed to limit human effort while increasing the recall of individual target species to >95%.

In **chapter II** ('Automatic segregation of pelagic habitats') an Autoencoder (AE) was used to detect patterns in a data set consisting of biotic and abiotic variables. Each variable contributed a single value to a multi-dimensional micro-habitat that was projected by the AE onto a two-dimensional plane. The projections were clustered and grouped into macro-habitats consisting of similar micro-habitats. The method consistently identified three distinct pelagic macro-habitats, a 'surface mixed layer', a 'bottom layer' and a 'productive layer'. Distinct plankton communities were observed

in the different macro-habitats. Furthermore, anthropogenic influences induced by an OWF were successfully identified. The predictive power of the variables varied between data sets from different cruises, an indication of the complexity of interacting factors shaping pelagic habitats.

In **chapter III** ('Identification of plankton habitats in the North Sea') I investigated the potential for predicting ecological niches from high-frequency multivariate datasets in the North Sea. The combination of an AE and a density-based clustering algorithm detected several complex habitat patterns, but niche segregation of plankton species at the sub-mesoscale was likely superimposed by local hydrography. Upwelling-downwelling dipoles (in the following simply dipoles) induced by offshore wind farms can develop similar characteristics as naturally occurring frontal systems and improve local productivity. Although of limited applicability at the sub-mesoscale, the model demonstrated the capability for rapid automated processing of multivariate datasets, a key requirement for future research given the increasing amount of data available to marine scientists and the complex dynamics of the pelagic zone.

Zusammenfassung

Die vorliegende Dissertation befasst sich mit der Analyse multidimensionaler Datensätze der pelagischen Zone der Nordsee. Im Gegensatz zu terrestrischen oder benthischen Habitaten gibt es im Pelagial nur wenige offensichtliche Unterschiede. Das macht es schwierig, verschiedene Lebensräume oder ökologische Nischen zu definieren und die Reaktion des Ökosystems oder seiner Bewohner auf anthropogene oder klimatische Einflüsse vorherzusagen. Dieses Wissen ist jedoch für ein nachhaltiges Management unerlässlich, insbesondere im Hinblick auf den Klimawandel und den Einfluss des Menschen auf Küstenökosysteme, z.B. durch Windparks. Die Ergebnisse dieser Dissertation zeigen das Potenzial, aber auch die Grenzen des maschinellen Lernens auf. Zu diesem Zweck habe ich vollautomatische Methoden entwickelt, die (a) *in-situ* Planktonbilder klassifizieren und (b) Unterschiede in der pelagischen Zone basierend auf physikalischen und biologischen Messungen detektieren können. In der anschließenden Diskussion wird auf das Potential der Methoden, ökologische Nischen und anthropogene Einflüsse zu erkennen, näher eingegangen.

In **Kapitel I** ('Automatic plankton image classification - can capsules and filters help cope with data set shift?') wurde das Potential einer neuartigen Modellstruktur, den 'Capsule Neural Networks' (CapsNet), zur Klassifikation von *in-situ* Planktonbildern untersucht. Der Schwerpunkt lag dabei auf der räumlichen und zeitlichen Variabilität von Planktongemeinschaften, ein Phänomen, das im Englischen als 'Data Set Shift' (DSS) bezeichnet wird. Obwohl das CapsNet weniger stark von DSS betroffen war als ein gewöhnliches 'Convolutional Neural Network' (CNN), hatte es einen geringeren Recall-Wert, insbesondere für seltene Gruppen. Die Ergebnisse des CNN, das stärker durch den DSS beeinflusst wurde, reichten dennoch aus, um die räumliche Verteilung der häufigsten Arten im Feld wiederzugeben. Für seltenere Arten wird eine alternative Methode vorgeschlagen, um den manuellen Aufwand zu minimieren und gleichzeitig den Recall-Wert zu maximieren. Diese 'top-3-accuracy'-Methode erreicht einen durchschnittlichen Recall-Wert von >95%.

In **Kapitel II** ('Automatic segregation of pelagic habitats') wurde ein Autoencoder (AE) verwendet, um Muster in einem Datensatz biologischer und physikalischer Variablen zu finden. Ein mehrdimensionales 'Mikrohabitat' wurde durch genau einen

Wert jeder Variablen in diesem Datensatz definiert und vom AE auf eine zweidimensionale Fläche projiziert. Die Projektionen wurden in Gruppen eingeteilt und zu Makrohabitaten zusammengefasst. Auf diese Weise wurden drei eindeutige pelagische Lebensräume identifiziert: eine 'Oberflächenmischschicht', eine 'Bodenschicht' und eine 'produktive Schicht'. In jedem dieser Habitate wurde eine andere Planktonzusammensetzung beobachtet. Zusätzlich konnten anthropogene Einflüsse durch einen Windpark nachgewiesen werden. Die große Variabilität der Relevanz verschiedener Variablen in zwei unabhängigen Datensätzen war ein Hinweis auf die Komplexität der Faktoren, die den pelagischen Lebensraum definieren.

In **Kapitel III** ('Identification of plankton habitats in the North Sea') untersuchte ich das Potenzial zur Identifizierung ökologischer Nischen in der Nordsee mit Hilfe hochfrequenter multidimensionaler Datensätze. Obwohl die Kombination eines AE mit einem dichte-basierten Clusteralgorithmus verschiedene komplexe Habitatmuster identifizierte, war es nicht möglich, Rückschlüsse auf ökologische Nischen für einzelne Planktonarten zu ziehen. Auf der betrachteten sub-meso Skala wurde die Einnischung der Arten wahrscheinlich durch die lokale Hydrographie verhindert. Darüber hinaus zeigte sich, dass durch Windparks erzeugte 'upwelling-downwelling Dipole' ähnliche Eigenschaften wie natürliche Frontensysteme entwickeln und somit die Produktivität des lokalen Systems erhöhen. Obwohl das Modell auf der hier verwendeten Skala nur von begrenztem Nutzen war, konnte ich erfolgreich seine Fähigkeit demonstrieren, multidimensionale Datensätze vollautomatisch zu analysieren. Gerade vor dem Hintergrund der stetig wachsenden Datenmengen, die den Wissenschaftlern zur Verfügung stehen, und der komplexen Wechselwirkungen in der pelagischen Zone des Ozeans ist diese Fähigkeit zur schnellen und automatisierten Datenverarbeitung von besonderem Interesse.

General Introduction

The North Sea is a shallow shelf sea (Huthnance, 1991; Lenhart *et al.*, 1995; Sündermann and Pohlmann, 2011) with a long history of oceanographic research (Lee, 1980; Anderson, 2002; Sündermann and Pohlmann, 2011). It occupies an area of ~575000 km², mainly between the European mainland and the United Kingdom (Eisma *et al.*, 1997). The depth increases from the southern parts towards the north (Huthnance, 1991; Lenhart *et al.*, 1995), but generally remains above 200 m (Eisma *et al.*, 1997). It can be roughly divided into northern, central and southern parts by depth and continentality (Otto *et al.*, 1990). The northern part is strongly influenced by Atlantic inflows (Huthnance, 1991; Lenhart *et al.*, 1995), while the southern North Sea is connected to the Atlantic via the English Channel (Otto *et al.*, 1990; Lenhart *et al.*, 1995; Eisma *et al.*, 1997) and increasingly influenced by river discharge towards the German Bight (Huthnance, 1991; Eisma *et al.*, 1997). The high level of discharge and the low salinity of the Baltic inflows create density differences that produce estuarine-like circulations (Huthnance, 1991; Eisma *et al.*, 1997), where the freshwater discharge flows seaward on top of the saltier marine water (Simpson and Nunes, 1981). This coastal zone is characterised by salinities <34 psu (Fransz *et al.*, 1991) and extends 50-200 km offshore (Lenhart *et al.*, 1995). Typical North Sea water is characterized by a salinity range between 34 and 35 psu and is the result of mixed coastal and Atlantic water (Böhnecke (1922) in Fransz *et al.*, 1991). Within the North Sea boundaries, an anticlockwise flow dominates (Otto *et al.*, 1990; Huthnance, 1991; Eisma *et al.*, 1997; Sündermann and Pohlmann, 2011) with the only outflow to the Atlantic along the Norwegian coast (Sündermann and Pohlmann, 2011). The tidal wave enters the North Sea from both the north and the south and again moves in an anticlockwise direction, with stronger currents and ranges in the shallower parts along the coast (Huthnance, 1991; Eisma *et al.*, 1997; Sündermann and Pohlmann, 2011). Partly due to the strong tidal current the coastal areas remain homogenized throughout the year whereas the deeper parts of the North Sea become seasonally stratified (Otto *et al.*, 1990; Huthnance, 1991; Eisma *et al.*, 1997; Sündermann and Pohlmann, 2011; Leeuwen *et al.*, 2015). In addition to tidal currents, wind-driven circulation can have a significant impact on the hydrography of the North Sea (Huthnance, 1991; Blaas and Swart, 2001; Sündermann and Pohlmann, 2011). Both thermal and haline fronts result in areas of high productivity,

which periodically alternate between high and low oxygen concentrations (Eisma *et al.*, 1997). The high productivity makes the North Sea a particularly important ecosystem for commercial fisheries (Eisma *et al.*, 1997; Kirby *et al.*, 2007).

The exploitation of natural resources was part of humankind since the beginning of our civilization (Gupta, 2004; Tauger, 2010). Marine, and in particular coastal, ecosystems already provide a wide range of ecosystem services of great importance to our society (Peterson and Lubchenco, 1997). More recently, the installation of offshore wind farms (OWFs) to meet the demand for renewable energy has further increased human interest in coastal areas (Carpenter *et al.*, 2016; Corbetta *et al.*, 2016). However, there is still a lack of tools to assess their potential impact on the marine environment (Burkhard *et al.*, 2011; ICES, 2022). Unlike many other marine organisms, plankton species are rarely commercially exploited and their dynamics exhibit a strong relation to environmental changes (Hays *et al.*, 2005). As a result, they have a unique potential to detect shifts in the marine ecosystem (Hays *et al.*, 2005; Lindegren *et al.*, 2012), which can have severe consequences for commercially relevant species (Sguotti *et al.*, 2022; Blöcker *et al.*, 2023).

Copepods are the numerically most abundant members of the holoplankton (Lombard *et al.*, 2010). A comprehensive review of North Sea copepod species is given in Fransz *et al.* (1991), but here I will focus on the most relevant. 85% of the entire copepod biomass in the North Sea is generated by only four species: *Acartia clausi* Giesbrecht, 1889, *Centropages hamatus* Liljeborg, 1853, *Temora longicornis* Müller, 1785, and *Pseudocalanus elongatus* Boeck, 1865 (Hickel, 1975). There are seasonal differences in the abundance of different species, but the by far most abundant is *P. elongatus* (Fransz *et al.*, 1991). *Calanus finmarchicus* Gunnerus, 1770 is another common copepod in the North Sea which has been observed to retreat in more northerly regions with increasing temperatures (Planque and Fromentin, 1996; Falkenhaug *et al.*, 2022). *Calanus helgolandicus* Claus, 1863 on the other hand shows an increasing trend in the North Sea (Planque and Fromentin, 1996; Falkenhaug *et al.*, 2022). It is therefore reasonable to assume that *C. helgolandicus* will become more important under global warming and could also be a good indicator of a regime shift in the North Sea (Beaugrand and Ibanez, 2004), next to other copepod species (Lindegren *et al.*, 2012). A possible mechanism beside

seasonality supporting the high diversity in copepods is species-specific prey selection (Fransz *et al.*, 1991). Copepods are important in the trophic cascade by linking primary producers to higher trophic levels (Cleary *et al.*, 2016; Winder *et al.*, 2017). They are the preferred prey for larvae of commercially important species such as cod (Munk *et al.*, 1995) and planktivorous fish and fish larvae in general (Möllmann *et al.*, 2003). Contrary to long-held beliefs, this position in the food chain is no longer the exclusive preserve of copepods (Harbison and Gilmer, 1976). Contributions of gelatinous zooplankton like salps (Harbison and Gilmer, 1976; Sutherland and Madin, 2010) or appendicularians (Alldredge, 1981) has long been underestimated.

After copepods, the next most abundant group of zooplankton are the appendicularians (Landry *et al.*, 1994). Appendicularia are filter feeders that generally feed on particles <20 μm (Alldredge, 1981). They produce gelatinous filter nets called 'houses' that are discarded several times a day due to clogging or growth of the individual under favorable conditions (Alldredge, 1976; Gorsky and Fenaux, 1998). Although appendicularia reduce their filtering activity at high particle densities, the houses clog more easily and therefore need to be discarded more often (Alldredge, 1976). The houses and fecal pellets sink rapidly and are therefore important for benthopelagic coupling and vertical carbon export in the open ocean (Lombard *et al.*, 2010; Winder *et al.*, 2017). The role of appendicularia in the ecosystem becomes even more interesting considering global warming. As temperatures rise, appendicularia have increased filtering rates (Alldredge, 1981) and shorter generation times (Paffenhöfer, 1976). In the North Sea, they had generation times of around 3 weeks (Paffenhöfer, 1976), allowing them to respond more quickly to rising temperatures than the copepods (Winder *et al.*, 2017) with which they compete for food (Landry *et al.*, 1994). While copepods have diurnal feeding rhythms (Landry *et al.*, 1994) appendicularia feed 24h a day (Alldredge, 1981). Thus, under climate warming and ocean acidification, appendicularia are likely to replace copepods as the most abundant species, with major implications for vertical particle fluxes (Winder *et al.*, 2017) and global ecosystems in general, as benthic communities will have a more ample food supply (Thorpe *et al.*, 2022).

Laakmann *et al.* (2016) identified 31 species of Echinoderms in the North Sea, all of which have a pelagic larvae called pluteus larvae. A shift in the pelagic community from copepod-dominated to pluteus-dominated systems was observed in the 1980s, due to an increase in the macrobenthic biomass (Lindley *et al.*, 1995). Such a regime would be strengthened in a system dominated by appendicularia. Furthermore, and despite the high diversity, most pluteus larvae in the North Sea belong to *Echinocardium cordatum* Pennant, 1777 (Kirby *et al.*, 2007). The adult form of *E. cordatum* benefits from warm winter and spring temperatures (Kirby and Lindley, 2005; Kirby *et al.*, 2007), which will become more likely under warming conditions. An increase in pluteus larvae in a warming ocean is therefore very likely and will further increase competition for copepods.

The examples above already indicate that global warming can have a major impact on pelagic communities and food webs in the oceans (Möller *et al.*, 2015; Möllmann *et al.*, 2021). Unfortunately, the processes that promote and maintain plankton diversity are largely understudied (Brun *et al.*, 2015; Lindegren *et al.*, 2020), and it is therefore essential that we gain a better understanding of the mechanisms that drive the spatio-temporal distribution of plankton (McGinty *et al.*, 2018). The ecological niche was first mentioned in Grinnell (1917) and understanding a species niche will help to understand and forecast their biogeographic distribution (McGinty *et al.*, 2018) along environmental gradients. The fundamental niche, also known as the Grinnellian niche, defines the area where a species can survive (Hutchinson, 1957). The realized or Eltonian niche, usually only a fraction of the fundamental niche, is where a species dominates over its competitors (Elton, 1927). The original assumption of the realized niche was, that if two species occupied the same niche, one would consequently go extinct (Hutchinson, 1961). However, this is not compatible with the idea of functional groups or traits. Functional traits strongly affect organismal performance and are applicable to multiple species at once (McGill *et al.*, 2006), defining their ecological role. Multiple species that can provide the same ecological service are a key requirement for redundancy (Leibold and McPeck, 2006), which is especially important considering the threat of global warming (McGill *et al.*, 2006). However, the investigation of functional traits and realized niches requires data sets that include both, environmental gradients and species compositions, preferably at large scales.

In recent decades, traditional net-based plankton sampling was complemented by modern optical sampling methods. Using net-based samples, skilled taxonomists can sometimes even provide information about zooplankton life stages and sex (Ston *et al.*, 2002; Johansson *et al.*, 2004; Vuorio *et al.*, 2005; Renz and Hirche, 2006; Peters *et al.*, 2013). However, the spatial resolution and coverage is severely limited and the process itself time-consuming (Benfield *et al.*, 2007; Irisson *et al.*, 2022). Fortunately, the strengths and weaknesses of the sampling methods are reversed for modern optical samples. While it is very difficult to obtain information at least down to the family level (Batten *et al.*, 2019), modern sampling devices provide information about plankton at spatial scales of cm and second (Davis *et al.*, 1992; Benfield *et al.*, 2007) along transects from 10's (Floeter *et al.*, 2017) to even 100's (Davis and McGillicuddy, 2006) of nautical miles long. Most recently, the first estimation of a global zooplankton biomass was attempted, thanks to an optical sampling device called UVP5 (Drago *et al.*, 2022). Still, the problem both methods had in common was the extremely long handling time. However, developments in machine learning models begin to erode this problem, at least for the modern optical samplers (Schmid *et al.*, 2020). Hand-sorting millions of images is not practical and the magical threshold of 95 % proposed by Culverhouse *et al.* (2003) for a machine learning model to rival expert plankton ecologists was inaccessible for even the most modern applications. But especially in the last decade there were some breakthroughs. While rare species remain problematic, an application of filtering thresholds can provide a researcher with a reduced, yet accurate data set regarding species composition and distribution for the more common taxa (Faillettaz *et al.*, 2016; Luo *et al.*, 2018). Other than that, the combination of a feature-extracting Convolutional Neural Network (CNN) with a density-based clustering algorithm (HDBSCAN) brings together the speed gained by using machine learning techniques and the knowledge provided by trained human personal (Schröder *et al.*, 2020). And further improvement is to be expected.

This advancement however brought down a new challenge upon the unsuspecting plankton ecologist: how to handle the enormous new resources? Even worse, remotely operated towed vehicles (ROTV) like the TRIAXUS could simultaneously sample multiple different physical parameters like temperature or salinity, generating an even bigger wealth of data (Floeter *et al.*, 2017). Extracting meaningful features

out of the 'sea of data' is a problem in its own right and requires the development of new tools and methods (Alvarez-Berastegui *et al.*, 2014; North *et al.*, 2016). Not only the amount of the data is a problem, but also the sparsity of plankton patches (Richerson *et al.*, 1978; Benoit-Bird *et al.*, 2013). The aggregation of particles around specific features like fronts or eddies is a long-known fact (Munk *et al.*, 1995; Höffle *et al.*, 2013; Munk, 2014; Swalethorp *et al.*, 2015). Still, this produces a similar problem as with the plankton images itself: ~ 90% of all the captured images usually show only detritus or 'marine snow' (Lee *et al.*, 2016; Luo *et al.*, 2018). This huge imbalance is a major problem, not only in classifying plankton images (He and Garcia, 2009). The similarity lies in the relatively small areas of peak abundances of plankton and the comparably vast areas of low abundance (Lombard *et al.*, 2019). On top of that, the pelagic realm is a very dynamic and unstable environment (Hinchey *et al.*, 2008) and simultaneously the downward shift from meso- to submeso- to micro-scale is accompanied by a reduction of the operative scales in time and space at which processes act. This seemingly chaotic system and the unexpected high diversity of zooplankton taxa led to the formulation of the now famous 'paradox of plankton' (Hutchinson, 1961). The seemingly contradiction of many coexisting plankton species despite the absence of relatively stable ecological niches to thrive in, which should therefore extinct each other in their daily struggle to survive.

These are the difficulties that arise if one deals with zooplankton and they all point towards the same fundamental question: is it possible to predict the distribution of plankton species? One possibility to answer this question are habitat maps. Habitat maps link species or communities to physically distinct areas (Harris and Baker, 2012) and are a valuable addition to ecosystem models that support management decisions (Püts *et al.*, 2020). Therefore, the general aim of this thesis was to extract habitat-based information from the high-resolution data available from the use of the ROTV Triaxus and similar devices.

In Chapter I a deep learning CNN was trained to classify plankton images generated by the Video Plankton Recorder (VPR) that was attached to the Triaxus. The performance of the CNN was compared to a new type of model, a Capsule Network (CapsNet), which was a recent alternation from basic NNs (Hinton *et al.*, 2011;

Sabour *et al.*, 2017). It was tested, which of the two models could provide more reliable classifications over an extended period of years since data set shift is a known issue in the processing of plankton images from real-world applications. An existing data set of ~ 124.000 hand sorted images was used to train both models similarly. Both models were based on an existing weight set for the imagenet data set (Deng *et al.*, 2009) which were adjusted to fit the task at hand. Three test sets with different degrees of data set shift were used to evaluate the performance of each model under realistic conditions.

In Chapter II a fully-connected Autoencoder (AE) was trained from scratch to extract meaningful features from a multidimensional data set generated with the Triaxus. Four environmental parameters were selected to generate maps of potential habitats in the pelagic realm. The aim of the chapter was to detect and segregate features like tidal mixing fronts or upwelling-downwelling dipoles in the North Sea. The communities in the predicted habitats were compared using different aggregation indices to see if the identified habitats were occupied by different communities.

Finally, in chapter III both approaches were combined to identify ecological niches. The model trained in chapter I was used to provide information about plankton abundances which were simultaneously processed with the environmental parameters by a new AE. The prospect was that the combination of species abundances and physical parameters would lead to the identification of different plankton communities along the physical gradients. This would help to understand the niches that species occupy and predict changes under changing environmental conditions (e.g. global warming). In addition, key parameters driving habitat segregation and the extent of their influence were identified.

References

- Alldredge, A. L. 1976 'Discarded appendicularian houses as sources of food, surface habitats, and particulate organic matter in planktonic environments: Discarded appendicularian houses', *Limnology and Oceanography*, 21(1), pp. 14–24. doi: [10.4319/lo.1976.21.1.0014](https://doi.org/10.4319/lo.1976.21.1.0014).
- Alldredge, A. L. 1981 'The impact of appendicularian grazing on natural food concentrations in situ', *Limnology and Oceanography*, 26(2), pp. 247–257. doi: [10.4319/lo.1981.26.2.0247](https://doi.org/10.4319/lo.1981.26.2.0247).
- Alvarez-Berastegui, D., Ciannelli, L., Aparicio-Gonzalez, A., Reglero, P., Hidalgo, M., López-Jurado, J. L., Tintoré, J., *et al.* 2014 'Spatial scale, means and gradients of hydrographic variables define pelagic seascapes of bluefin and bullet tuna spawning distribution', *PLOS ONE*, 9(10), p. e109338. doi: [10.1371/journal.pone.0109338](https://doi.org/10.1371/journal.pone.0109338).
- Anderson, E. (ed.). 2002 '100 years of science under ICES: A symposium held in helsinki, 1-4 august 2000', *ICES Marine Science Symposia*, 215. Available at: https://ices-library.figshare.com/articles/report/Report_of_the_ACFM_Study_Group_on_Standard_ICES_Assessment_Computer_Programs/19258286/2 (Accessed: 10 March 2023).
- Batten, S. D., Abu-Alhaija, R., Chiba, S., Edwards, M., Graham, G., Jyothibabu, R., Kitchener, J. A., *et al.* 2019 'A global plankton diversity monitoring program', *Frontiers in Marine Science*, 6. doi: [10.3389/fmars.2019.00321](https://doi.org/10.3389/fmars.2019.00321).
- Beaugrand, G. and Ibanez, F. 2004 'Monitoring marine plankton ecosystems. II: Long-term changes in north sea calanoid copepods in relation to hydro-climatic variability', *Marine Ecology Progress Series*, 284, pp. 35–47. doi: [10.3354/meps284035](https://doi.org/10.3354/meps284035).
- Benfield, M., Grosjean, P., Culverhouse, P., Irigolen, X., Sieracki, M., Lopez-Urrutia, A., Dam, H., *et al.* 2007 'RAPID: Research on automated plankton identification', *Oceanography*, 20(2), pp. 172–187. doi: [10.5670/oceanog.2007.63](https://doi.org/10.5670/oceanog.2007.63).
- Benoit-Bird, K. J., Battaile, B. C., Heppell, S. A., Hoover, B., Irons, D., Jones, N., Kuletz, K. J., *et al.* 2013 'Prey patch patterns predict habitat use by top marine predators with diverse foraging strategies', *PLOS ONE*, 8(1), p. e53348. doi: [10.1371/journal.pone.0053348](https://doi.org/10.1371/journal.pone.0053348).
- Blaas, M. and Swart, H. de 2001 'Large-scale circulation and flushing characteristics of the north sea under various climate forcings', *Climate Research*, 18. doi: [10.3354/cr018047](https://doi.org/10.3354/cr018047).
- Blöcker, A. M., Gutte, H. M., Bender, R. L., Otto, S. A., Sguotti, C. and Möllmann, C. 2023 'Regime shift dynamics, tipping points and the success of fisheries management', *Scientific Reports*, 13(1), p. 289. doi: [10.1038/s41598-022-27104-y](https://doi.org/10.1038/s41598-022-27104-y).
- Brun, P., Vogt, M., Payne, M. R., Gruber, N., O'Brien, C. J., Buitenhuis, E. T., Le Quéré, C., *et al.* 2015 'Ecological niches of open ocean phytoplankton taxa', *Limnology and Oceanography*, 60(3), pp. 1020–1038. doi: [10.1002/lno.10074](https://doi.org/10.1002/lno.10074).
- Burkhard, B., Opitz, S., Lenhart, H., Ahrendt, K., Garthe, S., Mendel, B. and Windhorst, W. 2011 'Ecosystem based modeling and indication of ecological integrity in the german north sea—case study offshore wind parks', *Ecological Indicators*, 11(1), pp. 168–174. doi: [10.1016/j.ecolind.2009.07.004](https://doi.org/10.1016/j.ecolind.2009.07.004).

- Carpenter, J. R., Merckelbach, L., Callies, U., Clark, S., Gaslikova, L. and Baschek, B. 2016 'Potential impacts of offshore wind farms on north sea stratification', *PLOS ONE*, 11(8), p. e0160830. doi: [10.1371/journal.pone.0160830](https://doi.org/10.1371/journal.pone.0160830).
- Cleary, A. C., Durbin, E. G., Ryneerson, T. A. and Bailey, J. 2016 'Feeding by pseudocalanus copepods in the bering sea: Trophic linkages and a potential mechanism of niche partitioning', *Deep Sea Research Part II: Topical Studies in Oceanography*. (Understanding ecosystem processes in the eastern bering sea IV), 134, pp. 181–189. doi: [10.1016/j.dsr2.2015.04.001](https://doi.org/10.1016/j.dsr2.2015.04.001).
- Corbetta, G., Mbistrova, A., Ho, A., Pineda, I. and Ruby, K. 2016 'Wind in power: 2015 european statistics', *European Wind Energy Association (EWEA)*, pp. 1–14. Available at: <https://windeurope.org/wp-content/uploads/files/about-wind/statistics/EWEA-Annual-Statistics-2015.pdf> (Accessed: 9 February 2023).
- Culverhouse, P., Williams, R., Reguera, B., Herry, V. and González-Gil, S. 2003 'Do experts make mistakes? A comparison of human and machine identification of dinoflagellates', *Marine Ecology Progress Series*, 247, pp. 17–25. doi: [10.3354/meps247017](https://doi.org/10.3354/meps247017).
- Davis, C. S., Gallager, S. M., Bermann, M. S., Haury, L. R. and Strickler, J. R. 1992 'The video plankton recorder (VPR): Design and initial results', *Archiv für Hydrobiologie–Beiheft Ergebnisse der Limnologie*, 36, pp. 67–81. Available at: https://www.researchgate.net/publication/284686405_The_Video_Plankton_Recorder_VPR_Design_and_initial_results (Accessed: 5 February 2020).
- Davis, C. S. and McGillicuddy, D. J. 2006 'Transatlantic abundance of the N₂-fixing colonial cyanobacterium trichodesmium', *Science*, 312(5779), pp. 1517–1520. doi: [10.1126/science.1123570](https://doi.org/10.1126/science.1123570).
- Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K. and Fei-Fei, L. 2009 'ImageNet: A large-scale hierarchical image database', in *2009 IEEE conference on computer vision and pattern recognition. 2009 IEEE conference on computer vision and pattern recognition*, pp. 248–255. doi: [10.1109/CVPR.2009.5206848](https://doi.org/10.1109/CVPR.2009.5206848).
- Drago, L., Panaiotis, T., Irisson, J.-O., Babin, M., Biard, T., Carlotti, F., Coppola, L., *et al.* 2022 'Global distribution of zooplankton biomass estimated by in situ imaging and machine learning', *Frontiers in Marine Science*, 9. Available at: <https://www.frontiersin.org/articles/10.3389/fmars.2022.894372> (Accessed: 26 October 2022).
- Eisma, D., Johnston, R., Cadogan, J. I. G., Cadogan, J. I. G., Clark, R. B. and Hartley, J. P. 1997 'The north sea: An overview', *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, 316(1181), pp. 461–485. doi: [10.1098/rstb.1987.0032](https://doi.org/10.1098/rstb.1987.0032).
- Elton, C. S. 1927 *Animal ecology*. New York: Macmillan Co.
- Faillottaz, R., Picheral, M., Luo, J. Y., Guigand, C., Cowen, R. K. and Irisson, J.-O. 2016 'Imperfect automatic image classification successfully describes plankton distribution patterns', *Methods in Oceanography*, 15-16, pp. 60–77. doi: [10.1016/j.mio.2016.04.003](https://doi.org/10.1016/j.mio.2016.04.003).

- Falkenhaus, T., Broms, C., Bagøien, E. and Nikolioudakis, N. 2022 'Temporal variability of co-occurring calanus finmarchicus and c. Helgolandicus in skagerrak', *Frontiers in Marine Science*, 9. Available at: <https://www.frontiersin.org/articles/10.3389/fmars.2022.779335> (Accessed: 10 March 2023).
- Floeter, J., Beusekom, J. E. E. van, Auch, D., Callies, U., Carpenter, J., Dudeck, T., Eberle, S., *et al.* 2017 'Pelagic effects of offshore wind farm foundations in the stratified north sea', *Progress in Oceanography*, 156, pp. 154–173. doi: [10.1016/j.pocean.2017.07.003](https://doi.org/10.1016/j.pocean.2017.07.003).
- Fransz, H. G., Colebrook, J. M., Gamble, J. C. and Krause, M. 1991 'The zooplankton of the north sea', *Netherlands Journal of Sea Research*, 28(1), pp. 1–52. doi: [10.1016/0077-7579\(91\)90003-J](https://doi.org/10.1016/0077-7579(91)90003-J).
- Gorsky, G. and Fenaux, R. 1998 'The role of appendicularia in marine food webs', in *The biology of pelagic tunicates*, pp. 161–169.
- Grinnell, J. 1917 'The niche-relationships of the california thrasher', *The Auk*, 34(4), pp. 427–433. doi: [10.2307/4072271](https://doi.org/10.2307/4072271).
- Gupta, A. K. 2004 'Origin of agriculture and domestication of plants and animals linked to early holocene climate amelioration', *Current Science*, 87(1), pp. 54–59. Available at: <https://www.jstor.org/stable/24107979> (Accessed: 9 February 2023).
- Harbison, G. R. and Gilmer, R. W. 1976 'The feeding rates of the pelagic tunicate pegea confederata and two other salps1', *Limnology and Oceanography*, 21(4), pp. 517–528. doi: [10.4319/lo.1976.21.4.0517](https://doi.org/10.4319/lo.1976.21.4.0517).
- Harris, P. T. and Baker, E. K. 2012 '1 - why map benthic habitats?', in Harris, P. T. and Baker, E. K. (eds) *Seafloor geomorphology as benthic habitat*. London: Elsevier, pp. 3–22. doi: [10.1016/B978-0-12-385140-6.00001-3](https://doi.org/10.1016/B978-0-12-385140-6.00001-3).
- Hays, G., Richardson, A. and Robinson, C. 2005 'Climate change and marine plankton', *Trends in ecology & evolution*, 20, pp. 337–44. doi: [10.1016/j.tree.2005.03.004](https://doi.org/10.1016/j.tree.2005.03.004).
- He, H. and Garcia, E. A. 2009 'Learning from imbalanced data', *IEEE Transactions on Knowledge and Data Engineering*, 21(9), pp. 1263–1284. doi: [10.1109/TKDE.2008.239](https://doi.org/10.1109/TKDE.2008.239).
- Hickel, W. 1975 'The mesozooplankton in the wadden sea of sylt (north sea)', *Helgoländer wissenschaftliche Meeresuntersuchungen*, 27(3), pp. 254–262. doi: [10.1007/BF01611694](https://doi.org/10.1007/BF01611694).
- Hinchey, E. K., Nicholson, M. C., Zajac, R. N. and Irlandi, E. A. 2008 'Preface: Marine and coastal applications in landscape ecology', *Landscape Ecology*, 23(1), pp. 1–5. doi: [10.1007/s10980-007-9141-3](https://doi.org/10.1007/s10980-007-9141-3).
- Hinton, G. E., Krizhevsky, A. and Wang, S. D. 2011 'Transforming auto-encoders', in Honkela, T., Duch, W., Girolami, M., and Kaski, S. (eds) *Artificial neural networks and machine learning – ICANN 2011*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 44–51. doi: [10.1007/978-3-642-21735-7_6](https://doi.org/10.1007/978-3-642-21735-7_6).
- Höfle, H., Nash, R. D. M., Falkenhaus, T. and Munk, P. 2013 'Differences in vertical and horizontal distribution of fish larvae and zooplankton, related to hydrography', *Marine Biology Research*, 9(7), pp. 629–644. doi: [10.1080/17451000.2013.765576](https://doi.org/10.1080/17451000.2013.765576).

- Hutchinson, G. E. 1957 'Concluding remarks', *Cold Spring Harbor Symposia on Quantitative Biology*, 22(0), pp. 415–427. doi: [10.1101/SQB.1957.022.01.039](https://doi.org/10.1101/SQB.1957.022.01.039).
- Hutchinson, G. E. 1961 'The paradox of the plankton', *THE AMERICAN NATURALIST*, p. 10.
- Huthnance, J. M. 1991 'Physical oceanography of the north sea', *Ocean and Shoreline Management*. (North sea: Environment and sea use planning), 16(3), pp. 199–231. doi: [10.1016/0951-8312\(91\)90005-M](https://doi.org/10.1016/0951-8312(91)90005-M).
- ICES 2022 *Working group on offshore wind development and fisheries (WGOWDF)*. report. ICES Scientific Reports. doi: [10.17895/ices.pub.21750458.v2](https://doi.org/10.17895/ices.pub.21750458.v2).
- Irisson, J.-O., Ayata, S.-D., Lindsay, D. J., Karp-Boss, L. and Stemmann, L. 2022 'Machine learning for the study of plankton and marine snow from images', *Annual Review of Marine Science*, 14(1), pp. 277–301. doi: [10.1146/annurev-marine-041921-013023](https://doi.org/10.1146/annurev-marine-041921-013023).
- Johansson, M., Gorokhova, E. and Larsson, U. 2004 'Annual variability in ciliate community structure, potential prey and predators in the open northern baltic sea proper', *Journal of Plankton Research*, 26(1), pp. 67–80. doi: [10.1093/plankt/fbg115](https://doi.org/10.1093/plankt/fbg115).
- Kirby, R. R., Beaugrand, G., Lindley, J. A., Richardson, A. J., Edwards, M. and Reid, P. C. 2007 'Climate effects and benthic–pelagic coupling in the north sea', *Marine Ecology Progress Series*, 330, pp. 31–38. Available at: <https://www.jstor.org/stable/24871087> (Accessed: 2 March 2023).
- Kirby, R. R. and Lindley, J. A. 2005 'Molecular analysis of continuous plankton recorder samples, an examination of echinoderm larvae in the north sea', *Journal of the Marine Biological Association of the United Kingdom*, 85(3), pp. 451–459. doi: [10.1017/S0025315405011392](https://doi.org/10.1017/S0025315405011392).
- Laakmann, S., Boos, K., Kneibelsberger, T., Raupach, M. J. and Neumann, H. 2016 'Species identification of echinoderms from the north sea by combining morphology and molecular data', *Helgoland Marine Research*, 70(1), p. 18. doi: [10.1186/s10152-016-0468-5](https://doi.org/10.1186/s10152-016-0468-5).
- Landry, M., Peterson, W. and Fagerness, V. 1994 'Mesozooplankton grazing in the southern california bight. I. Population abundances and gut pigment contents', *Marine Ecology Progress Series*, 115, pp. 55–71. doi: [10.3354/meps115055](https://doi.org/10.3354/meps115055).
- Lee, A. J. 1980 'Chapter 14 north sea: Physical oceanography', in Banner, F. T., Collins, M. B., and Massie, K. S. (eds) *Elsevier oceanography series*. Elsevier (The north-west european shelf seas: The sea bed and the sea in motion II. Physical and chemical oceanography, and physical resources), pp. 467–493. doi: [10.1016/S0422-9894\(08\)71359-X](https://doi.org/10.1016/S0422-9894(08)71359-X).
- Lee, H., Park, M. and Kim, J. 2016 'Plankton classification on imbalanced large scale database via convolutional neural networks with transfer learning', in *2016 IEEE international conference on image processing (ICIP)*. *2016 IEEE international conference on image processing (ICIP)*, pp. 3713–3717. doi: [10.1109/ICIP.2016.7533053](https://doi.org/10.1109/ICIP.2016.7533053).
- Leeuwen, S. van, Tett, P., Mills, D. and Molen, J. van der 2015 'Stratified and nonstratified areas in the north sea: Long-term variability and biological and policy implications', *Journal of Geophysical Research: Oceans*, 120(7), pp. 4670–4686. doi: [10.1002/2014JC010485](https://doi.org/10.1002/2014JC010485).

- Leibold, M. A. and McPeck, M. A. 2006 'COEXISTENCE OF THE NICHE AND NEUTRAL PERSPECTIVES IN COMMUNITY ECOLOGY', *Ecology*, 87(6), pp. 1399–1410. doi: [10.1890/0012-9658\(2006\)87\[1399:COTNAN\]2.0.CO;2](https://doi.org/10.1890/0012-9658(2006)87[1399:COTNAN]2.0.CO;2).
- Lenhart, H. J., Radach, G., Backhaus, J. O. and Pohlmann, T. 1995 'Simulations of the north sea circulation, its variability, and its implementation as hydrodynamical forcing in ERSEM', *Netherlands Journal of Sea Research*, 33(3), pp. 271–299. doi: [10.1016/0077-7579\(95\)90050-0](https://doi.org/10.1016/0077-7579(95)90050-0).
- Lindgren, M., Dakos, V., Gröger, J. P., Gardmark, A., Kornilovs, G., Otto, S. A. and Möllmann, C. 2012 'Early detection of ecosystem regime shifts: a multiple method evaluation for management application', *PLOS ONE*, 7(7), p. e38410. doi: [10.1371/journal.pone.0038410](https://doi.org/10.1371/journal.pone.0038410).
- Lindgren, M., Thomas, M. K., Jónasdóttir, S. H., Nielsen, T. G. and Munk, P. 2020 'Environmental niche separation promotes coexistence among ecologically similar zooplankton species—north sea copepods as a case study', *Limnology and Oceanography*, 65(3), pp. 545–556. doi: [10.1002/lno.11322](https://doi.org/10.1002/lno.11322).
- Lindley, J. A., Gamble, J. C. and Hunt, H. G. 1995 'A change in the zooplankton of the central north sea (55° to 58° n): A possible consequence of changes in the benthos', *Marine Ecology Progress Series*, 119(1), pp. 299–303. Available at: <https://www.jstor.org/stable/24849819> (Accessed: 1 March 2023).
- Lombard, F., Boss, E., Waite, A. M., Vogt, M., Uitz, J., Stemmann, L., Sosik, H. M., *et al.* 2019 'Globally consistent quantitative observations of planktonic ecosystems', *Frontiers in Marine Science*, 6. Available at: <https://www.frontiersin.org/articles/10.3389/fmars.2019.00196> (Accessed: 8 November 2022).
- Lombard, F., Legendre, L., Picheral, M., Sciandra, A. and Gorsky, G. 2010 'Prediction of ecological niches and carbon export by appendicularians using a new multispecies ecophysiological model', *Marine Ecology Progress Series*, 398, pp. 109–125. doi: [10.3354/meps08273](https://doi.org/10.3354/meps08273).
- Luo, J. Y., Irisson, J.-O., Graham, B., Guigand, C., Sarafraz, A., Mader, C. and Cowen, R. K. 2018 'Automated plankton image analysis using convolutional neural networks', *Limnology and Oceanography: Methods*, 16(12), pp. 814–827. doi: [10.1002/lom3.10285](https://doi.org/10.1002/lom3.10285).
- McGill, B. J., Enquist, B. J., Weiher, E. and Westoby, M. 2006 'Rebuilding community ecology from functional traits', *Trends in Ecology & Evolution*, 21(4), pp. 178–185. doi: [10.1016/j.tree.2006.02.002](https://doi.org/10.1016/j.tree.2006.02.002).
- McGinty, N., Barton, A. D., Record, N. R., Finkel, Z. V. and Irwin, A. J. 2018 'Traits structure copepod niches in the north atlantic and southern ocean', *Marine Ecology Progress Series*, 601, pp. 109–126. doi: [10.3354/meps12660](https://doi.org/10.3354/meps12660).
- Möller, K. O., Schmidt, J. O., St.John, M., Temming, A., Diekmann, R., Peters, J., Floeter, J., *et al.* 2015 'Effects of climate-induced habitat changes on a key zooplankton species', *Journal of Plankton Research*, 37(3), pp. 530–541. doi: [10.1093/plankt/fbv033](https://doi.org/10.1093/plankt/fbv033).
- Möllmann, C., Cormon, X., Funk, S., Otto, S. A., Schmidt, J. O., Schwermer, H., Sguotti, C., *et al.* 2021 'Tipping point realized in cod fishery', *Scientific Reports*, 11(1), p. 14259. doi: [10.1038/s41598-021-93843-z](https://doi.org/10.1038/s41598-021-93843-z).
- Möllmann, C., Kornilovs, G., Fetter, M., Köster, F. W. and Hinrichsen, H.-H. 2003 'The marine copepod, *pseudocalanus elongatus*, as a mediator between climate variability and fisheries in the central baltic

- sea: *Copepod mediates climate to fishery*, *Fisheries Oceanography*, 12(4), pp. 360–368. doi: [10.1046/j.1365-2419.2003.00257.x](https://doi.org/10.1046/j.1365-2419.2003.00257.x).
- Munk, P. 2014 'Fish larvae at fronts: Horizontal and vertical distributions of gadoid fish larvae across a frontal zone at the norwegian trench', *Deep Sea Research Part II: Topical Studies in Oceanography*. (Fronts, fish and top predators), 107, pp. 3–14. doi: [10.1016/j.dsr2.2014.01.016](https://doi.org/10.1016/j.dsr2.2014.01.016).
- Munk, P., Larsson, P., Danielsen, D. and Moksness, E. 1995 'Larval and small juvenile cod gadus morhua concentrated in the highly productive areas of a shelf break front', *Marine Ecology Progress Series*, 125, pp. 21–30. doi: [10.3354/meps125021](https://doi.org/10.3354/meps125021).
- North, R. P., Riethmüller, R. and Baschek, B. 2016 'Detecting small-scale horizontal gradients in the upper ocean using wavelet analysis', *Estuarine, Coastal and Shelf Science*, 180, pp. 221–229. doi: [10.1016/j.ecss.2016.06.031](https://doi.org/10.1016/j.ecss.2016.06.031).
- Otto, L., Zimmerman, J. T. F., Furnes, G. K., Mork, M., Saetre, R. and Becker, G. 1990 'Review of the physical oceanography of the north sea', *Netherlands Journal of Sea Research*, 26(2), pp. 161–238. doi: [10.1016/0077-7579\(90\)90091-T](https://doi.org/10.1016/0077-7579(90)90091-T).
- Paffenhöfer, G.-A. 1976 'On the biology of Appendicularia of the southeastern North Sea'. Available at: <https://www.vliz.be/nl/personen-opzoeken?module=ref&refid=4869&basketaction=add> (Accessed: 7 March 2023).
- Peters, J., Dutz, J. and Hagen, W. 2013 'Trophodynamics and life-cycle strategies of the copepods temora longicornis and acartia longiremis in the central baltic sea', *Journal of Plankton Research*, 35(3), pp. 595–609. doi: [10.1093/plankt/fbt004](https://doi.org/10.1093/plankt/fbt004).
- Peterson, C. H. and Lubchenco, J. 1997 'Marine ecosystem services.', *Nature's services: societal dependence on natural ecosystems.*, pp. 177–194. Available at: <https://www.cabdirect.org/cabdirect/abstract/19991803936> (Accessed: 9 February 2023).
- Planque, B. and Fromentin, J.-M. 1996 'Calanus and environment in the eastern north atlantic. I. Spatial and temporal patterns of c. Finmarchicus and c. helgolandicus', *Marine Ecology Progress Series*, 134(1), pp. 101–109. Available at: <https://www.jstor.org/stable/24856130> (Accessed: 2 March 2023).
- Püts, M., Taylor, M., Núñez-Riboni, I., Steenbeek, J., Stäbler, M., Möllmann, C. and Kempf, A. 2020 'Insights on integrating habitat preferences in process-oriented ecological models – a case study of the southern north sea', *Ecological Modelling*, 431, p. 109189. doi: [10.1016/j.ecolmodel.2020.109189](https://doi.org/10.1016/j.ecolmodel.2020.109189).
- Renz, J. and Hirche, H.-J. 2006 'Life cycle of pseudocalanus acuspes giesbrecht (copepoda, calanoida) in the central baltic sea: I. Seasonal and spatial distribution', *Marine Biology*, 148(3), pp. 567–580. doi: [10.1007/s00227-005-0103-5](https://doi.org/10.1007/s00227-005-0103-5).
- Richerson, P. J., Powell, T. M., Leigh-Abbott, M. R. and Coil, J. A. 1978 'Spatial heterogeneity in closed basins', in Steele, J. H. (ed.) *Spatial pattern in plankton communities*. Boston, MA: Springer US (NATO conference series), pp. 239–276. doi: [10.1007/978-1-4899-2195-6_11](https://doi.org/10.1007/978-1-4899-2195-6_11).
- Sabour, S., Frosst, N. and Hinton, G. E. 2017 'Dynamic routing between capsules', *arXiv:1710.09829 [cs]*. Available at: <http://arxiv.org/abs/1710.09829> (Accessed: 1 April 2019).

- Schmid, M. S., Cowen, R. K., Robinson, K., Luo, J. Y., Briseño-Avena, C. and Sponaugle, S. 2020 'Prey and predator overlap at the edge of a mesoscale eddy: Fine-scale, in-situ distributions to inform our understanding of oceanographic processes', *Scientific Reports*, 10(1), p. 921. doi: [10.1038/s41598-020-57879-x](https://doi.org/10.1038/s41598-020-57879-x).
- Schröder, S.-M., Kiko, R. and Koch, R. 2020 'MorphoCluster: Efficient annotation of plankton images by clustering', *Sensors*, 20(11), p. 3060. doi: [10.3390/s20113060](https://doi.org/10.3390/s20113060).
- Sguotti, C., Blöcker, A. M., Färber, L., Blanz, B., Cormier, R., Diekmann, R., Letschert, J., *et al.* 2022 'Irreversibility of regime shifts in the north sea', *Frontiers in Marine Science*, 9. Available at: <https://www.frontiersin.org/articles/10.3389/fmars.2022.945204> (Accessed: 10 March 2023).
- Simpson, J. H. and Nunes, R. A. 1981 'The tidal intrusion front: An estuarine convergence zone', *Estuarine, Coastal and Shelf Science*, 13(3), pp. 257–IN1. doi: [10.1016/S0302-3524\(81\)80024-2](https://doi.org/10.1016/S0302-3524(81)80024-2).
- Ston, J., Kosakowska, A., Lotocka, M. and Lysiak-Pastuszek, E. 2002 'Pigment composition in relation to phytoplankton community structure and nutrient content in the baltic sea', *Oceanologia*, 44(4). Available at: <http://yadda.icm.edu.pl/yadda/element/bwmeta1.element.agro-article-190160a5-a0a9-4092-8f28-dda8582dec6f> (Accessed: 30 March 2020).
- Sündermann, J. and Pohlmann, T. 2011 'A brief analysis of North Sea physics', *Oceanologia*, 53(3), pp. 663–689. doi: [10.5697/oc.53-3.663](https://doi.org/10.5697/oc.53-3.663).
- Sutherland, K. R. and Madin, L. P. 2010 'A comparison of filtration rates among pelagic tunicates using kinematic measurements', *Marine Biology*, 157(4), pp. 755–764. doi: [10.1007/s00227-009-1359-y](https://doi.org/10.1007/s00227-009-1359-y).
- Swalethorp, R., Malanski, E., Dalgaard Agersted, M., Gissel Nielsen, T. and Munk, P. 2015 'Structuring of zooplankton and fish larvae assemblages in a freshwater-influenced greenlandic fjord: Influence from hydrography and prey availability', *Journal of Plankton Research*, 37(1), pp. 102–119. doi: [10.1093/plankt/fbu099](https://doi.org/10.1093/plankt/fbu099).
- Tauger, M. B. 2010 *Agriculture in world history*. London: Routledge. doi: [10.4324/9780203847480](https://doi.org/10.4324/9780203847480).
- Thorpe, R. B., Arroyo, N. L., Safi, G., Niquil, N., Preciado, I., Heath, M., Pace, M. C., *et al.* 2022 'The response of north sea ecosystem functional groups to warming and changes in fishing', *Frontiers in Marine Science*, 9. Available at: <https://www.frontiersin.org/articles/10.3389/fmars.2022.841909> (Accessed: 30 March 2023).
- Vuorio, K., Lagus, A., Lehtimäki, J. M., Suomela, J. and Helminen, H. 2005 'Phytoplankton community responses to nutrient and iron enrichment under different nitrogen to phosphorus ratios in the northern baltic sea', *Journal of Experimental Marine Biology and Ecology*, 322(1), pp. 39–52. doi: [10.1016/j.jembe.2005.02.006](https://doi.org/10.1016/j.jembe.2005.02.006).
- Winder, M., Bouquet, J., Rafael Bermúdez, J., Berger, S. A., Hansen, T., Brandes, J., Sazhin, A. F., *et al.* 2017 'Increased appendicularian zooplankton alter carbon cycling under warmer more acidified ocean conditions', *Limnology and Oceanography*, 62(4), pp. 1541–1551. doi: [10.1002/lno.10516](https://doi.org/10.1002/lno.10516).

Chapter I

Automatic plankton image classification - can capsules and filters help cope with data set shift?

Rene-Marcel Plonus*^a, Jan Conradt, André Harmer, Silke Janßen, Jens Floeter

Institute of Marine Ecosystem and Fishery Science, Faculty of Mathematics, Informatics and Natural Sciences, University of Hamburg, Hamburg, Germany

Running head: Can capsules and filters cope with DSS?

Keywords: automatic classification, deep learning, data-set shift, Capsule Network, VPR, plankton image data sets, North Sea

*Correspondence: rene-marcel.plonus@uni-hamburg.de

^aPresent address: Institute of Marine Ecosystem and Fishery Science, Hamburg, Germany

Abstract

The general task of image classification seems to be solved due to the development of modern convolutional neural networks (CNNs). However, the high intra-class variability and inter-class similarity of plankton images still prevents the practical identification of morphologically similar organisms. This prevails especially for rare organisms. Every CNN requires a vast amount of manually validated training images which renders it inefficient to train study-specific classifiers. In most follow-up studies the plankton community is different than before and this data set shift (DSS) reduces the correct classification rates. A common solution is to discard all uncertain images and hope that the remains still resemble the true field situation. The intention of this North Sea Video Plankton Recorder (VPR) study is to assess if a combination of a Capsule Neural Network (CapsNet) with probability filters can improve the classification success in applications with DSS. Second, to provide a guideline how to customize automated CNN & CapsNet deep learning image analysis methods according to specific research objectives. In community analyses our approach achieved a discard of uncertain predictions of only 5 %. CapsNet and CNN reach similar precision scores, but the CapsNet has lower recall scores despite similar discard ratios. This is due to a higher discard ratio in rare classes. The recall advantage of the CNN decreases with increasing DSS. We present an alternative method to handle rare classes with a CNN achieving a mean recall of 96 % by manually validating an average of 6.5 % of the original images.

Introduction

State-of-the-art sampling with towed optical devices provides anthropogenic marine planktologists with a wealth of data that even their most recent ancestors could only have dreamed of. Old-school planktologists had to spend hours sitting over the microscope hand-sorting net samples. They were rewarded with snapshots of plankton communities in space and time at the highest possible taxonomic level, sometimes even down to ontogenetic life stages, sex and clutch sizes (Hansson et al., 1990; Ston et al., 2002; Johansson et al., 2004; Vuorio et al., 2005; Renz and Hirche, 2006; Peters et al., 2013).

Modern plankton sampling devices provide information from the other ends of these scales: millions of images at a spatio-temporal resolution of cm and seconds (Davis

et al., 1992; Wiebe and Benfield, 2003; Benfield et al., 2007) sampled continuously over transects 10's (Floeter et al., 2017) or even 100's (Davis and McGillicuddy, 2006) of nautical miles long.

The subsequently necessary automatic plankton image classification has followed the trends in machine learning from Support Vector Machines (SVMs, Hu and Davis, 2005; Sosik and Olson, 2007), later on Neural Networks (NNs, Tang and Stewart, 1996) to modern Random Forest (RF, Bell and Hopcroft, 2008; Orenstein et al., 2015; Faillettaz et al., 2016) and Convolutional Neural Networks (CNNs, LeCun et al., 2015; Krizhevsky et al., 2017), though the use of manually engineered features such as in SVMs is still relatively common (e.g. Nanni et al., 2019). Since the year 2015, when the Microsoft Research Asia team (He et al., 2015) had won the annual ImageNet challenge (Russakovsky et al., 2015) by reaching an accuracy of 96.4 % in classifying high-resolution color images into 1,000 different categories, image classification seemed to be solved (Chollet, 2017). At first sight, plankton images are no exception, because recent efforts have resulted in > 90% average classification accuracy (Al-Barazanchi et al., 2016; Luo et al., 2018).

However, the taxonomic resolution is also almost always diametrically opposed to the increasing scales, providing densities for coarse zooplankton groups such as “jellyfish” or “calanoid copepods” and reaching the family-, or for very distinct organisms the genus-level at best (e.g., *Pseudocalanus* spp. in the Baltic Sea - Möller et al., 2015; Pitois et al., 2018). This is certainly not sufficient for biodiversity monitoring (Batten et al., 2019). However, in many cases coarse groups are suitable for ecological process studies, especially targeting the meso- (Floeter et al., 2017) and microscale (Möller et al., 2012; Ohman et al., 2019).

Further on, the specific success of an automatic plankton image classification task depends on a number of factors: first on the desired taxonomic resolution, i.e. the research question, and second on technical properties as the number of training images and their distribution among classes (e.g. Orenstein et al., 2015).

Additionally, the image quality can have an effect (e.g., how many suspended particles have scattered the flashlight), as the GIGO Principle (Garbage In - Garbage Out) still prevails in times when machines are learning.

Some plankton classes are very abundant while others are scientifically more in focus but rare. Coupled with the usually high intra-class variability and inter-class similarity this leads to the first unsolved problem in real world applications of automatic plankton image classification: the correct identification of rare and/or morphologically similar organisms (Culverhouse et al., 2003; Benfield et al., 2007; Bell and Hopcroft, 2008). The second remaining problem of plankton classification with machine learning methods in production mode applications is related to data set shift (DSS, Moreno-Torres et al., 2012), more specifically in form of “covariate shift” (Moreno-Torres et al., 2012; Webb et al., 2018). DSS can be a problem when, e.g., a machine learning model fitted to images of one region such as the North Atlantic is applied in an apparently similar region in the adjacent North Sea (Webb et al., 2018). Covariate shift is a specification of DSS and can occur when a model that is fitted to images sampled from one plankton distribution needs to be applied to another plankton community sampled some weeks later at the same location.

One approach to cope with these challenges in the production mode application of machine learning methods in plankton image classification is the introduction of probability thresholds, which discards images with uncertain (i.e., likely erroneous) classifications (Faillettaz et al., 2016). This method leads to considerable improvements in average precision but simultaneously to high discard rates of 30-70% of the original images, which artificially changes their abundances (Luo et al., 2018). As some of the discarded images were correctly identified objects, the recall (i.e., the proportion of the true total number of objects of a class that are correctly predicted in that class) is reduced. The resulting key question is whether any subsequent analyses still yield ecological patterns that resemble the truth (Faillettaz et al., 2016; Luo et al., 2018). This is usually fulfilled for research questions that target common taxa at coarse spatial resolutions. When validated images in the order of magnitude of the test data set are easily obtainable for each new data set, a multiplication factor can be computed from the F-score based confusion matrix to calculate post-filtering corrected concentrations (Hu and Davis, 2006; Briseño-Avena et al., 2020; Schmid et al., 2020).

However, when the scientific focus is on rare organisms or alpha biodiversity, recall is more important than mean precision and a filtering method may be impedimental.

The second main challenge is the consistency of model performance over time and space, i.e. data set drift (Bell and Hopcroft, 2008; Al-Barazanchi et al., 2016; González et al., 2017). The Capsule Network (CapsNet) is a recently developed machine learning architecture (Hinton et al., 2011; Sabour et al., 2017), which could overcome this issue. CapsNets group neurons into so called Capsules, which learn specific properties of an object or segment such as size or rotation. This makes the predictions of a CapsNet invariant to the viewpoint, i.e., variations in position and orientation, and to variations in scale and lighting. It can theoretically improve the performance on overlapping objects, thus it could be useful to detect e.g. grazing interactions with marine snow particles (Möller et al., 2012). Instead of dropout layers a CapsNet uses a reconstruction autoencoder for regularization. This autoencoder should be able to reconstruct an object of the predicted class based on the features learned for that class (Sabour et al., 2017; Xi et al., 2017). So far CapsNets have been successfully applied to “baseline” data sets such as MNIST or CIFAR10 (Sabour et al., 2017; Xi et al., 2017; Rajasegaran et al., 2019) but only to a limited number of “real-world” applications such as brain tumor recognition (Afshar et al., 2018).

The theoretical advantages of the CapsNet over a common CNN led us to the assumption, that a CapsNet should be able to adapt better to changing field conditions and therefore yield better results in production mode applications. By following González et al. (2017) recommendations for the development of unbiased input data sets reflecting the class distribution in the field we describe the whole training process for a deep learning CNN and a CapsNet to classify plankton images in 26 different classes. This includes pre-processing, classification and postprocessing of the images. Subsequently, we apply our models in production mode, i.e. without updating the training, to three different North Sea field data sets with increasing temporal and structural distance.

In our analysis we demonstrate how the filtering method and a CapsNets can help coping with DSS in automatic plankton image classification. Specific research tasks typically focus on predicting broad scale plankton community properties in unseen samples or on classifying each image correctly, also for rare organisms.

To assess whether filtering methods and CapsNets can be customized to successfully cope with data set and covariate shift, we compare two different scenarios: a baseline scenario (BL) without any filtering and a high precision scenario (P95) with probability filters aimed to maximise precision in a fully automated analyses of plankton communities. Second, we show how to customize the method to maximise the recall for classes, individually, supporting studies focusing on specific classes exclusively. To measure the potential advantage of the CapsNet, we compare the results of a simultaneously trained CNN with those of our CapsNet.

Material and procedures

Description of instrument

We used a Video Plankton Recorder (VPR, Seascan Inc., Falmouth, MA02540, USA) mounted on a MacArtney TRIAXUS ROTV which was connected to a research vessel with a fibre optic cable to record high resolution images of in-situ plankton organisms. The ROTV was towed at a speed of 8 knots (4.1 m s^{-1}) with a three-degree lateral offset to lessen any disturbance from the vessels wake. During most transects the ROTV was undulating with a vertical speed of 0.1 m s^{-1} from $\sim 4 \text{ m}$ below the sea surface to $\sim 8 \text{ m}$ above the sea floor. The VPR was equipped with a high-resolution digital camera (Pulnix TM-1040) that records up to 25 fps. A synchronized strobe (Seascan - 20 W Hamamatsu xenon bulb) provided the illumination for the images at a pulse of $1 \mu\text{s}$. The resulting images consist of 1392×1024 pixels with a size of $9.0 \times 9.0 \mu\text{m}$. The chosen field of view was $24 \times 24 \text{ mm}$ with a focal depth of $\sim 60 \text{ mm}$ at 246 mm from the lens. The image volume was thus 34.93 mL . Imaged particles were extracted as regions of interest (ROIs) by the Autodeck image analysis software (Seascan Inc.) and saved to the computer hard drive as TIFF files.

Description of hard- and software

The German Climate Computing Center (DKRZ) provided computing time with the High Performance Computer System for Earth System Research (HLRE-3, Mistral), which consists of more than 3,000 compute nodes, providing a peak compute performance of 3.6 PFLOPs and was used to train our models. We used two Mistral computing nodes (2x 18-core Intel Xeon (E5-2695 v4) with a single Nvidia Tesla

V100 GPU and 512 GB RAM) for the training of our models

(<https://www.dkrz.de/up/systems/mistral/configuration>, 2020 August 04. 16:35:40).

Training and application of deep learning models were done with a GPU supported Tensorflow (Abadi et al., 2015) backend for Keras (Chollet, 2015) under Python 3.7 (Van Rossum and Drake, 2009). Subsequent data analyses were done with the statistical package R (RCoreTeam, 2020). Visualisations were created using ggplot2 (Wickham, 2016) while data management was mainly done with dplyr (Wickham et al., 2020). We calculated the t-test modified by Dutilleul (Dutilleul et al., 1993) using SpatialPack (Osorio and Vallejos, 2019). The Bray-Curtis dissimilarity (Bray and Curtis, 1957) between the validated training set and the predicted field sets was calculated using the implementation in vegan (Oksanen et al., 2019).

Field sampling

We used ~ 124,000 hand sorted and labelled images to train our models, of which ~ 90 % were sampled on the FS Heincke cruise HE446 on the 4th of June 2015 between 07:00 and 13:00 (UTC). The remaining 10 % of the images originate from the period June to August of 4 years (2014 - 2017) and cover all 24 hours of a day. Most of our images (94 %) were sampled in the inner German Bight of the North Sea, including the three unlabelled field data sets (Figure 1) which we used in our production mode analysis. The remaining 6 % originated from the Baltic Sea and provided images for the classes 'eggs' and 'larvae' which were not represented otherwise.

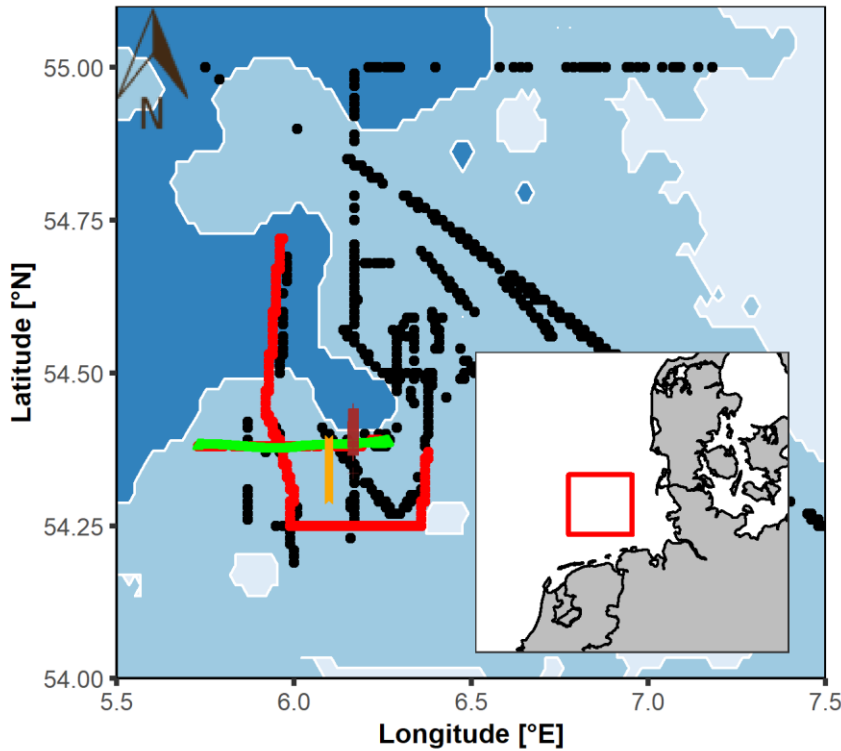
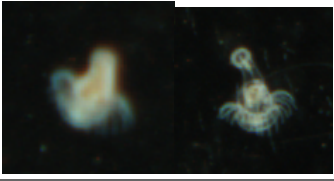
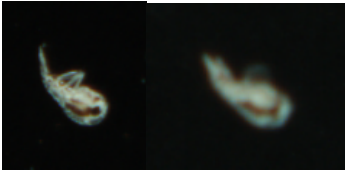
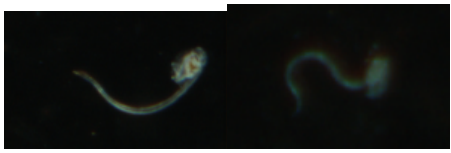
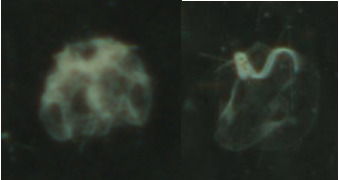
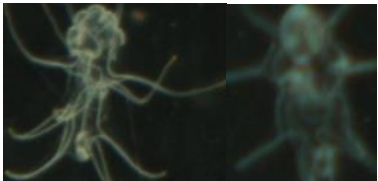
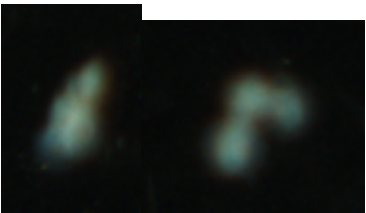
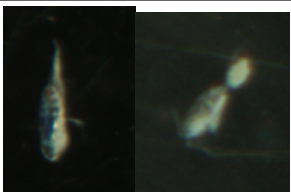

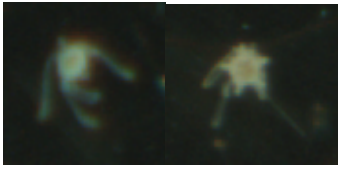


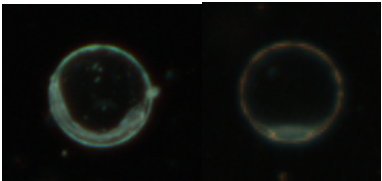
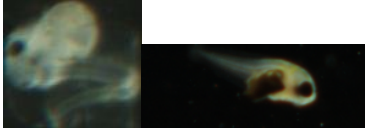
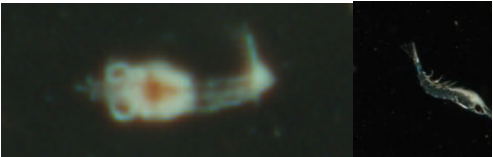
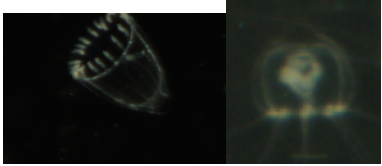

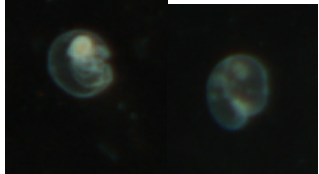
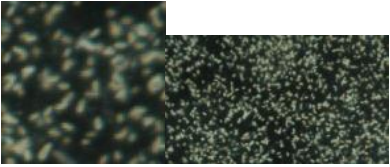
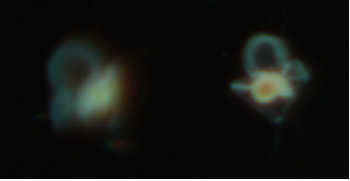

Figure 1: Core area of the sampling transects with the VPR in the North Sea. The red transect provided ~90 % of the labelled training images. black: remaining training set; field sets - green: FS446 (#1; 2015; 55,302 images); brown: FS466 (#2; 2016; 7,798 images); orange: FS534 (#3; 2019; 31,848 images). Blue shading: depth contours from 20 to 50 m.

Evaluation of the performance consistency of our final classifiers in production mode was done using the three field data sets. The similarity of the unvalidated classifications (i.e. predictions) of the field data sets and the training set (TS) was assessed calculating the Bray-Curtis dissimilarity. Field set 1 (FS446) originated from the same HE446 cruise in 2015 as the majority of our training images. However, the field set was sampled in the morning from 06:00 to 09:00 (UTC, 55,302 images). The second field set (HE466) was sampled in June 2016 between 18:00 and 19:00 (UTC, 7,798 images). The third field set (HE534) was sampled in June 2019 from 11:00 to 12:00 (UTC, 31,848 images). All field set model predictions were manually checked and if necessary corrected by a human zooplankton expert to obtain the “true” classification.

Table 1: Models were trained on 26 classes, including 2 classes with none-living objects (‘marine snow’ and ‘rod’) and 2 classes for unrecognized objects (‘unknown’

and 'blurry'). The numbers for the training- (TS) and field sets (FS) correspond to the 'true' distribution obtained by manual classifications.

class	label	TS [N]	FS446 [N]	FS466 [N]	FS534 [N]	example image
actinotrocha	act	208	0	5	11	
amphipods	amp	241	0	0	0	
appendicularia	app	545	28	147	198	
appendicularia with house	app	837	270	204	899	
bipinnaria	bip	473	79	11	8	
blurry	blu	1,187	101	1,313	1,014	
copepods	cop	2,258	151	627	2,219	
diatoms	dia	6,984	3,190	346	16	
echinodermata	ech	100	0	0	0	

class	label	TS [N]	FS446 [N]	FS466 [N]	FS534 [N]	example image
eggs	egg	416	5	0	0	
larvae	lar	230	15	7	1	
malacostraca	mal	376	22	43	87	
medusae	med	394	76	144	24	
mnemiopsis	mne	739	144	1	7	
noctiluca	noc	834	348	20	3,696	
phaeocystis	pha	224	0	0	0	
pilidium	pil	142	0	56	19	
pluteus	plu	14,713	9,861	212	1,343	

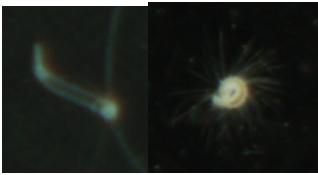

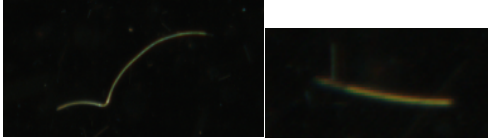
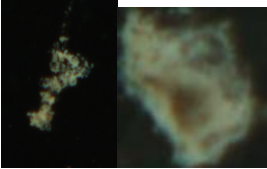
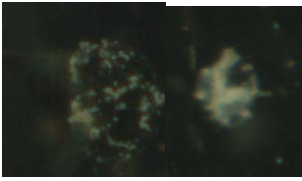
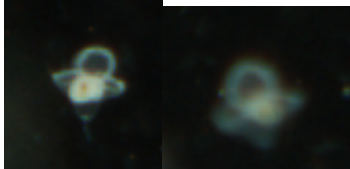
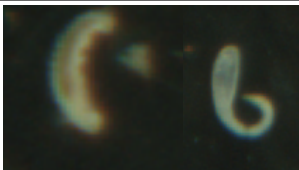
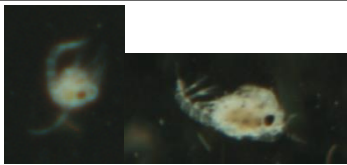
class	label	TS [N]	FS446 [N]	FS466 [N]	FS534 [N]	example image
polychaeta	pol	802	363	16	22	
pteropods	pte	587	0	0	0	
rod	rod	264	2,034	814	22	
marine snow	sno	68,311	37,675	3,578	20,519	
unknown	unk	509	208	176	758	
veliger	vel	249	0	34	47	
worms	wor	2,103	705	37	913	
zoea	zoe	274	27	7	25	

Image pre-processing

As CNNs require equally sized images, based on our most common regions of interest (ROI) size we chose a size of 240 x 240 pixel. Smaller images were extended by placing the original ROI image in the center and adding pixels. The new pixel values were set to the median of the ROI pixel values since most ROI images

were mainly filled with black background. ROI images that were greater than 240 pixels in one or both dimensions were first squared by adding pixels with median values to keep the proportionality of the objects before rescaling them to 240 x 240 pixels. Initial experiments showed that classification accuracy did not benefit from coloured images, so we transformed our images to greyscale by multiplying the rgb values with 0.299 (r), 0.587 (g) and 0.144 (b). The resulting matrix was replicated two times to create the required three channel image input format. When fed to the model, images greater than 240 pixels were reduced to 240 pixels.

Images were fed to the model in small batches using an Image Data Generator function provided by Keras. The batch size was adapted to the respective model and image set (1 - 40). Since deep learning models usually perform better with homogeneous, small values (Bishop, 1995) all pixel values were divided by 255. Data augmentation was applied during the training but not in the validation and test step. Images were rotated, shifted in both directions, sheered, zoomed and horizontally flipped randomly with fill mode set to "nearest". This was done to increase the generalization of the deep learning model by providing slightly altered images during each training cycle (= epoch).

While CapsNets don't necessarily need data augmentation to achieve the performance of similar CNNs trained with data augmentation (Jiménez-Sánchez et al., 2018), data augmentation nevertheless can increase the performance especially for small classes (Toraman et al., 2020). Thus, we also applied data augmentation during the training with the CapsNet.

Automated image classification

Workflow

We combined a two-step training procedure suggested by Lee et al. (2016) and the application of different filtering thresholds, as suggested by Faillettaz et al. (2016), to optimize our model performance. In step 1, the model was trained with a balanced data set and rated according to the performance with a balanced test set, both of which were subsets of the imbalanced labelled data set. On this basis we continued to train the same model with an imbalanced training set, using the final weights from step 1 in the initialization and all available labelled images. As is common practice in deep learning, we split the entire data sets into training-, validation- and test-subsets.

Based on the predictions for the imbalanced test set, we calculated filter values which can be applied to tailor the results for specific research questions in production mode, i.e., application to new field data sets without updating the training procedure.

Convolutional neural network

We used the convolutional base (ConvBase) of the Xception V1 model with weights pre-trained on ImageNet available for download using the Tensorflow backend from Keras (<https://keras.io/api/applications/xception/>, 2019 December 10. 15:06:15). The input size was changed to 240 x 240 pixel from 299 x 299 pixel. We added an additional convolutional layer (ConvLayer - SeparableConv2D) using the Keras functional API before the flatten operation and the final Dense-Layer. The ConvLayer had a convolutional window with kernel size 3 x 3 and padding set to “same”. We chose “Rectified Linear Unit” as activation function. The resulting filter stack of 2,560 filter maps with size 8 x 8 was flattened and the final Dense-Layer with softmax activation was used to classify the images into 26 different classes. The final model had 30,385,218 parameters (Figure 2).

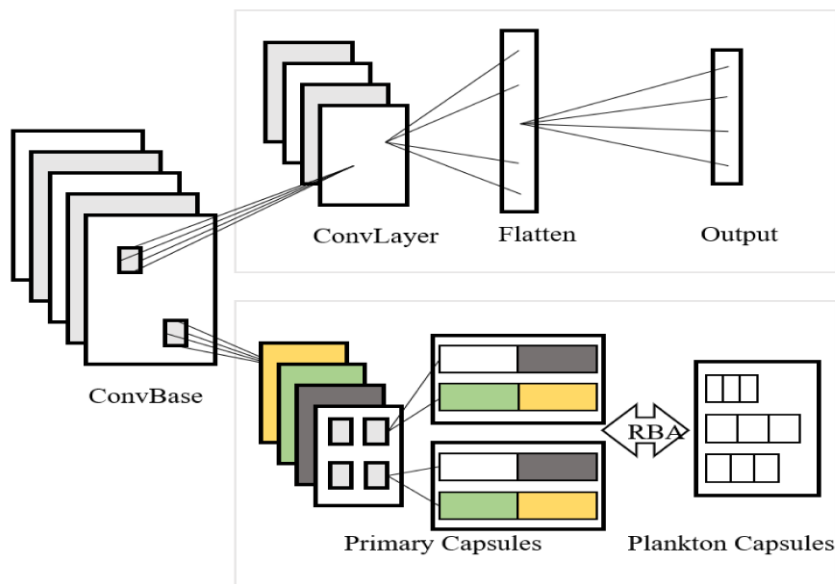


Figure 2: Our model architectures. Both models were based on the ConvBase of the Xception V1 (for details see Chollet (2017)). We changed the input size to 240 x 240, so the output of the ConvBase was a feature stack of 2,048 filter maps with size 8 x 8. CNN: a convolutional layer with kernel size 3 x 3 returning 2,560 filter maps

followed by a flatten layer. The final softmax layer had length 26 (for 26 classes). CapsNet: a convolutional layer with 240 kernels of 6 x 6 and strides 2. The output was grouped into 80 primary capsules with 24 dimensions, each of which represented 1 property of the feature learned by the respective capsule. Plankton Capsules returned 26 capsules (one for each class) with 16 dimensions (one for each property) after 3 cycles of routing-by-agreement (RBA).

Training the models

In a first training step we used only 100 images of each class (2,600 images in total) for the training set and 10 images for validation (260) and testing (259) respectively. The smallest class had only 9 test images. The validation set is used to monitor the ability of the model to generalize during the training process, while the test set is a final evaluation prior to the application in production mode. ConvLayers “learn” by applying small weights to each input. Those weights store the “learned” information. We successively adapted more layers of the pre-trained ConvBase to our images during training, starting with the topmost (last) layers, going deeper in each successive phase (Table 2). This is called “transfer learning” (Pan and Yang, 2010; Kornblith et al., 2019). The Keras callback “ReduceLROnPlateau” was used with patience 2 and factor 0.6 and the weights of the best model achieved during training were saved by another callback “ModelCheckpoint”. Using the Adam optimizer (Kingma and Ba, 2014) and a categorical crossentropy loss function the model was trained with an initial learning rate of 2×10^{-5} (CNN) or 5×10^{-5} (CapsNet), using accuracy for evaluation.

Table 2: Training procedure for the CNN and the CapsNet. Training the CNN was initialized using weights pre-trained on ImageNet, while the training of the CapsNet was initialized using the weights received at the end of step 1 with the CNN. Due to overfitting, the CapsNet was trained only for 4 phases in step 2, while the CNN was trained for 5 phases. The ConvBase had 40 layers in total.

model	training step	training phase	epochs	ConvBase trainable layers
CNN	1	1	7	2
CNN	1	2	7	11
CNN	1	3	7	20
CNN	1	4	7	29

model	training step	training phase	epochs	ConvBase trainable layers
CNN	2	1	3	2
CNN	2	2	3	11
CNN	2	3	3	20
CNN	2	4	3	29
CNN	2	5	5	35
Cap	1	1	7	2
Cap	2	1	3	2
Cap	2	2	3	11
Cap	2	3	4	20
Cap	2	4	3	29

In a second training step the same model was initialized with the final weights from step 1 and trained on a heterogenous data set. The distribution of the training images represented the distribution observed in the labelled data set (Table 1). 84 % of all images in a class were used as training set and 8 % as validation and test set respectively. The smallest class (echinodermata) had 100 (0.1 %) training images while the largest (marine snow) had 68,311 (65.7%). We used class weights (CW, eq. 1) to account for this imbalance:

$$(1) CW_i = \log\left(\frac{N_{max}}{N_i}\right)$$

The CW of class i was calculated as natural logarithm of the ratio of the maximum number of images over all classes (N_{max}) and the number of training images of class i (N_i). The CW of the largest class marine snow was set to 1 and the CWs of the other 25 classes increased logarithmically with decreasing number of available training images up to a factor of 6.5 for the smallest class 'echinodermata'. Again, we used 'transfer learning' to benefit from the features learned during the first training step, especially in less abundant classes.

The first training needed a computing time of ~ 20 minutes while second training required ~ 24 hours (1,440 minutes) for the CNN and ~ 21 hours for the CapsNet.

As CNNs are a gradient-based method, the chosen starting point may be crucial for the final fit of the model, and one way of assessing and reducing the effect of start

conditions are multi-start approaches (Subbey, 2018). We repeated the first training step 100 times with randomly changed image sequences fed into the CNN. The second training was only performed once based on the model from step 1 which achieved the best test accuracy. Repeating the second training step was not feasible due to long computing times.

The same best step 1 CNN model was used to train the CapsNet. Before we started the training with the heterogenous data set in the same way as described for the CNN, we repeated step 1 once in a reduced form (Figure 3) to adjust the weights of the last three layers of the ConvBase to the new Capsule-Layers, which replaced the Dense- and Flatten-Layers used in the CNN (Figure 2).

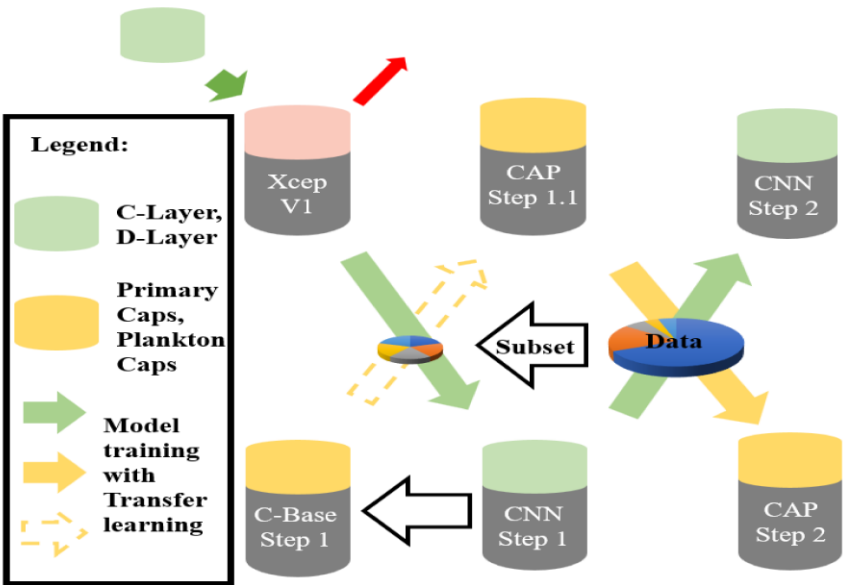


Figure 3: Schematic visualisation of model training. Training step 1 was repeated for the Capsule Network in a reduced form (Table 2), which is indicated by the dashed arrow. Both our models shared the ConvBase with the Xception V1 model, but we replaced the final classifying layers with our own choice of layers as indicated by the coloured part of the cylinders and the red and green arrow in the top left corner. C-Base: Convolutional Base, C-Layer: Convolutional-Layer, D-Layer: Dense-Layer, Xcep V1: Xception V1, CNN: Convolutional Neural Network, CAP: Capsule Network.

Model performance

To compare the different models, we calculated the class-wise as well as average training-, validation- and test-accuracies (Acc), which is the percentage of correctly classified images. Training-, validation- and test-accuracies are related to the respective image subsets. In case of the balanced data set for the first training step that means 100 images per class for training and 10 images for each, validation and testing. While this is sufficient for homogenous data sets such as the one used during the first training, accuracy fails to account for the imbalance in a heterogenous data set as used during the second training step (González et al., 2017). Therefore, we also calculated the F1 score (eq. 2), which is the harmonic mean of the classification metrics precision (P - “purity”, eq. 3) and recall (R - “completeness”, eq. 4) and is more sensitive to wrong predictions in highly skewed data (He and Garcia, 2009).

Precision and recall are calculated using true positives (TP), false positives (FP, type I error) and false negatives (FN, type II error). A correctly identified copepod image in the copepod class is a TP. A copepod identified as a diatom is a FN for the copepod class and at the same time a FP for the diatom class. FPs and FNs are class specific and make sense only from the viewpoint of the respective class. Images which truly belong to a class, though they are sorted into other classes, count as FNs, while all images which don't belong to a class, though they are sorted into that class, count as FPs. Precision is the proportion of correctly classified objects in a predicted class and recall is the proportion of the true (i.e., manually labelled) number of objects of a class that are correctly predicted in that class.

$$(2) F1_i = \frac{2 \times P_i \times R_i}{P_i + R_i}$$

$$(3) P_i = \frac{TP_i}{TP_i + FP_i}$$

$$(4) R_i = \frac{TP_i}{TP_i + FN_i} (= Acc_i)$$

Probability filtering and top-k predictions

All 100 models of the first training step were used to create predictions for an identical test set of 259 images to assess the final model performance. After the second training step, a labelled test set of 9,903 images was used to validate the

model performance on new, unseen data sets. The three unlabeled field sets were only predicted (i.e. classified) once by the final model after the second training step, since initial results suggested that field set predictions after step 1 were not meaningful. For a given image, CNNs compute a probability for each class. In an ideal case, the class with the highest probability resembles the true taxonomic class of the imaged object. The filtering method of Faillettaz et al. (2016) takes advantage of those probabilities and accepts only predictions above a user specified threshold. The assumption is, that TPs have a higher probability than FPs and thus more wrong than correct predictions are discarded, ultimately increasing the precision. We used the labelled test set to calculate probability filters for each class individually. All images assigned to a class (TP + FP) were sorted in increasing order of their probability to belong to that class. All images with a lower probability than the chosen threshold but correctly classified as class i were then nevertheless treated as FN's and thus decreased the recall and subsequently the F1-score. Precision was calculated using only images with a higher probability than the chosen threshold, since the "purity" of a class can only be affected by images assigned to this class. Therefore, FP's with a lower probability than the chosen threshold had no influence on the calculated precision. This method can of course only be applied with data sets that have been manually validated and labelled to obtain the "true" classifications.

Tailoring to specific research questions

During the validation of the final model the filters were stepwise increased from the lowest to the highest probability and the corresponding classification metrics were calculated. This enables the researchers to pick their favorite set of class-specific filters along the trade-off continuum between the best average precision and the best recall. In Luo et al. (2018), classes with $n < 25$ of 75,000 randomly drawn images were excluded to achieve a mean precision of 90.7 %. This threshold ($n = 25$) divided classes into a 'pure' (precision > 90 %) and an 'uncertain' (precision < 90 %) group. In addition to the class-wise filters aiming to maximise the precision (P95) we chose for each model and field set a class unspecific threshold (t) of n images to separate a 'pure' group of classes (mean precision > 90 %) from a 'uncertain' group of classes (mean precision < 90 %). In a larger scale community distribution-oriented research question, this sorts classes classified on a human-like level into the 'pure' group and leaves classes with poor performances in the 'uncertain' group (Luo et al., 2018). The

threshold was chosen based on the sum of the TP's and the FP's of the predictions of the three field sets ($n = TP + FP$). 100 bootstraps were performed using randomly chosen 75 % of the images that remained after P95-filtering to increase the confidence in the chosen threshold.

However, a reduced recall is problematic if rare taxa like fish larvae are specifically in the focus of the research question. Thus, we assessed whether the deep-learning practice of the Top-5-Accuracy can be used to increase the recall and significantly reduce the time needed for manual classification. We accepted the k highest predictions for each image, stepwise increasing k from 2 to 5, and treated an image as 'correct classified' if the correct class was assigned within one of the top-k probabilities. Subsequently, the user has to manually classify all top-k images in the classes which are in focus of the research question. In this case, the trade-off between the recall and the number of images that have to be manually classified is of particular interest.

Representativeness of field set classifications

For research questions involving the detection of ecological patterns in high frequency data sets, particularly for common taxa, precision could be more important than recall (Faillettaz et al., 2016). This arises because the distribution of images could resemble the field plankton community even when large fractions of images that cannot be classified with sufficient certainty are discarded.

As in Faillettaz et al. (2016), we tested the spatial distributions of our filtered predictions against the spatial distribution of the manual classification using the t-test modified by Dutilleul (Dutilleul et al., 1993). We aggregated our data in 1 m depth bins and by Latitude (0.01 decimal degree (DD) bins) for North-South transects or Longitude (0.01 DD bins) for West-East transects. Since the filtered predictions are per definition a subset of the original data set, we compared relative abundances instead of the absolute ones as Faillettaz et al. (2016) suggested.

Results

Model training

The dynamics of training and validation accuracy of the 100 models during the first CNN training step were slightly different each time, despite the fact that all were trained in the same way and with the same training images. At the end of step 1, the training accuracies (mean: 75 %) usually slightly exceeded the validation accuracies (mean: 71 %; Figure 4). Further increasing the number of adjustable layers or training epochs only led to strong overdispersion, which indicates decreasing generalization of the model. Even though the general trends during the 100 training runs were similar, the final test accuracies ranged from 54.4 % to 84.6 %, indicating that different runs produced different convergence progressions and therefore different outcomes. The oppositional pattern of validation- and trainings-accuracy between phase 1 and all following phases was probably due to the fact that the model was trained to the verge of overfitting in each phase and thus was already close to overfitting when training started in phase 2 (and following phases).

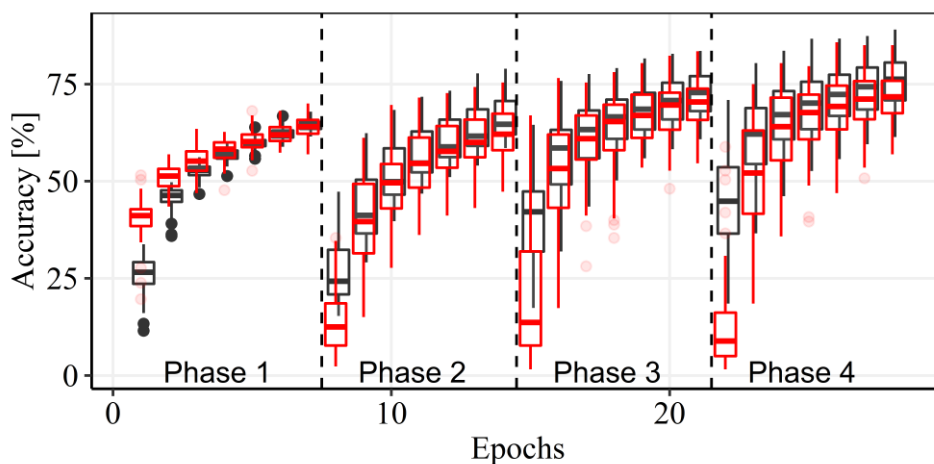


Figure 4: Convolutional neural network training step 1. The vertical lines separate the different training phases 1-4 during the first step (training with homogenous data set), where successively more layers were trained in each phase. Black boxplots: training accuracy of 100 models; red boxplots: validation accuracy of 100 models.

The training progress during the second step using the heterogenous data set differed between the CapsNet and the CNN. The training and validation accuracy of the CNN increased gradually to ~ 95 % ($Acc_{Tr} = 95.2$; $Acc_{Val} = 94.9$). The final test

accuracy was 94.67 %. The CapsNet instead deteriorated during the last of the five phases, even though more weights were released for training compared to the phase before. While training ($Acc_{Tr} = 95.9\%$) and validation ($Acc_{Val} = 94.3\%$) accuracy were similar at the end of phase 4, the validation accuracy decreased to $Acc_{Val} = 89.5\%$ in phase 5 while the training accuracy increased further to $Acc_{Tr} = 96.3\%$ (Figure 5). The CapsNet reached a slightly lower test accuracy of 89.72 %.

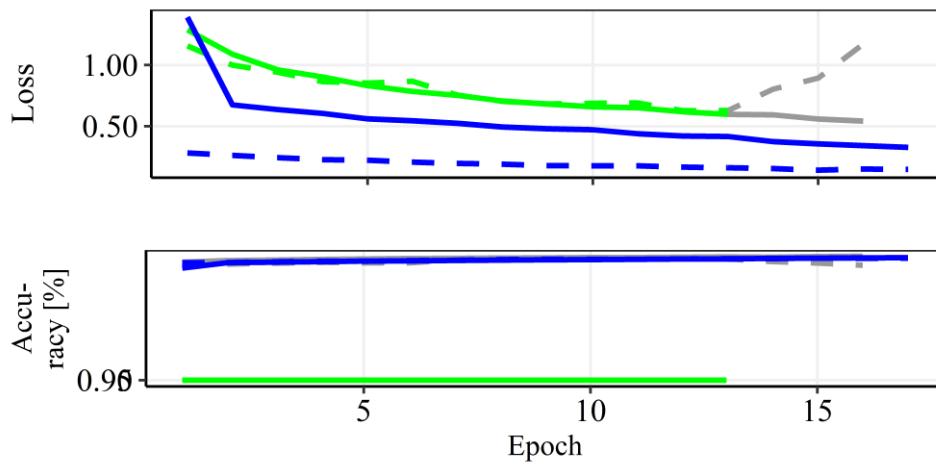


Figure 5: Training step 2, heterogenous data set. Upper panel shows the loss, which is an index of the difference between prediction and truth. The lower panel shows the accuracies as for training step 1. Solid: during training; dashed: during validation; green: CapsNet; blue: CNN; grey: period of overdispersion, the changes during these epochs were not included in the final model.

Test set predictions

Classes with a high mean test accuracy over all 100 models in step 1 had small confidence intervals (CI; 95 %), while classes with lower mean test accuracies could range from 0 to 80 % correctly classified images, depending on the model run, even though all classes were trained with the same amount of 100 images (Figure 6).

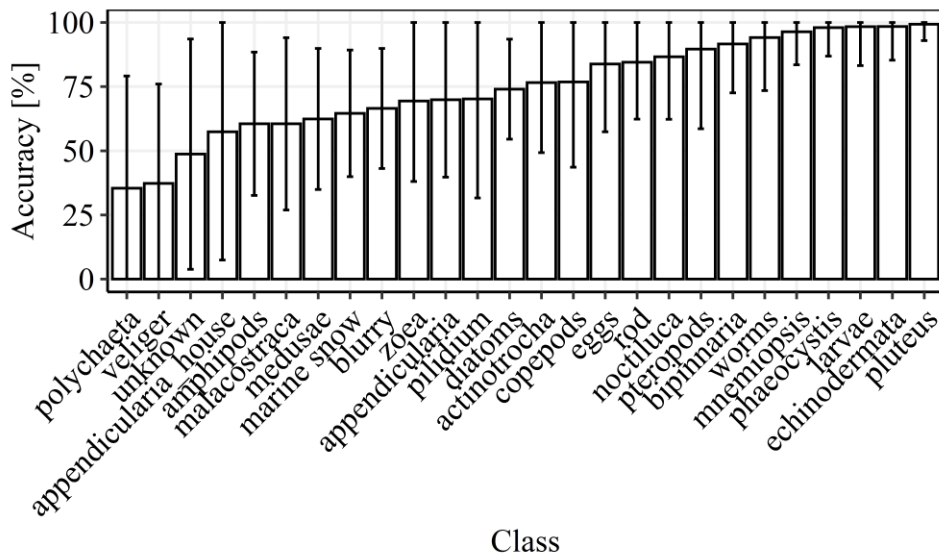


Figure 6: Increasingly sorted mean test accuracies for the 26 classes based on the predictions of the 100 models trained during step 1. Error bars: 95% confidence interval (CI).

In step 2, classes with high abundances generally achieved high F1-scores, whereas the opposite was not true as low abundant classes could have low, medium or high F1-scores (Figure 7). In general, the CNN achieved better results than the CapsNet after step 2. However, the CapsNet outperformed the CNN in four classes ('diatoms', 'echinodermata', 'noctiluca', and 'pteropods'), at least in precision and the F1-score. Only the 'marine snow' recall of the CapsNet was superior to the CNN. Both models had difficulties with the class 'rod', which contains unidentified elongated objects. Another common weakness was the 'unknown' class with low recall scores (Figure 7). Most of the images labelled as 'unknown' by a human are recognized as a specific class by both models, mainly as 'marine snow' or 'appendicularia with houses'.

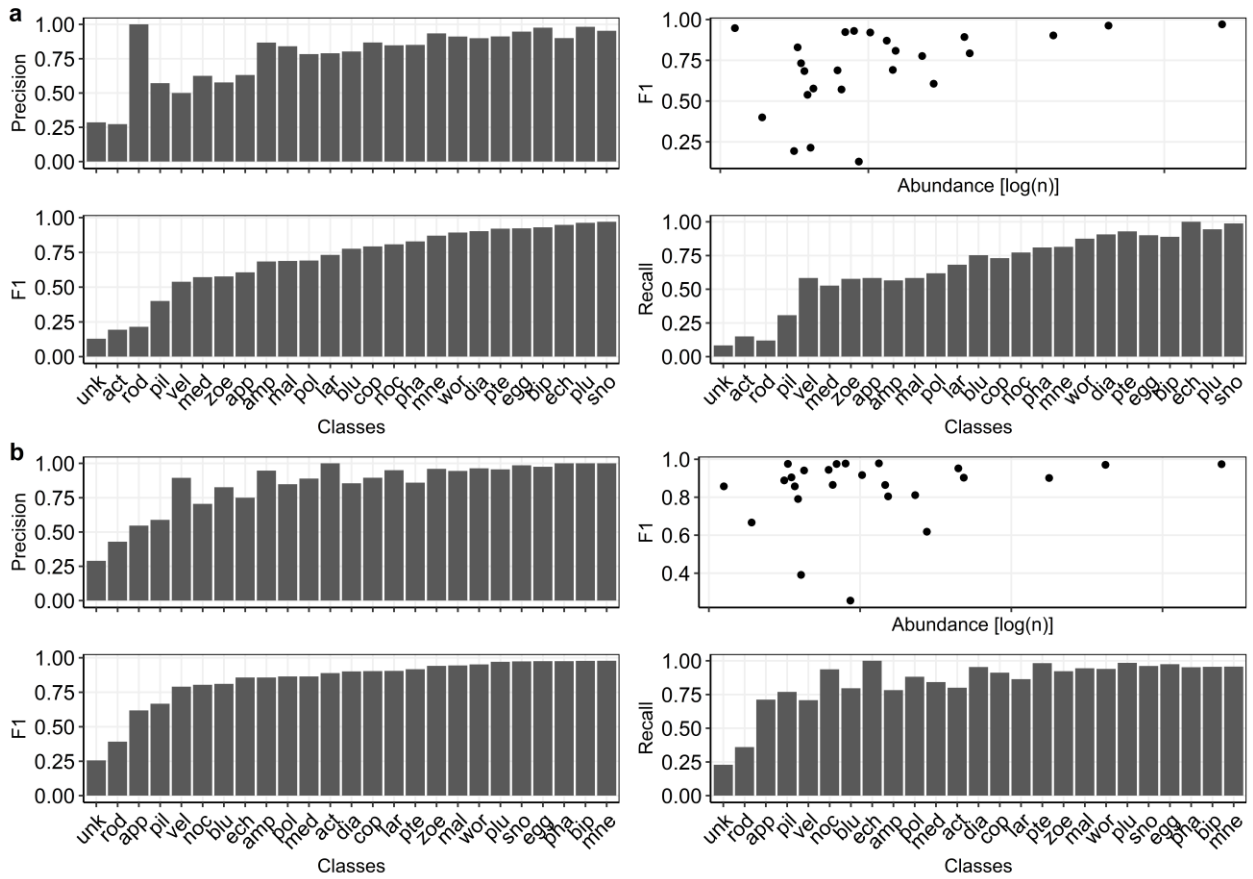


Figure 7: F-metrics (F1, precision, and recall) by class for the CapsNet (a) and the CNN (b). Classes were sorted by increasing F1-score from left to right. Class keys were presented in Table 1. The upper right panel in each plot presents the F1-score in relation to the abundance.

Image filtering

All classes shared a common pattern in regard to the assigned probability filters: at a high threshold, precision was high while recall was low. Thus, only correct classifications were accepted at the cost of discarding most of the correct, less confident classifications together with the wrong classifications. With decreasing probability filters, this was reversed at some point since more and more correctly identified images of the respective class were kept, while simultaneously the chance increased that incorrect classifications were kept as well. As long as the recall was close to 0, the F1-score tended to follow the recall. This was due to the fact, that the harmonic mean (F1) tends to be 0 as soon as one of the components is 0 (recall, Figure 8). We selected class-specific filters aiming to achieve at least 95 % precision in each class. All of the following results were based on these set of filters (P95).

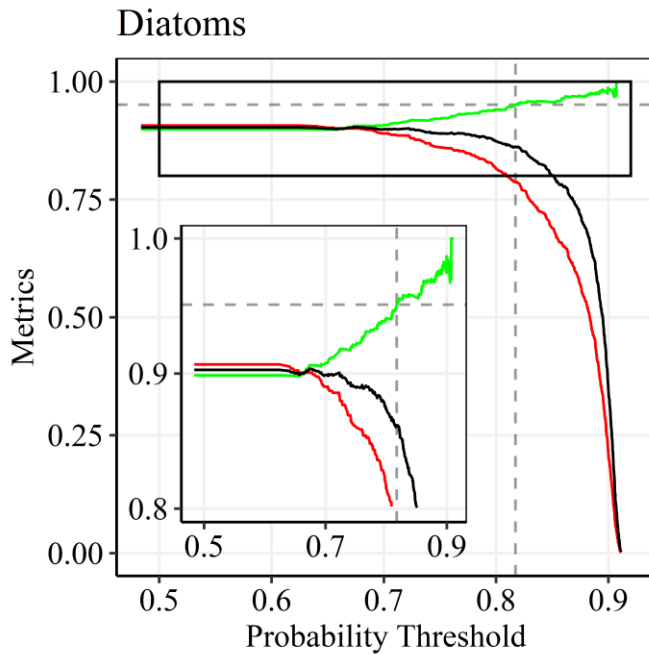


Figure 8: Example plot (class 'diatoms') of a filter selection. Probability filters reflect the confidence of the model in the predicted class. The sub panel is the sub - section from the whole plot where the metrics reach values $\geq 80\%$. Green: precision; red: recall; black: F1-score; grey dashed lines: probability at which precision reaches 95 %.

The final filters varied from 36.44 - 92.36 % (CapsNet) and from 21.17 - 98.65 % (CNN), depending on the class. In general, after filtering the CNN was still superior to the CapsNet. However, for single classes the results of the CapsNet could overcome those of the CNN (Figure 9).

Applying the filters to our test set increased the mean precision of the CNN by 14 % from 84 % to 98 % and of the CapsNet by 15 % from 78 % to only 93 %, as 6 classes did not achieve the target of 95 % precision. Seven percent of all predictions had to be discarded using the CapsNet to maximise precision (5 % for the CNN).

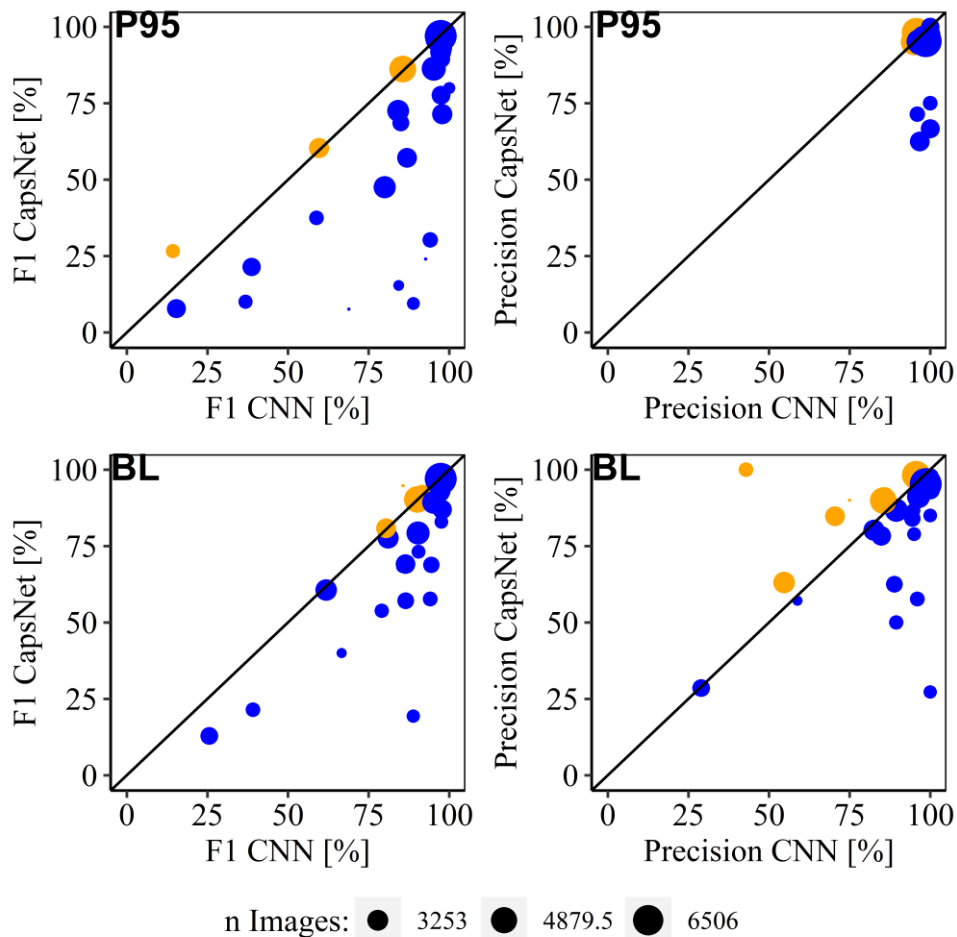


Figure 9: Comparing the F1-score and the precision between the CapsNet and the CNN with (P95) and without (BL) filtering the predictions. The size of each dot represents the logarithms of the abundance of the respective class in the labelled test set. Upper panels: with filtering; lower panels: without filtering; left panels: F1-score; right panels: precision; orange: CapsNet superior to CNN; blue: CNN superior to CapsNet.

Field set predictions

The Bray-Curtis (BC) dissimilarity for the field set predictions, not the manually validated FSs, confirmed as expected, that FS446 ($BC_{\text{CNN}} = 0.45$; $BC_{\text{Cap}} = 0.44$) was closest to the TS as it was sampled in the same geographical region 12 hours after the majority of our training images. FS534 ($BC_{\text{CNN}} = 0.69$; $BC_{\text{Cap}} = 0.67$) was closer to our training data and to FS446 than FS466 ($BC_{\text{CNN}} = 0.92$; $BC_{\text{Cap}} = 0.92$). Thus, DSS is highest for FS466, lowest for FS446 and in-between for FS534 (Supplementary Figure S 1).

With increasing DSS the threshold (t) to separate ‘pure’ from ‘uncertain’ classes for the CapsNet increased according to an exponential function of distance (Figure 10):

$$(5) t = a \times e^{(b \times BC)}$$

with $a = 0.64$ and $b = 5.21$. The simulated thresholds for FS466 followed a bimodal distribution. As the two groups were clearly separated, we chose to include only the higher group of thresholds in the estimation of the model. Therefore, it is less likely for the model to underestimate the true threshold. The observed thresholds ($t_{446} = 5$; $t_{466} = 65$; $t_{534} = 25$) were close to the average simulated thresholds ($t_{446} = 3$; $t_{466} = 65$; $t_{534} = 23$). No reasonable relationship could be established for the CNN and simulated and observed thresholds did not match either.

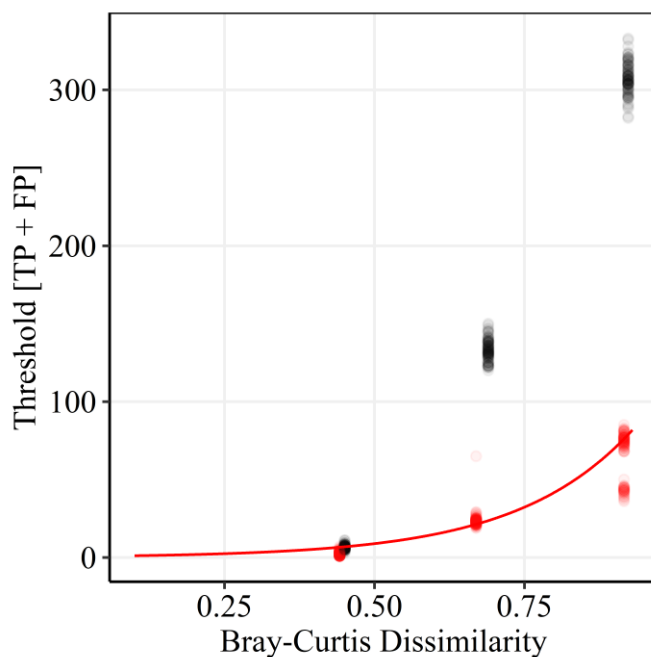


Figure 10: Thresholds to differentiate between ‘pure’ and ‘uncertain’ classes based on the Bray-Curtis dissimilarity between the three field sets and the training set. For each field set, 100 bootstraps were performed using randomly chosen 75 % of the images remaining after P95-filtering. The labels are valid for the black and the red dots alike. Black: CNN; red: CapsNet.

Filtering generally increased the mean precision and reduced the mean recall as expected. Excluding three none-biological classes from the analyses, namely ‘blurry’, ‘unknown’, ‘rod’ and additionally ‘marine snow’, the thresholds between ‘pure’ (mean precision > 90 %) and ‘uncertain’ classes were always ~ 3 times higher for the CNN

compared to the CapsNet, e.g. in FS446 all classes with 5 assigned images by the CapsNet already belonged to the 'pure' group while the CNN had to assign at least 15 images to all classes to reach a mean precision > 90 % in the 'pure' group (Table 3).

Both models successfully detected a similar amount of classes in the field sets, but the CapsNet predictions had more classes contribution to the 'pure' group compared to the CNN. So overall, the CapsNet was better in the generation of 'pure' groups. The CapsNet predicted less images in classes, which were not occupied in the field set ($n_{\text{true}} = 0$ & $n_{\text{pred}} > 0$), thereby creating so-called empty classes with only FPs. In the predictions of the field set least similar to the training set (FS466) neither model achieved a mean precision > 90 %. The selected threshold only maximised the mean precision to 87 % for the CNN (1 'pure' class) and to 76 % for the CapsNet (3 'pure' classes, Table 3).

The recall of the CapsNet was always lower compared to the CNN, but this CNN advantage was reduced by increasing DSS. Neither model dominated the other one regarding the discard ratio, i.e. the number of images that had a lower probability than the class specific filter for certain predictions (P95) compared to uncertain predictions. The discards ranged from 3 % to 45 % for the CNN and from 5 % to 41 % for the CapsNet (Table 3).

For illustrative purposes, we will give an interpretation of the first row in Table 3 (predictions of FS446 by the CNN). The field set included 18 classes that were also present in the training set: prior to P95-filtering, the model predicted 22 classes including 4 empty classes ($TP = 0$). 12 classes included more than 15 predicted images ($t = 15$). Those 12 classes had a mean F1-score of 77 %, a mean precision of 77 % and a mean recall of 80 %. After applying the P95-filter set, 17 classes remained including now only 3 empty classes. The 'pure' group of classes ($TP + FP > 15$) included 8 classes of which none were empty ($TP = 0$). Those had a mean F1-score of 89 %, a mean precision of 94 % and a mean recall of 87 %. Only 3 % of the images belonging to the 18 'true' classes ($TP > 0$) were discarded after filtering.

Table 3: Changes in model performance for biological classes induced by class-wise P95-filtering. The numbers give the actual result after filtering, while the numbers in the brackets give the difference from pre- to post-filtering. Mean F-scores (F1,

precision, and recall) were calculated using only classes with $TP + FP > t$ (group of ‘pure’ classes after filtering). True and empty refer to the respective number of classes predicted by the model for the respective FS. Discard is the percentage of images predicted with insufficient certainty for the P95-filter set.

mod	data	classes		pure classes		t	mean F1	mean Precision	mean Recall	discard [%]
		true	empty	true	empty					
CNN	446	14 (+0)	3 (-1)	8 (-3)	0 (-1)	15	0.89 (+0.12)	0.94 (+0.17)	0.87 (+0.07)	3
CNN	466	16 (+0)	4 (+0)	1 (-2)	0 (+0)	190	0.74 (+0.09)	0.87 (+0.29)	0.64 (-0.16)	33
CNN	534	15 (-1)	5 (+0)	5 (+0)	0 (-2)	85	0.68 (+0.05)	0.98 (+0.35)	0.61 (-0.03)	45
CAP	446	13 (-1)	1 (-1)	8 (-5)	0 (+0)	5	0.59 (-0.13)	0.94 (+0.06)	0.53 (-0.15)	5
CAP	466	16 (+0)	3 (-1)	3 (-1)	0 (-1)	60	0.55 (+0.05)	0.76 (+0.19)	0.5 (-0.02)	41
CAP	534	15 (-1)	3 (-2)	6 (-4)	0 (-2)	25	0.61 (+0.03)	0.91 (+0.28)	0.59 (-0.01)	33

Top-k predictions

We investigated the relationship between k and the mean recall based on the predictions for FS446 (n = 55,302). The recall scores of the CNN always exceeded those of the CapsNet and simultaneously the number of images to validate manually was always lower. We therefore only present the results for the CNN. With k = 2 the mean recall increased from 63 % (Supplementary Table S 1) to 93 % (Table 4), while on average 7.8 % of the images had to be validated. Only three classes required manual classification of more than 10 % of the original data set images (‘diatoms’, ‘pluteus’, and ‘snow’), but those were the most abundant classes. The majority (12 classes) required less than 3 % of the original data set to be manually classified. With k = 3, the increase in recall (+ 2.9 %) was similar to the increase in images (+ 3.3 %), but further increasing k was less effective. We therefore selected k = 3 for all field sets.

Table 4: The development of the mean recall and the mean percentage of images to validate with increasing k for FS446, where k was the number of most likely predictions accepted for each image. If the true class belonged to the k accepted predictions, an image was counted as ‘true positive’. N images: 55,302.

N k = 2 [%]	recall k = 2	N k = 3 [%]	recall k = 3	N k = 4 [%]	recall k = 4	N k = 5 [%]	recall k = 5
7.8	93.4	11.1 (+3.3)	96.3 (+2.9)	14.5 (+3.5)	97.2 (+0.9)	18.6 (+4.1)	97.6 (+0.4)

The classes ‘blurry’ and ‘unknown’ included per definition a wide range of different, unidentified objects which made them scientifically uninteresting. They were therefore not included in the analyses. We also excluded the class ‘marine snow’ because of the extraordinary size. For the remaining classes the mean recall with k = 3 exceeded 90 % for the low- (FS446; 96 %, n = 17,318) and medium-shift data sets (FS534; 95 %, n = 9,557), while the mean recall for the high-shift data set (FS466; n = 2,731) was only 86 %. The supplementary material includes a complete table with all classes and field sets (Supplementary Table S 2), here we described only the results for the first field set (FS446) in detail.

Only two classes had a recall below 90 % (‘malacostraca’ and ‘medusae’). All other classes, even the rarest, had a recall above 95 %. For especially rare classes like ‘eggs’ (n = 5) and ‘larvae’ (n = 15) less than 1 % of the original data set needed manual validation to achieve a recall score of 100 % at k = 3. However, rare classes usually had lower ratios between TP’s and FP’s compared to more abundant classes. An exception from this trend was the class ‘polychaeta’ with n = 363 and a ratio between TP and FP of 1 : 29. Thus, ~ 20 % of the original data set needed to be manually validated in order to achieve 97.5 % recall for this class (Table 5).

Table 5: Results of the Top-k-method with the CNN and $k = 3$ applied to FS446. The variable 'n' is the number of images that need to be manually validated to maximize the recall. The variable 'N' is the number of images in FS446 (55,302) including the three classes 'unknown', 'blurry', and 'marine snow'. When recall is empty, the model sorted images in a class which was not occupied in the field data set ($TP = 0$). The ratio 'TP:FP' provides the number of false images to be manually sorted for one true image found.

class	recall	CC	n	n / N [%]	CC:FC
amp		0	4824	8.7	
app	0.997	298	6331	11.4	1:20.2
bip	0.987	79	171	0.3	1:1.2
cop	0.98	151	3601	6.5	1:22.8
dia	1	3190	11172	20.2	1:2.5
ech		0	321	0.6	
egg	1	5	346	0.6	1:68.2
lar	1	15	85	0.2	1:4.7
mal	0.773	22	122	0.2	1:4.5
med	0.789	76	4353	7.9	1:56.3
mne	0.986	144	184	0.3	1:0.3
noc	0.971	348	5197	9.4	1:13.9
plu	0.998	9861	11675	21.1	1:0.2
pol	0.975	363	10896	19.7	1:29
pte		0	81	0.1	
rod	0.954	2034	5684	10.3	1:1.8
vel		0	90	0.2	
wor	0.991	705	2731	4.9	1:2.9
zoe	1	27	77	0.1	1:1.9
Mean:	0.96 ± NA	1155	3576	6.46	1:15

class	recall	CC	n	n / N [%]	CC:FC
Median:	0.987	79	2731	4.9	1:13.9

Spatial distributions

We calculated Dutilleuls modified t-test to assess whether our P95 filtered model predictions were representative for the true plankton community in our field sets. While $p < 0.05$ was sufficient to accept the representativeness of a class prediction, we generally assumed the model with the lower p -value to be superior. While the CapsNet was superior to the CNN in 11.1 % of all classes in FS446 (low DSS), this increased to 21.1 % in FS534 (medium DSS). However, the CNN was superior to the CapsNet in 50 % of all classes in FS446 and in 36.8 % classes in FS534. While this gave hope for a trend reversal in high DSS situations, in the high-shift field set FS466 the CNN is still superior in 55 % of all classes and the CapsNet is only superior in 10 % of the classes. Spatial distributions predicted by both models did not show any significant deviations from those of manually validated images, when two conditions were met: $n_{\text{true}} > 50$ images and recall > 20 %, regardless of the level of DSS (Supplementary Table S 3). Figure 11 shows two exemplary distributions of copepods predicted by our models in the field sets FS446 and FS534, demonstrating the difficulties the CapsNet had with low abundant classes (Supplementary Table S 3 a).

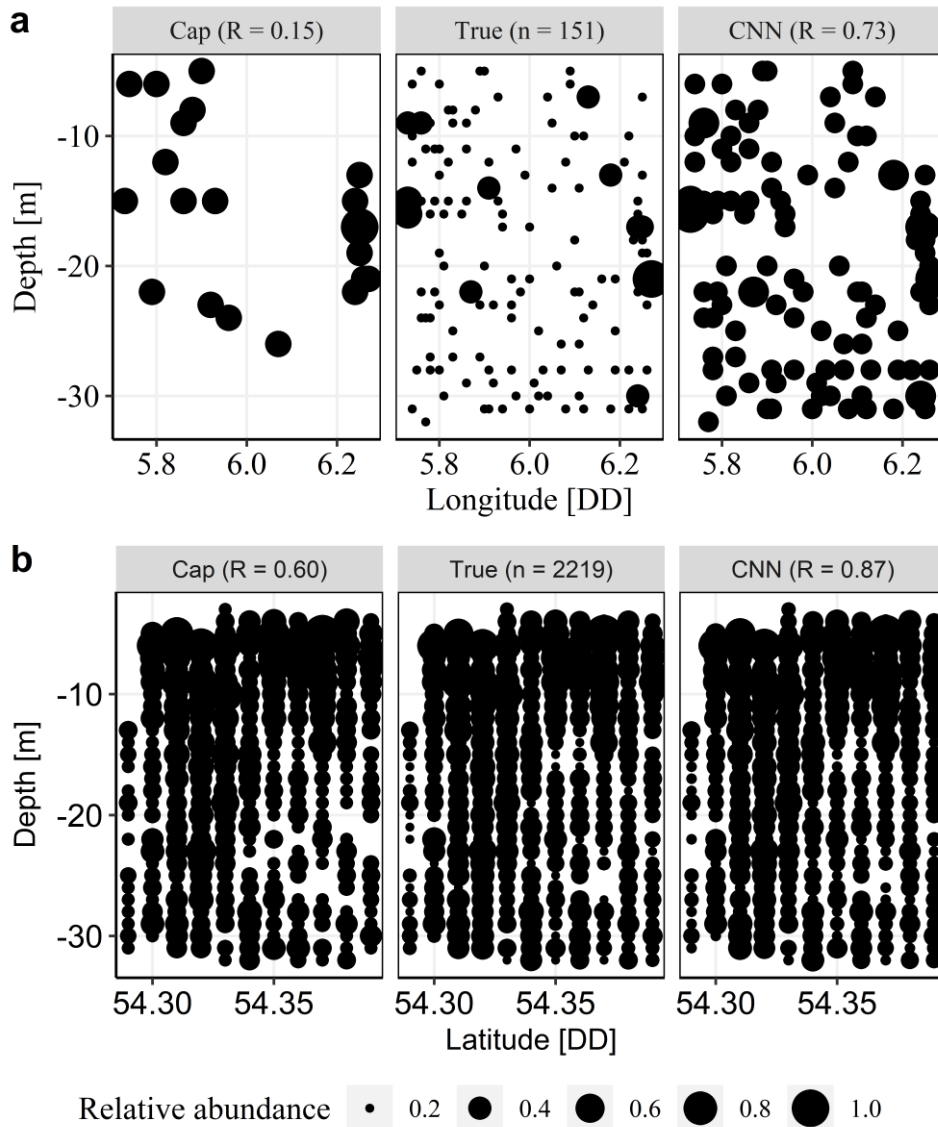


Figure 11: Spatial distribution of the relative copepod abundances in FS446 (a) and FS534 (b) aggregated by Longitude/ Latitude (0.01 DD) and depth (m). The predictions of the CapsNet (left panel) and the CNN (right panel) were compared to the manual validation (central panel). *R*: recall; *n*: absolute abundance.

Discussion

The intention of this study was to provide a guideline to efficiently process not only common, but also rare biotic taxa using automated analyses of in situ plankton images, which is even more challenging than laboratory imagery of plankton according to Faillettaz et al. (2016).

During the first step of the model training we observed a great variability of the final test accuracies, even though the procedure was exactly the same each time, except for the sequence of the images. All classes were trained with the same number of images during step 1, but some classes had persistently lower mean test accuracies than others. However, the lower the mean test accuracy, the greater were the CI's. For those classes, e.g. 'polychaeta', 100 images were clearly not sufficient to reflect the class variability of the full data set. Some of the 100 models probably learned more relevant patterns, most likely by chance (González et al., 2017). This highlights one of the drawbacks of gradient-based algorithms as described by Subbey (2018) and the importance of a vast amount of training images, especially if different classes contain similar organisms (like 'amphipods' and 'copepods') and additionally one or more similar classes have high intra-class variability (e.g. already due to frontal, dorsal or lateral viewpoints). E.g., images of 'veliger' were frequently misinterpreted as 'pilidium' and in case of the CapsNet even vice versa (Supplementary Table S 4).

For ecological studies it is sometimes more important that an image is correctly classified into a certain group rather than the exact class (González et al., 2017). E.g., one way to cope with the high intraspecific variability in plankton classes is to divide images of a single species into multiple classes according to morphological distinctions (Luo et al., 2018). In this study, initial experiments with our classifier showed that the separation of the images with appendicularians in 'appendicularia' and 'appendicularia with house' yielded much better results compared to the classification of a single, combined class. Since both classes were treated as one in the subsequent analyses, a misclassification of an 'appendicularia' into 'appendicularia with house' was ultimately a correct classification, thus increasing the performance of our model in a way of a Top-2 accuracy. This method is probably even more relevant for more detailed image sets that allow for a higher taxonomic resolution than our images (e.g. flowcam images), but is somewhat limited by the number of available training images. Within this context, different sizes of plankton can have different biological meaning, like ontogenetic stages, and thus could be worth including. However, the true size information is unfortunately not available from VPR-images as the distance to the lens is unknown.

Filtering

Our approach reduced the discard of uncertain predicted images, which are removed by filtering, from 35.7 % in Luo et al. (2018) to 5 %. In Luo et al. (2018), classes with $n < 25$ of 75,000 randomly drawn images were excluded to achieve a mean precision of 90.7 %. This threshold ($n = 25$) divided classes into a 'pure' (precision > 90 %) and an 'uncertain' (precision < 90 %) group. However, an evaluation using n is probably misleading since some classes in the field set had a true $n = 0$, which was below the threshold of $n < 5$, and simultaneously had $FP > 5$, which was above that threshold. Thus, without human interference (i.e. validation), classes with $n < 5$ could be erroneously categorized as 'pure' (and vice versa). Therefore, we used the sum of TP's and FP's instead to divide between 'pure' (i.e. trustworthy) and 'uncertain' classes (i.e. classes that need to be validated). The major advantage of this approach is that this threshold is applicable without knowledge of the true distribution of the classes since it is based on the predictions instead of the true abundance. Furthermore, we found a correlation between the threshold and the Bray-Curtis dissimilarity that separates the distribution of the TS from the distribution of the new field set. Remarkably, this new method increased the threshold from $n < 5$ to $TP + FP < 15$ for the CNN while it decreased the threshold for the CapsNet ($n < 25$; $TP + FP < 5$). This threshold to separate 'pure' from 'uncertain' classes using the CapsNet was below the threshold of the CNN and therefore the CapsNet is superior in extended production mode applications. Summarizing, the CapsNet had similar discard ratios but lower mean recall scores compared to the CNN. Thus, while it produced more 'pure' classes, the drawback was a stronger filter pressure on rare classes.

However, each optical sampler is designed to target a different component of the zooplankton (Owens et al., 2013) which encompasses organisms that vary greatly in terms of size, shape and behavior (Pitois et al., 2018). This and varying environmental conditions and ecosystem compositions may affect the difficulty of classification tasks (Luo et al., 2018) and might contribute to the differences found between different studies.

The fraction of 'pure' classes decreased for the CapsNet from 57 % in FS446 (CNN: 57 %) with a low amount of DSS to 38 % (CNN: 31 %) in FS534 with the medium amount of DSS. The high amount of DSS in FS466 overcharged both models

similarly, as no class actually reached > 90 % precision. This trend observed between the threshold 't' and DSS (measured as Bray-Curtis dissimilarity) was not reflected in the F-metrics. No obvious pattern for biological classes was observed regarding the F-score, recall, or precision, indicating that the community structure, and thus the difficulty of the classification task, is probably an evenly strong driver of model performance as DSS. Considering all classes, the CapsNet could not achieve as high scores as the CNN under a low amount of DSS, but the decrease with increasing DSS was also lower. However, for the CapsNet to overcome the CNN an amount of DSS is probably necessary that precludes a practical application of either model.

The performance of regionally trained classifiers tends to decline with increasing environmental dissimilarity, whereas a globally trained classifier achieved similar results in all areas, but at the cost of lower accuracies for rare taxa (Chang et al., 2012). Continually increasing the training set and adapting the model to new situations could therefore help coping with unknown community structures and keeping the amount of DSS at a lower level, but this was not investigated in this study.

Spatial distributions

Both models accurately predicted the spatial distributions of filtered classes with $n_{\text{true}} > 50$ and recall > 20 %, regardless of the amount of DSS. Due to better recall scores, the CNN could predict the spatial distributions of more classes compared to the CapsNet, especially more smaller classes. Most classes which spatial distributions were correctly predicted belonged to the 'pure' group. However, the opposite was not given. Due to low recall scores not all predictions for classes from the 'pure' group reflected the spatial distribution in the field, as some 'pure' classes had a high precision but a low recall (e.g. CapsNet FS446 'appendicularia': precision = 100 %, recall < 1 %). Thus, a categorization in trustworthy and misleading predicted spatial distributions generally requires knowledge of the recall and therefore manual validation.

Top-k predictions

The application of filters enables a user to automatically detect a wide range of taxa with a high precision. However, less abundant classes are still difficult to predict,

especially if DSS occurs. Since filters are per definition not appropriate to increase the recall and make it more likely to detect rare classes, we instead employed the concept of the Top-5-Accuracy used in machine learning. Here, the correct class of an image doesn't have to have the highest probability. Instead, the correctly predicted class is among the five highest probabilities. Using already the Top-3-Accuracy, we increased the mean recall by 33.3 % to 96.3 % and reduced the average number of images that needed manual validation to 6.5 % of the original amount. This method significantly reduces the required human efforts if the research focuses on rare classes, like fish larvae, that are most likely not detected at a sufficient rate using only the highest probability. Such an approach is certainly limited by the size of a potential data set and the number of classes, but so far it is probably the most effective way if spatial distributions are equally important as total abundances.

Comments and recommendations

The effectivity of a classifier is not solely determined by the final model performance. Particularly, the specific objectives of a research task need to be considered to tailor the best model. Research targeting rare classes usually requires a quantitative classification (i.e. high recall) rather than a qualitative classification (i.e. high precision) which is more important for a study of community structures and biodiversity. The assessment of spatial distributions requires qualitative and quantitative classifications which is, without manual validation, currently limited to dominant classes (> 1 % of the whole data set). As a general guideline we propose the following scheme (Figure 12): either model can be used to classify a given data set using the P95-filter set. Subsequently, a cluster analyses to estimate the similarity of the new data set (predictions) and the training data set (validated) is needed. As we have not investigated how the thresholds behave for increasing data set sizes, currently a subset of the original data set of 30,000 - 50,000 images is recommended to estimate the threshold for 'pure' classes based on Eqn. 5. Recall scores of the CapsNet fluctuate less strong with varying levels of DSS and the threshold 't' can be adapted dynamically, which is important for the comparison of different samples. The spatial distributions of a particular sample should be investigated using a CNN due to better recall scores in general, but without manual validation this is limited to the

dominant taxa. Rare taxa should be targeted using a CNN and the Top-3-Accuracy to maximise the recall at a limited amount of human effort. Thus, the only benefit in using a CapsNet in fact arises under the presumption of DSS.

Predicted data set - TP + FP for each class (no validation necessary!)		
Target: single class	Target: multiple classes	Target: spatial distributions
CNN + Top-3-Accuracy	use SUBSET for cluster analyses and select threshold	threshold: classes with abundances > 1 %
→ high recall → low effort	CapsNet + Filter → preferably presence/ absence	CNN + Filter → limited to dominant taxa

Figure 12: Suggested workflows for three research target specific automated plankton image analysis methods.

Data availability statement

All manually classified images from the full training set and test sets (124 K probability filtering set and 94 K random field sets), as well as text files containing predicted and validated classes for all test sets will be available on Zenodo.org (doi: 10. 5281/zenodo.4431509).

Acknowledgments

Many thanks to Ángel López-Urrutia for his insightful comments and helpful discussions. We thank Mathis Rosenhauer for the introduction in how to operate the high performance computing system 'Mistral' of the German Climate Computing Center and Inga Hense for sharing valuable computing hours of her project. The three RV Heincke research cruises were supported by grant numbers AWI_HE446_00, AWI_HE466_00, and AWI_HE534_00. Finally, we would like to thank an anonymous reviewer whose comments helped to improve our manuscript.

References

- Abadi, M., Agarwal, A., Barham, P., et al. 2015 *TensorFlow: Large-scale machine learning on heterogeneous distributed systems*. Available at: <https://www.tensorflow.org/>.
- Afshar, P., Mohammadi, A. and Plataniotis, K. N. 2018 'Brain tumor type classification via capsule networks', *arXiv:1802.10200 [cs]*. Available at: <http://arxiv.org/abs/1802.10200> (Accessed: 2 April 2019).
- Al-Barazanchi, H., Verma, A. and Wang, S. X. 2016 'Intelligent plankton image classification with deep learning', p. 11.
- Batten, S. D., Abu-Alhajja, R., Chiba, S., et al. 2019 'A global plankton diversity monitoring program', *Frontiers in Marine Science*, 6. doi: [10.3389/fmars.2019.00321](https://doi.org/10.3389/fmars.2019.00321).
- Bell, J. L. and Hopcroft, R. R. 2008 'Assessment of ZoolImage as a tool for the classification of zooplankton', *Journal of Plankton Research*, 30(12), pp. 1351–1367. doi: [10.1093/plankt/fbn092](https://doi.org/10.1093/plankt/fbn092).
- Benfield, M., Grosjean, P., Culverhouse, P., et al. 2007 'RAPID: Research on automated plankton identification', *Oceanography*, 20(2), pp. 172–187. doi: [10.5670/oceanog.2007.63](https://doi.org/10.5670/oceanog.2007.63).
- Bishop, C. M. 1995 *Neural networks for pattern recognition*. Oxford: Oxford University Press.
- Bray, J. R. and Curtis, J. T. 1957 'An ordination of the upland forest communities of southern wisconsin', *Ecological monographs*, 27(4), pp. 326–349.
- Briseño-Avena, C., Schmid, M. S., Swieca, K., Sponaugle, S., Brodeur, R. D. and Cowen, R. K. 2020 'Three-dimensional cross-shelf zooplankton distributions off the central oregon coast during anomalous oceanographic conditions', *Progress in Oceanography*, 188, p. 102436. doi: [10.1016/j.pocean.2020.102436](https://doi.org/10.1016/j.pocean.2020.102436).
- Chang, C.-Y., Ho, P.-C., Sastri, A. R., Lee, Y.-C., Gong, G.-C. and Hsieh, C.-h. 2012 'Methods of training set construction: Towards improving performance for automated mesozooplankton image classification systems', *Continental Shelf Research*, 36, pp. 19–28. doi: [10.1016/j.csr.2012.01.005](https://doi.org/10.1016/j.csr.2012.01.005).
- Chollet, F. 2015 *Keras*. GitHub. Available at: <https://github.com/fchollet/keras>.
- Chollet, F. 2017 *Deep learning with python*. 1st edn. Shelter Island, New York: Manning Publications.
- Culverhouse, P., Williams, R., Reguera, B., Herry, V. and González-Gil, S. 2003 'Do experts make mistakes? A comparison of human and machine identification of dinoflagellates', *Marine Ecology Progress Series*, 247, pp. 17–25. doi: [10.3354/meps247017](https://doi.org/10.3354/meps247017).
- Davis, C. S., Gallager, S. M., Bermann, M. S., Haury, L. R. and Strickler, J. R. 1992 'The video plankton recorder (VPR): Design and initial results', *Archiv für Hydrobiologie–Beiheft Ergebnisse der Limnologie*, 36, pp. 67–81. Available at: https://www.researchgate.net/publication/284686405_The_Video_Plankton_Recorder_VPR_Design_and_initial_results (Accessed: 5 February 2020).
- Davis, C. S. and McGillicuddy, D. J. 2006 'Transatlantic abundance of the n₂-fixing colonial cyanobacterium trichodesmium', *Science*, 312(5779), pp. 1517–1520. doi: [10.1126/science.1123570](https://doi.org/10.1126/science.1123570).

- Dutilleul, P., Clifford, P., Richardson, S. and Hemon, D. 1993 'Modifying the t test for assessing the correlation between two spatial processes', *Biometrics*, 49(1), pp. 305–314. doi: [10.2307/2532625](https://doi.org/10.2307/2532625).
- Faillietaz, R., Picheral, M., Luo, J. Y., Guigand, C., Cowen, R. K. and Irisson, J.-O. 2016 'Imperfect automatic image classification successfully describes plankton distribution patterns', *Methods in Oceanography*, 15-16, pp. 60–77. doi: [10.1016/j.mio.2016.04.003](https://doi.org/10.1016/j.mio.2016.04.003).
- Floeter, J., Beusekom, J. E. E. van, Auch, D., et al. 2017 'Pelagic effects of offshore wind farm foundations in the stratified north sea', *Progress in Oceanography*, 156, pp. 154–173. doi: [10.1016/j.pocean.2017.07.003](https://doi.org/10.1016/j.pocean.2017.07.003).
- González, P., Álvarez, E., Díez, J., López-Urrutia, Á. and Coz, J. J. del 2017 'Validation methods for plankton image classification systems', *Limnology and Oceanography: Methods*, 15(3), pp. 221–237. doi: [10.1002/lom3.10151](https://doi.org/10.1002/lom3.10151).
- Hansson, S., Larsson, U. and Johansson, S. 1990 'Selective predation by herring and mysids, and zooplankton community structure in a baltic sea coastal area', *Journal of Plankton Research*, 12(5), pp. 1099–1116. doi: [10.1093/plankt/12.5.1099](https://doi.org/10.1093/plankt/12.5.1099).
- He, H. and Garcia, E. A. 2009 'Learning from imbalanced data', *IEEE Transactions on Knowledge and Data Engineering*, 21(9), pp. 1263–1284. doi: [10.1109/TKDE.2008.239](https://doi.org/10.1109/TKDE.2008.239).
- He, K., Zhang, X., Ren, S. and Sun, J. 2015 'Deep residual learning for image recognition', *arXiv:1512.03385 [cs]*. Available at: <http://arxiv.org/abs/1512.03385> (Accessed: 26 March 2020).
- Hinton, G. E., Krizhevsky, A. and Wang, S. D. 2011 'Transforming auto-encoders', in Honkela, T., Duch, W., Girolami, M., and Kaski, S. (eds) *Artificial neural networks and machine learning – ICANN 2011*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 44–51. doi: [10.1007/978-3-642-21735-7_6](https://doi.org/10.1007/978-3-642-21735-7_6).
- Hu, Q. and Davis, C. 2005 'Automatic plankton image recognition with co-occurrence matrices and support vector machine', *Marine Ecology Progress Series*, 295, pp. 21–31. doi: [10.3354/meps295021](https://doi.org/10.3354/meps295021).
- Hu, Q. and Davis, C. 2006 'Accurate automatic quantification of taxa-specific plankton abundance using dual classification with correction'. doi: [10.3354/MEPS306051](https://doi.org/10.3354/MEPS306051).
- Jiménez-Sánchez, A., Albarqouni, S. and Mateus, D. 2018 'Capsule networks against medical imaging data challenges', *arXiv:1807.07559 [cs]*, 11043, pp. 150–160. doi: [10.1007/978-3-030-01364-6_17](https://doi.org/10.1007/978-3-030-01364-6_17).
- Johansson, M., Gorokhova, E. and Larsson, U. 2004 'Annual variability in ciliate community structure, potential prey and predators in the open northern baltic sea proper', *Journal of Plankton Research*, 26(1), pp. 67–80. doi: [10.1093/plankt/fbg115](https://doi.org/10.1093/plankt/fbg115).
- Kingma, D. P. and Ba, J. 2014 'Adam: A method for stochastic optimization', *arXiv:1412.6980 [cs]*. Available at: <http://arxiv.org/abs/1412.6980> (Accessed: 11 April 2019).
- Kornblith, S., Shlens, J. and Le, Q. V. 2019 'Do better ImageNet models transfer better?', in *2019 IEEE/CVF conference on computer vision and pattern recognition (CVPR)*. *Conference on computer vision and pattern recognition (CVPR)*, pp. 2656–2666. doi: [10.1109/CVPR.2019.00277](https://doi.org/10.1109/CVPR.2019.00277).

- Krizhevsky, A., Sutskever, I. and Hinton, G. E. 2017 'ImageNet classification with deep convolutional neural networks', *Communications of the ACM*, 60(6), pp. 84–90. doi: [10.1145/3065386](https://doi.org/10.1145/3065386).
- LeCun, Y., Bengio, Y. and Hinton, G. 2015 'Deep learning', *Nature*, 521(7553), pp. 436–444. doi: [10.1038/nature14539](https://doi.org/10.1038/nature14539).
- Lee, H., Park, M. and Kim, J. 2016 'Plankton classification on imbalanced large scale database via convolutional neural networks with transfer learning', in *2016 IEEE international conference on image processing (ICIP). 2016 IEEE international conference on image processing (ICIP)*, pp. 3713–3717. doi: [10.1109/ICIP.2016.7533053](https://doi.org/10.1109/ICIP.2016.7533053).
- Luo, J. Y., Irisson, J.-O., Graham, B., Guigand, C., Sarafraz, A., Mader, C. and Cowen, R. K. 2018 'Automated plankton image analysis using convolutional neural networks', *Limnology and Oceanography: Methods*, 16(12), pp. 814–827. doi: [10.1002/lom3.10285](https://doi.org/10.1002/lom3.10285).
- Moreno-Torres, J. G., Raeder, T., Alaiz-Rodríguez, R., Chawla, N. V. and Herrera, F. 2012 'A unifying view on dataset shift in classification', *Pattern Recognition*, 45(1), pp. 521–530. doi: [10.1016/j.patcog.2011.06.019](https://doi.org/10.1016/j.patcog.2011.06.019).
- Möller, K. O., Schmidt, J. O., St.John, M., et al. 2015 'Effects of climate-induced habitat changes on a key zooplankton species', *Journal of Plankton Research*, 37(3), pp. 530–541. doi: [10.1093/plankt/fbv033](https://doi.org/10.1093/plankt/fbv033).
- Möller, K. O., St John, M., Temming, A., Floeter, J., Sell, A., Hermann, J. and Möllmann, C. 2012 'Marine snow, zooplankton and thin layers: Indications of a trophic link from small-scale sampling with the video plankton recorder', *Marine Ecology Progress Series*, 468, pp. 57–69. doi: [10.3354/meps09984](https://doi.org/10.3354/meps09984).
- Nanni, L., Brahmam, S., Ghidoni, S. and Lumini, A. 2019 'Bioimage classification with handcrafted and learned features', *IEEE Computer Society Press*. Available at: <https://doi.org/10.1109/TCBB.2018.2821127> (Accessed: 26 March 2020).
- Ohman, M. D., Davis, R. E., Sherman, J. T., Grindley, K. R., Whitmore, B. M., Nickels, C. F. and Ellen, J. S. 2019 'ZooGLIDER: An autonomous vehicle for optical and acoustic sensing of zooplankton', *Limnology and Oceanography: Methods*, 17(1), pp. 69–86. doi: [10.1002/lom3.10301](https://doi.org/10.1002/lom3.10301).
- Oksanen, J., Blanchet, F. G., Friendly, M., et al. 2019 *Vegan: Community ecology package*. Available at: <https://CRAN.R-project.org/package=vegan>.
- Orenstein, E. C., Beijbom, O., Peacock, E. E. and Sosik, H. M. 2015 'WHOI-plankton- a large scale fine grained visual recognition benchmark dataset for plankton classification', *arXiv:1510.00745 [cs]*. Available at: <http://arxiv.org/abs/1510.00745> (Accessed: 25 April 2019).
- Osorio, F. and Vallejos, R. 2019 *Tools for assessment the association between two spatial processes*. Available at: <http://spatialpack.mat.utfsm.cl>.
- Owens, N. J. P., Hosie, G. W., Batten, S. D., Edwards, M., Johns, D. G. and Beaugrand, G. 2013 'All plankton sampling systems underestimate abundance: Response to "continuous plankton recorder underestimates zooplankton abundance" by j.w. Dippner and m. Krause', *Journal of Marine Systems*, 128, pp. 240–242. doi: [10.1016/j.jmarsys.2013.05.003](https://doi.org/10.1016/j.jmarsys.2013.05.003).

- Pan, S. J. and Yang, Q. 2010 'A survey on transfer learning', *IEEE Transactions on Knowledge and Data Engineering*, 22(10), pp. 1345–1359. doi: [10.1109/TKDE.2009.191](https://doi.org/10.1109/TKDE.2009.191).
- Peters, J., Dutz, J. and Hagen, W. 2013 'Trophodynamics and life-cycle strategies of the copepods temora longicornis and acartia longiremis in the central baltic sea', *Journal of Plankton Research*, 35(3), pp. 595–609. doi: [10.1093/plankt/fbt004](https://doi.org/10.1093/plankt/fbt004).
- Pitois, S. G., Tilbury, J., Bouch, P., Close, H., Barnett, S. and Culverhouse, P. F. 2018 'Comparison of a cost-effective integrated plankton sampling and imaging instrument with traditional systems for mesozooplankton sampling in the celtic sea', *Frontiers in Marine Science*, 5. doi: [10.3389/fmars.2018.00005](https://doi.org/10.3389/fmars.2018.00005).
- Rajasegaran, J., Jayasundara, V., Jayasekara, S., Jayasekara, H., Seneviratne, S. and Rodrigo, R. 2019 'DeepCaps: Going deeper with capsule networks', *arXiv:1904.09546 [cs]*. Available at: <http://arxiv.org/abs/1904.09546> (Accessed: 29 January 2020).
- RCoreTeam 2020 *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Available at: <https://www.R-project.org/>.
- Renz, J. and Hirche, H.-J. 2006 'Life cycle of pseudocalanus acuspes giesbrecht (copepoda, calanoida) in the central baltic sea: I. Seasonal and spatial distribution', *Marine Biology*, 148(3), pp. 567–580. doi: [10.1007/s00227-005-0103-5](https://doi.org/10.1007/s00227-005-0103-5).
- Russakovsky, O., Deng, J., Su, H., et al. 2015 'ImageNet large scale visual recognition challenge', *International Journal of Computer Vision*, 115(3), pp. 211–252. doi: [10.1007/s11263-015-0816-y](https://doi.org/10.1007/s11263-015-0816-y).
- Sabour, S., Frosst, N. and Hinton, G. E. 2017 'Dynamic routing between capsules', *arXiv:1710.09829 [cs]*. Available at: <http://arxiv.org/abs/1710.09829> (Accessed: 1 April 2019).
- Schmid, M. S., Cowen, R. K., Robinson, K., Luo, J. Y., Briseño-Avena, C. and Sponaugle, S. 2020 'Prey and predator overlap at the edge of a mesoscale eddy: Fine-scale, in-situ distributions to inform our understanding of oceanographic processes', *Scientific Reports*, 10(1), p. 921. doi: [10.1038/s41598-020-57879-x](https://doi.org/10.1038/s41598-020-57879-x).
- Sosik, H. M. and Olson, R. J. 2007 'Automated taxonomic classification of phytoplankton sampled with imaging-in-flow cytometry: Phytoplankton image classification', *Limnology and Oceanography: Methods*, 5(6), pp. 204–216. doi: [10.4319/lom.2007.5.204](https://doi.org/10.4319/lom.2007.5.204).
- Ston, J., Kosakowska, A., Lotocka, M. and Lysiak-Pastuszak, E. 2002 'Pigment composition in relation to phytoplankton community structure and nutrient content in the baltic sea', *Oceanologia*, 44(4). Available at: <http://yadda.icm.edu.pl/yadda/element/bwmeta1.element.agro-article-190160a5-a0a9-4092-8f28-dda8582dec6f> (Accessed: 30 March 2020).
- Subbey, S. 2018 'Parameter estimation in stock assessment modelling: Caveats with gradient-based algorithms', *ICES Journal of Marine Science*, 75(4), pp. 1511–1511. doi: [10.1093/icesjms/fsy060](https://doi.org/10.1093/icesjms/fsy060).
- Tang, X. and Stewart, W. K. 1996 'Plankton image classification using novel parallel-training learning vector quantization network', in *OCEANS 96 MTS/IEEE conference proceedings. The coastal ocean -*

- prospects for the 21st century. OCEANS 96 MTS/IEEE conference proceedings. The coastal ocean - prospects for the 21st century*, pp. 1227–1236 vol.3. doi: [10.1109/OCEANS.1996.569077](https://doi.org/10.1109/OCEANS.1996.569077).
- Toraman, S., Alakus, T. B. and Turkoglu, I. 2020 'Convolutional capsnet: A novel artificial neural network approach to detect COVID-19 disease from x-ray images using capsule networks', *Chaos, Solitons & Fractals*, 140, p. 110122. doi: [10.1016/j.chaos.2020.110122](https://doi.org/10.1016/j.chaos.2020.110122).
- Van Rossum, G. and Drake, F. L. 2009 *Python 3 reference manual*. Scotts Valley, CA: CreateSpace.
- Vuorio, K., Lagus, A., Lehtimäki, J. M., Suomela, J. and Helminen, H. 2005 'Phytoplankton community responses to nutrient and iron enrichment under different nitrogen to phosphorus ratios in the northern baltic sea', *Journal of Experimental Marine Biology and Ecology*, 322(1), pp. 39–52. doi: [10.1016/j.jembe.2005.02.006](https://doi.org/10.1016/j.jembe.2005.02.006).
- Webb, G. I., Lee, L. K., Goethals, B. and Petitjean, F. 2018 'Analyzing concept drift and shift from sample data', *Data Mining and Knowledge Discovery*, 32(5), pp. 1179–1199. doi: [10.1007/s10618-018-0554-1](https://doi.org/10.1007/s10618-018-0554-1).
- Wickham, H. 2016 *Ggplot2: Elegant graphics for data analysis*. New York: Springer-Verlag. Available at: <https://ggplot2.tidyverse.org>.
- Wickham, H., François, R., Henry, L. and Müller, K. 2020 *Dplyr: A grammar of data manipulation*. Available at: <https://CRAN.R-project.org/package=dplyr>.
- Wiebe, P. H. and Benfield, M. C. 2003 'From the hensen net toward four-dimensional biological oceanography', *Progress in Oceanography*, 56(1), pp. 7–136. doi: [10.1016/S0079-6611\(02\)00140-4](https://doi.org/10.1016/S0079-6611(02)00140-4).
- Xi, E., Bing, S. and Jin, Y. 2017 'Capsule network performance on complex data', *arXiv:1712.03480 [cs, stat]*. Available at: <http://arxiv.org/abs/1712.03480> (Accessed: 8 April 2019).

Supplementary material

Supplementary Table S 1: Changes in model performance induced by class-wise P95-filtering. The numbers give the actual result after filtering, while the numbers in the brackets give the difference from pre- to post-filtering. Mean F-scores (F1, precision, and recall) were calculated using only classes with $TP + FP > t$ (group of 'pure' classes after filtering). True and empty refer to the respective number of classes predicted by the model for the respective FS. Discard is the percentage of images predicted with insufficient certainty for the P95-filter set.

mod	data	classes		pure classes		thres	mean F1	mean Precision	mean Recall	discard [%]	
		true	empty	true	empty						
CNN	446	18 (+0)	3 (-1)	-	-	0	0.53 (-0.09)	0.78 (+0.12)	0.5 (-0.13)	5	ALL CLASSES
CNN	466	20 (+0)	4 (+0)	-	-	0	0.34 (-0.12)	0.52 (+0.03)	0.33 (-0.18)	29	
CNN	534	19 (-1)	5 (+0)	-	-	0	0.45 (-0.09)	0.64 (+0.11)	0.41 (-0.2)	20	
CAP	446	17 (-1)	1 (-1)	-	-	0	0.33 (-0.25)	0.75 (+0.02)	0.3 (-0.24)	3	
CAP	466	20 (+0)	3 (-1)	-	-	0	0.26 (-0.13)	0.53 (+0.02)	0.23 (-0.19)	22	
CAP	534	19 (-1)	3 (-2)	-	-	0	0.34 (-0.15)	0.62 (+0.08)	0.3 (-0.21)	14	
CNN	446	18 (+0)	3 (-1)	18 (+0)	0 (-2)	5	0.61 (-0.07)	0.91 (+0.19)	0.58 (-0.12)	5	PURE CLASSES (TP + FP > Thres)
CNN	466	20 (+0)	4 (+0)	3 (-3)	0 (+0)	190	0.71 (+0.02)	0.92 (+0.18)	0.63 (-0.09)	17	
CNN	534	19 (-1)	5 (+0)	8 (+0)	0 (-2)	85	0.69 (+0)	0.98 (+0.28)	0.6 (-0.09)	20	
CAP	446	17 (-1)	1 (-1)	10 (-6)	0 (+0)	5	0.58 (-0.1)	0.95 (+0.09)	0.53 (-0.1)	2	
CAP	466	20 (+0)	3 (-1)	5 (-1)	0 (-1)	60	0.53 (-0.02)	0.79 (+0.16)	0.52 (-0.04)	16	
CAP	534	19 (-1)	3 (-2)	9 (-4)	0 (-2)	25	0.56 (-0.05)	0.91 (+0.25)	0.53 (-0.09)	13	
CNN	446	14 (+0)	3 (-1)	-	-	0	0.55 (-0.07)	0.74 (+0.12)	0.54 (-0.11)	7	BIOLOGICAL CLASSES
CNN	466	16 (+0)	4 (+0)	-	-	0	0.32 (-0.12)	0.47 (+0.02)	0.33 (-0.18)	45	

mod	data	classes		pure classes		thres	mean F1	mean Precision	mean Recall	discard [%]	
		true	empty	true	empty						
CNN	534	15 (-1)	5 (+0)	-	-	0	0.43 (-0.08)	0.57 (+0.1)	0.4 (-0.2)	45	
CAP	446	13 (-1)	1 (-1)	-	-	0	0.35 (-0.29)	0.75 (-0.03)	0.31 (-0.29)	7	
CAP	466	16 (+0)	3 (-1)	-	-	0	0.26 (-0.14)	0.5 (+0.02)	0.22 (-0.22)	45	
CAP	534	15 (-1)	3 (-2)	-	-	0	0.34 (-0.14)	0.6 (+0.07)	0.29 (-0.22)	34	
CNN	446	14 (+0)	3 (-1)	8 (-3)	0 (-1)	15	0.89 (+0.12)	0.94 (+0.17)	0.87 (+0.07)	3	
CNN	466	16 (+0)	4 (+0)	1 (-2)	0 (+0)	190	0.74 (+0.09)	0.87 (+0.29)	0.64 (-0.16)	33	
CNN	534	15 (-1)	5 (+0)	5 (+0)	0 (-2)	85	0.68 (+0.05)	0.98 (+0.35)	0.61 (-0.03)	45	
CAP	446	13 (-1)	1 (-1)	8 (-5)	0 (+0)	5	0.59 (-0.13)	0.94 (+0.06)	0.53 (-0.15)	5	
CAP	466	16 (+0)	3 (-1)	3 (-1)	0 (-1)	60	0.55 (+0.05)	0.76 (+0.19)	0.5 (-0.02)	41	
CAP	534	15 (-1)	3 (-2)	6 (-4)	0 (-2)	25	0.61 (+0.03)	0.91 (+0.28)	0.59 (-0.01)	33	

PURE BIOLOGICAL C. (TP + FP > Thres)

Supplementary Table S 2: Results of the Top-k-method with the CNN and k = 3. The variable 'n' is the number of images that need to be manually validated to maximize the recall. The variable 'N' is the number of images that belong to the respective field set (446: 55,302; 466: 7,798; 534: 31,848) including the three classes 'unknown', 'blurry', and 'marine snow'. When recall is empty, the model sorted images in a class which was not occupied in the field data set (TP = 0). The ratio 'CC:FC' (correct classified:false classified) provides the number of false images to be manually sorted for one true image found.

class	FS446					FS466					FS534				
	recall	TP	n	n / N [%]	CC:FC	recall	TP	n	n / N [%]	CC:FC	recall	TP	n	n / N [%]	CC:FC
amp		0	4824	8.7			0	353	4.5			0	5641	17.7	
app	0.997	298	6331	11.4	1:20.2	0.932	351	1050	13.5	1:2	0.987	1097	4760	14.9	1:3.3
bip	0.987	79	171	0.3	1:1.2	0.818	11	38	0.5	1:2.5	1	8	56	0.2	1:6
cop	0.98	151	3601	6.5	1:22.8	0.992	627	3364	43.1	1:4.4	0.999	2219	5638	17.7	1:1.5
dia	1	3190	11172	20.2	1:2.5	0.991	346	1982	25.4	1:4.7	1	16	588	1.8	1:35.8
ech		0	321	0.6			0	58	0.7			0	72	0.2	
egg	1	5	346	0.6	1:68.2		0					0	5034	15.8	
lar	1	15	85	0.2	1:4.7	0.571	7	14	0.2	1:1	1	1	27	0.1	1:26
mal	0.773	22	122	0.2	1:4.5	0.907	43	308	3.9	1:6.2	0.908	87	571	1.8	1:5.6
med	0.789	76	4353	7.9	1:56.3	0.847	144	1164	14.9	1:7.1	0.875	24	1124	3.5	1:45.8
mne	0.986	144	184	0.3	1:0.3	1	1	19	0.2	1:18	0.857	7	35	0.1	1:4
noc	0.971	348	5197	9.4	1:13.9	0.9	20	100	1.3	1:4	0.996	3696	16092	50.5	1:3.4
plu	0.998	9861	11675	21.1	1:0.2	0.939	212	510	6.5	1:1.4	0.967	1343	2804	8.8	1:1.1
pol	0.975	363	10896	19.7	1:29	0.625	16	328	4.2	1:19.5	0.818	22	977	3.1	1:43.4
pte		0	81	0.1			0	558	7.2			0	1787	5.6	
rod	0.954	2034	5684	10.3	1:1.8	0.914	814	1193	15.3	1:0.5	1	22	395	1.2	1:17
vel		0	90	0.2		1	34	98	1.3	1:1.9	0.936	47	1599	5	1:33
wor	0.991	705	2731	4.9	1:2.9	0.811	37	702	9	1:18	0.993	913	2435	7.6	1:1.7
zoe	1	27	77	0.1	1:1.9	1	7	260	3.3	1:36.1	0.92	25	126	0.4	1:4
Mean :	0.96	1155	3576	6.46	1:15	0.86	161	588	7.53	1:8	0.95	562	2272	7.12	1:14

Supplementary Table S 3: Statistical comparison between model predictions and field data sets regarding the spatial distribution of plankton organisms. F-statistic, recomputed degrees of freedom and p-value of Dutilleuls modified t-test are presented for the CNN and the CapsNet. 'n' is the true abundance of the respective class (manual classification), while the recall was estimated after filtering (automatic classification). Class keys refer to Table 1.

Dataset	class	CNN			CapsNet			n	recall	
		F	df	p	F	df	p		CNN	CapsNet
FS446	egg	0.48	297.10	0.000	NA	NA	NA	5	0.80	0.00
FS446	lar	0.00	266.64	0.824	NA	NA	NA	15	0.73	0.00
FS446	mal	0.07	89.90	0.017	NA	NA	NA	22	0.36	0.00
FS446	zoe	0.46	125.43	0.000	0.08	126.81	0.001	27	0.93	0.11
FS446	med	0.01	72.98	0.535	0.01	68.48	0.403	76	0.17	0.11
FS446	bip	1.78	69.50	0.000	0.50	69.83	0.000	79	0.96	0.80
FS446	blu	0.07	89.90	0.017	0.00	92.00	0.778	101	0.29	0.01
FS446	mne	4.16	108.19	0.000	0.30	113.74	0.000	144	0.95	0.58
FS446	cop	0.46	125.43	0.000	0.08	126.81	0.001	151	0.73	0.15
FS446	unk	0.01	177.45	0.293	0.01	171.36	0.348	208	0.13	0.04
FS446	app	0.00	266.64	0.824	0.00	252.42	0.580	298	0.05	0.01
FS446	noc	0.00	283.83	0.492	0.00	263.99	0.642	348	0.03	0.01
FS446	pol	0.48	297.10	0.000	0.22	290.81	0.000	363	0.53	0.02
FS446	wor	163.52	298.45	0.000	25.71	298.62	0.000	705	0.99	0.80
FS446	rod	0.07	991.48	0.000	NA	NA	NA	2034	0.01	0.00
FS446	dia	65.29	741.96	0.000	33.21	730.92	0.000	3190	0.89	0.85
FS446	plu	273.12	41.07	0.000	80.81	39.58	0.000	9861	0.98	0.93
FS446	sno	1321.22	453.18	0.000	442.50	480.06	0.000	37675	0.99	1.00
FS466	mne	0.45	233.10	0.000	0.01	30.98	0.680	1	1.00	1.00
FS466	act	0.45	233.10	0.000	NA	NA	NA	5	0.40	0.00
FS466	zoe	0.45	233.10	0.000	0.01	30.98	0.680	7	0.86	0.14

Dataset	class	CNN			CapsNet			n	recall	
		F	df	p	F	df	p		CNN	CapsNet
FS466	lar	0.45	233.10	0.000	NA	NA	NA	7	0.14	0.00
FS466	bip	0.45	233.10	0.000	0.01	30.98	0.680	11	0.82	0.73
FS466	pol	NA	NA	NA	NA	NA	NA	16	0.00	0.00
FS466	noc	0.45	233.10	0.000	0.01	30.98	0.680	20	0.10	0.25
FS466	vel	NA	NA	NA	0.01	30.98	0.680	34	0.00	0.15
FS466	wor	0.01	35.17	0.589	NA	NA	NA	37	0.35	0.00
FS466	mal	0.01	42.21	0.498	NA	NA	NA	43	0.67	0.00
FS466	pil	0.32	51.61	0.000	0.32	51.61	0.000	56	0.05	0.05
FS466	med	0.20	127.62	0.000	0.27	119.79	0.000	144	0.38	0.31
FS466	unk	0.00	141.64	0.566	NA	NA	NA	176	0.02	0.00
FS466	plu	0.70	158.16	0.000	0.53	163.26	0.000	212	0.65	0.61
FS466	dia	0.45	233.10	0.000	0.83	203.56	0.000	346	0.31	0.59
FS466	app	0.03	231.85	0.015	0.01	272.00	0.052	351	0.25	0.02
FS466	cop	0.38	483.81	0.000	0.05	485.90	0.000	627	0.64	0.31
FS466	rod	0.00	235.48	0.895	0.00	254.22	0.887	814	0.11	0.02
FS466	blu	0.75	463.74	0.000	0.00	444.78	0.398	1313	0.31	0.07
FS466	sno	18.35	126.00	0.000	9.52	143.48	0.000	3578	0.95	0.99
FS534	mne	0.06	17.29	0.341	0.05	17.81	0.368	7	0.86	0.43
FS534	bip	0.02	14.35	0.617	0.05	34.31	0.197	8	0.87	0.62
FS534	act	0.02	14.35	0.617	NA	NA	NA	11	0.55	0.00
FS534	dia	0.02	14.35	0.617	0.05	34.31	0.197	16	0.50	0.87
FS534	pil	0.02	14.35	0.617	0.05	34.31	0.197	19	0.16	0.16
FS534	rod	0.02	14.35	0.617	NA	NA	NA	22	0.14	0.00
FS534	pol	0.03	16.58	0.518	NA	NA	NA	22	0.27	0.00
FS534	med	0.06	17.29	0.341	0.05	17.81	0.368	24	0.42	0.37
FS534	zoe	0.24	14.82	0.081	0.02	12.95	0.646	25	0.80	0.16
FS534	vel	NA	NA	NA	0.05	34.31	0.197	47	0.00	0.04

Dataset	class	CNN			CapsNet			<i>n</i>	recall	
		F	<i>df</i>	<i>p</i>	F	<i>df</i>	<i>p</i>		CNN	CapsNet
FS534	mal	0.43	55.28	0.000	NA	NA	NA	87	0.53	0.00
FS534	unk	0.49	223.96	0.000	0.12	242.92	0.000	758	0.34	0.11
FS534	wor	92.62	22.50	0.000	51.49	22.34	0.000	913	0.93	0.89
FS534	blu	1.19	251.94	0.000	0.31	203.43	0.000	1014	0.50	0.16
FS534	app	0.45	141.98	0.000	0.09	199.37	0.000	1097	0.15	0.03
FS534	plu	54.66	16.62	0.000	47.46	16.48	0.000	1343	0.92	0.90
FS534	cop	25.66	28.53	0.000	5.72	27.47	0.000	2219	0.87	0.60
FS534	noc	0.68	58.03	0.000	1.13	47.73	0.000	3696	0.16	0.21
FS534	sno	214.65	10.09	0.000	142.02	11.22	0.000	20519	0.94	0.99

Supplementary Table S 4: Most frequent misinterpretations for each class pooled for all three field data sets. FP_{max} is the number of images of 'Class' that were predicted as 'Pred', while FP_{total} is the number of all wrong predictions for 'Class' (including FP_{max}).

class	CNN			CapsNet		
	pred	FP_{max}	FP_{total}	pred	FP_{max}	FP_{total}
actinotrocha	bipinnaria	2	5	marine snow	6	11
appendicularia	unknown	324	503	marine snow	594	799
bipinnaria	polychaeta	2	6	marine snow	15	20
blurry	marine snow	324	895	marine snow	1084	1276
copepods	blurry	43	123	marine snow	209	445
diatoms	marine snow	118	178	marine snow	317	317
eggs	mnemiopsis	1	1	mnemiopsis	1	1
larvae	appendicularia	2	8	marine snow	3	5
malacostraca	copepods	19	55	marine snow	22	69
medusae	marine snow	43	115	marine snow	114	182
mnemiopsis	eggs	3	7	marine snow	7	15
noctiluca	marine snow	203	284	marine snow	1740	1778
pilidium	pteropods	16	37	veliger	27	62
pluteus	marine snow	145	343	marine snow	709	851
polychaeta	marine snow	17	60	marine snow	166	182
rod	diatoms	2004	2260	diatoms	2334	2851
marine snow	noctiluca	758	1789	appendicularia	101	242
unknown	pteropods	138	423	marine snow	417	792
veliger	pilidium	43	55	pilidium	15	41
worms	appendicularia	27	93	diatoms	86	156
zoa	bipinnaria	2	8	marine snow	10	11

Chapter II

Automatic segregation of pelagic habitats

Rene-Marcel Plonus*¹, Stefanie Vogl², Jens Floeter¹

Keywords: machine learning, North Sea, sub-mesoscale, pelagic habitats, plankton patchiness

*Correspondence: rene-marcel.plonus@uni-hamburg.de

¹Institute of marine ecosystem and fishery science, University of Hamburg, Hamburg, Germany.

²Department of Informatics and Mathematics, Hochschule München, München, Germany.

Abstract

It remains difficult to segregate pelagic habitats since structuring processes are dynamic on a wide range of scales and clear boundaries in the open ocean are non-existent. However, to improve our knowledge about existing ecological niches and the processes shaping the enormous diversity of marine plankton, we need a better understanding of the driving forces behind plankton patchiness. Here we describe a new machine-learning method to detect and quantify pelagic habitats based on hydrographic measurements. An Autoencoder learns two-dimensional, meaningful representations of higher-dimensional micro-habitats, which are characterized by a variety of biotic and abiotic measurements from a high-speed ROTV. Subsequently, we apply a density-based clustering algorithm to group similar micro-habitats into associated pelagic macro-habitats in the German Bight of the North Sea. Three distinct macro-habitats, a 'surface mixed layer', a 'bottom layer' and an exceptionally 'productive layer' are consistently identified, each with its distinct plankton community. We provide evidence that the model detects relevant features like the doming of the thermocline within an Offshore Wind Farm or the presence of a tidal mixing front.

Introduction

Submesoscale features like eddies, fronts or filaments structure the pelagic realm at spatial scales of $O(1-10\text{km})$ (Lévy *et al.*, 2012; Shulman *et al.*, 2015; Buckingham *et al.*, 2016) and temporal scales that range from several hours to a few days (Baschek and Maarten Molemaker, 2010; Thompson *et al.*, 2016). Associated processes determine nutrient fluxes (Omand *et al.*, 2015; Thompson *et al.*, 2016) as well as plankton patchiness (Levy and Martin, 2013; Shulman *et al.*, 2015; Lévy *et al.*, 2018) and thereby even shape the seascape for top predators like sea birds (Bertrand *et al.*, 2014).

Recent advances in marine remote sensing technology (Wedding *et al.*, 2011) enabled scientists to separate benthic structures into mosaic-like patterns of different habitat classes (Hinchey *et al.*, 2008; Pittman *et al.*, 2011) following the role model of terrestrial ecosystems. However, what is well known and trivial in landscape ecology can be quite challenging in seascape ecology. While it remains difficult to segregate pelagic habitats, which exhibit no clear boundaries (Hinchey *et al.*, 2008; Pittman *et*

al., 2011; Wedding *et al.*, 2011) and can be quite dynamic on a wide range of scales, benthic habitat maps can give an impression of physically distinct areas that consistently occur together with particular species communities (Harris and Baker, 2012). Some effort has been undertaken to characterize fish habitats (e.g. Amorim *et al.*, 2018; Bellido *et al.*, 2008; Giannoulaki *et al.*, 2011; Tugores *et al.*, 2011; Laman *et al.*, 2017; Friedland *et al.*, 2020; Funk *et al.*, 2020), but fewer studies focused on zooplankton (e.g. Alvarez-Berastegui *et al.*, 2014; Labat *et al.*, 2009; Espinasse *et al.*, 2014). Thus, mechanisms contributing to the enormous diversity of plankton, a fundamental component of pelagic food webs, are still not fully understood (Sano *et al.*, 2013; North *et al.*, 2016). Understanding the processes shaping plankton communities is essential to improve our knowledge of existing ecological niches (Houliez *et al.*, 2021). Despite the growing awareness of the importance of spatial structure for ecology and management (Pittman *et al.*, 2011; Wedding *et al.*, 2011), there is still a lack of concepts and techniques applicable to characterize the spatial structure of the seascape in pelagic environments (Alvarez-Berastegui *et al.*, 2014). Mainly, because traditional oceanographic methods are inadequate for observing the submesoscale (Baschek and Maarten Molemaker, 2010) due to insufficient resolution and range (Marmorino *et al.*, 2018). Recent advances in instrumentation partially closed this gap, but there still is a need for novel analysis methods to take advantage of the existing data (North *et al.*, 2016). Some machine learning techniques are specifically designed to identify and characterize features in a 'sea of data', which makes it very promising to apply them also in this challenging field of research.

Autoencoders (AE) are a common tool in the machine learning community which consist of an encoding and a decoding part (Hinton, 2006). Initially devised to reduce (Encoder) and recover (Decoder) the dimensionality of their inputs (Hinton, 2006), they have been soon applied to a wide range of tasks like denoising (e.g. Vincent *et al.*, 2010) or anomaly detection (e.g. Chen *et al.*, 2018; Zhao *et al.*, 2017).

AEs do not classify or detect specific elements or objects in their inputs, but learn meaningful low dimensional representations, i.e. relevant high-level abstractions, of their inputs (Bengio *et al.*, 2006) so that the original data can be reconstructed as similar as possible by the decoder part. The input data don't need any pre-processing, e.g. labelling of subsets, by humans, since the target the network aims to

reconstruct is basically the original input. The compressed representations of the encoder can also be used as input for subsequent modelling, e.g. in a Convolutional Neural Network (CNN) application. In that case the unsupervised pre-training of a CNN embedded in an AE can help to capture more intricate dependencies (Erhan *et al.*, 2009) and better initialize the weights of the extended model (Bengio *et al.*, 2006). Thus, the (local) minimum in the loss surface of the AE corresponds to a good transformation of a high dimensional input to a lower dimensional intermediate output (output of the Encoder-part) (Bengio *et al.*, 2006), which would become the input for the classifier in a CNN. In this setting, the final output of the AE, the reconstructions, are secondary. However, a low reconstruction error of the AE ensures that the compressed signal incorporates the important features of the original high dimensional input data.

In this study we take advantage of this specific application of AEs. Instead of substituting the decoder part with a classification or regression network we use the compressed signal of the encoder as input for a subsequent clustering algorithm. We use a fully connected AE to reduce a high dimensional input consisting of a variety of abiotic and biotic oceanographic measurements to a lower dimensional meaningful representation (intermediate output), skip the decoding part after the training is completed and cluster the encoded features to macro-habitats. Similar micro-habitats lead to similar representations and therefore regions with different characteristics are segregated as different macro-habitats. These macro-habitats correspond to distinct pelagic habitats in the southern North Sea, whose plankton communities are compared and analyzed.

Material and Procedures

Data acquisition and preparation

Physical and biological oceanographic measurements were recorded on a North Sea summer cruise with the RV Heincke (HE429, July 19-24, 2014) with a MacArtney TRIAXUS Remotely Operated Towed Vehicle (ROTV). For a detailed description of the device see Plonus *et al.* (2021). The ROTV transects were located in the direct vicinity of two Offshore Wind Farms (OWF) BARD Offshore 1 (BARD) and Global Tech I (GTI) (Figure 1). The map was generated using QGIS v3.18 (QGIS.org, 2021)

with bathymetric metadata and Digital Terrain Model data products from the EMODnet Bathymetry portal (<http://www.emodnet-bathymetry.eu>, 15.7.21). The ROTV was towed at a speed of 8 knots (4.1 m s^{-1}) with a three-degree lateral offset to lessen any disturbance from the vessels wake. During most transects the ROTV was undulating with a vertical speed of 0.1 m s^{-1} from $\sim 4 \text{ m}$ below the sea surface to $\sim 8 \text{ m}$ above the sea floor. The horizontal resolution between two surface peaks was $\sim 560 \text{ m}$, while the vertical resolution was $\sim 0.3 \text{ m}$. The ROTV measured water temperature, salinity, oxygen, and chlorophyll-a at a frequency of 1 Hz and was equipped with a Video Plankton Recorder (VPR, Seascan Inc., Falmouth, MA02540, USA) which provided zoo- and phytoplankton densities on the taxonomic family-, and sometimes even genus-level. For a detailed description of the VPR plankton image classification see Floeter *et al.* (2017). We used a similar summer cruise with the RV Heincke five years later (HE534, June 16-21, 2019) as a test data set. For our analyses we selected the following variables: temperature ($^{\circ}\text{C}$), salinity (PSU), oxygen ($\mu\text{mol}\cdot\text{l}^{-1}$), density ($\text{kg} \cdot \text{m}^{-3}$), and chlorophyll-a (RFU). For each of the variables, we calculated the horizontal (grid cell to the left, i.e. $\sim 25 \text{ m}$) and vertical (grid cell above, i.e. 1 m) gradient. Furthermore, we had sufficient density data ($\text{N}\cdot\text{l}^{-1}$) available for the taxa 'Appendicularia', 'Copepoda', 'Dinoflagellates', 'Gastropoda', 'Jelly', 'Marine snow', 'Nauplii', 'Ophiuroida', 'Pilidium', 'Pluteus', and 'Polychaeta'.

Transect diagrams were generated using Ocean Data View (ODV, Schlitzer, 2020) with the embedded spatial interpolation software DIVA (Troupin *et al.*, 2012) and exported as grids with a resolution of $\sim 25 \text{ m}$ length x 1 m depth. Abiotic measurements as well as density values were normalized and rescaled to range from -1 to 1. This was necessary since deep learning models generally perform better with homogeneous, small values (Bishop, 1995).

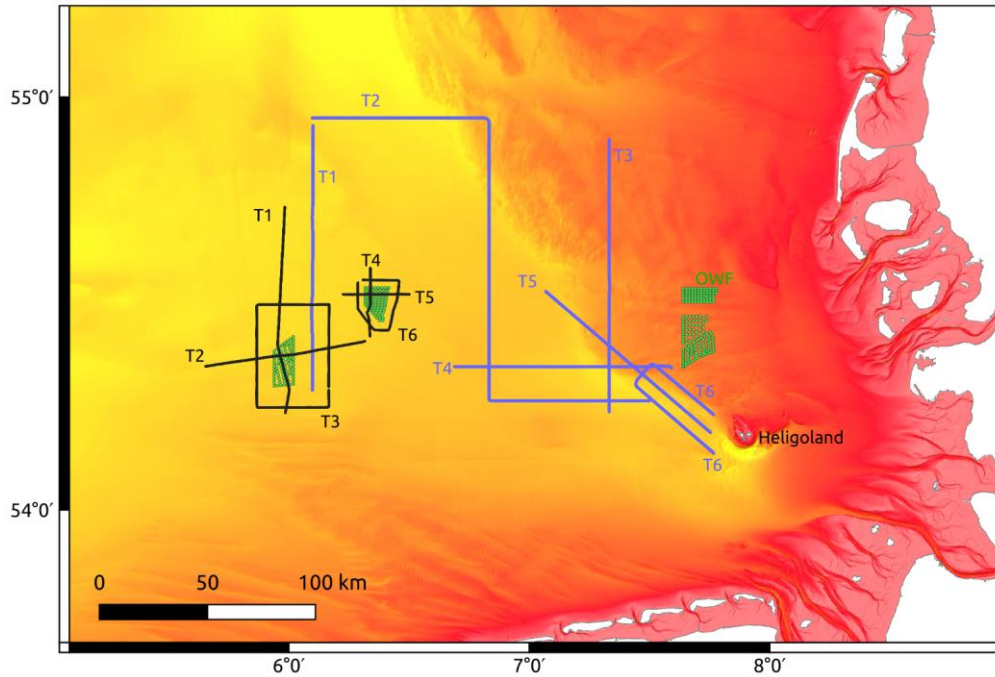


Figure 1: Sampling transects from HE429 (black) and HE534 (blue) in the German Bight of the North Sea. Green dots: Wind turbines. Depth ranges from 10 m (red) to 50 m (yellow).

To check for multicollinearity between our variables we calculated the variance inflation factor (VIF) in R (RCoreTeam, 2020) using functions provided by Zuur *et al.* (2009). A threshold of $VIF > 3$ was applied to identify highly collinear variables and exclude them from further analyses (Zuur *et al.*, 2010). The exported grids for each selected parameter were stacked and transformed into feature-vectors where each grid cell became one vector with 4 features (1 parameter = 1 feature). In our definition, a pelagic micro-habitat with a spatial extent of $\sim 25 \text{ m} \times 1 \text{ m}$ corresponds to one of those feature-vectors (Figure 2).

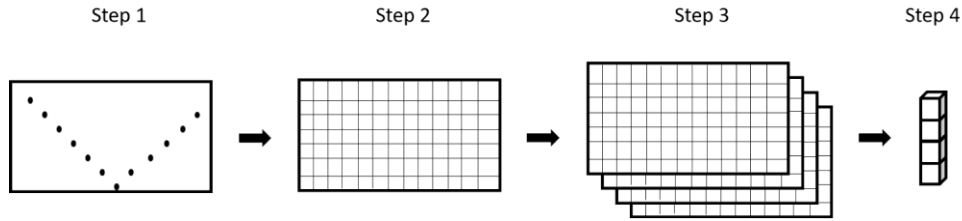


Figure 2: Schematic data processing from measurements to feature-vector. Step 1: original measurements; Step 2: gridded parameter table; Step 3: stacked grids; Step 4: feature-vector for 1 of the 91 grid-cells.

Based on these feature-vectors the AE was trained to reconstruct the original micro-habitats and thereby learn relevant abstractions that represent important patterns in the pelagic environment. We used a GPU supported Tensorflow backend (Abadi *et al.*, 2015) for Keras (Chollet, 2015) under Python 3.7 (Van Rossum and Drake, 2009) to build and train our AE.

Model description

The AE consisted of two fully connected layers in the Encoder and Decoder, respectively. The Decoder used the transposed weights of the Encoder in reversed order, e.g. the weights of the first Encoder-Layer were shared with the last Decoder-Layer. The first layer of the Encoder inflated the 4-dimensional feature-vector to a 100-dimensional feature-vector, which was reduced to a 2-dimensional feature-vector by the second layer (4 - 100 - 2). The Decoder did the same in reverse (2 - 100 - 4). The batch size (number of inputs that are processed simultaneously) was set to 38 and the learning rate followed a sawtooth-like scheme, initialized at $5e^{-8}$. Each input feature-vector corresponds to one micro-habitat and includes 1 measurement of each parameter selected for the analyses. The model was trained using the data from HE429 exclusively. Approximately ~ 13 % of the data was separated to validate the training process based on the remaining 87 %. Data from HE534 was used as a final test set. As an AE is a gradient-based method, the chosen starting point may be crucial for the final fit of the model (Hinton, 2006), and one way of assessing and reducing the effect of start conditions are multi-start approaches (Subbey, 2018). Therefore, we trained multiple models and selected the one with the smallest final validation RMSE.

Habitat segregation

By applying the trained Encoder only, we projected all micro-habitats into a xy-coordinate system using the 2-dimensional intermediate output. We will refer to the encoded outputs as 'Encoded Components' in the following. Micro-habitats with similar characteristics were projected closer to each other than micro-habitats with different characteristics. We used the Euclidean distance to calculate the dissimilarity matrix for the Encoded Components of the micro-habitats, which was clustered by the HDBSCAN algorithm (McInnes *et al.*, 2017). HDBSCAN uses a density-based linkage function, defining clusters by the size of the area in which a certain number of neighbors is found. Micro-habitats in 'sufficiently dense' regions were assigned to a macro-habitat. Obviously, the parameters 'size of the neighborhood' (Epsilon) and the 'critical number of neighbors' (min_samples) are determining the resulting clusters (dense areas) with HDBSCAN. Thus, we checked the resulting macro-habitats for multiple different combinations of these two parameters as well as 'min_cluster_size'. The parameter 'min_cluster_size' is the threshold that separates 'sufficiently dense' regions (clusters) from the random background noise. All micro-habitats that were not assigned to a specific macro-habitat by HDBSCAN got the label '-1'. Homogeneous regions in the transect produced more dense regions in the 2-dimensional surface that were more likely to trespass the 'min_cluster_size' threshold and were separated from other homogeneous water masses by less dense regions. We used the silhouette method (Rousseeuw, 1987) to select the best segregation of micro-habitats. The silhouette score ranges from '-1' to '1' and indicates how well each point fits into the assigned cluster (macro-habitat) and is one of the best performing indices available (Arbelaitz *et al.*, 2013). '-1' is probably wrong labeled, '0' is close to the decision boundary of two clusters and '1' means this specific point is far away from points of other clusters. The silhouette scores were calculated using the scikit-learn module (Pedregosa *et al.*, 2011) for python.

Analyses

We used ODV to add a transect plot of the identified macro-habitats to the original measurements and plankton densities. Isolines of selected parameter measurements were overlaid on the macro-habitat plots to investigate which feature characteristics contributed to the segregation and to assess the associated plankton communities.

We furthermore described the macro-habitat plankton communities by modified Species-Abundance-Plots (SAP). We calculated the relative number of micro-habitats by plankton density and taxonomic group for each cluster. As is common for SAPs we used a log2 scale for density. That way we visualized the shift in specific species densities between the macro-habitats of different segregations of the same ROTV survey transect.

Pelagic submesoscale features often are highly productive areas and aggregate particles (Levy and Martin, 2013; Lévy *et al.*, 2018). Therefore, we calculated Lloyd's mean crowding (Lloyd's MC) and Lloyd's index of patchiness (Lloyd's IP) with the R-function 'agg_index' from the 'epiphy' package (Gigot, 2018) and compared the results for different segregations of the same transects. Lloyd's index is > 1 if species were aggregated, 1 if the distribution is random and < 1 indicates an overdispersed distribution compared to a homogeneous distribution. The Index of aggregation proposed by Bez (2000) (Bez's IoA) was calculated in addition to Lloyd's IP.

Data handling was done with R (RCoreTeam, 2020) and some tidyverse packages (Wickham *et al.*, 2019), namely purrr, tibble, dplyr, ggplot2, and tidyr.

Results

In the initial VIF analysis with the full dataset a couple of parameters had $VIF > 3$. After removing 'density' which had the highest score, no further parameter exceeded this threshold (Supplementary Table S 1). After a detailed analysis of model sensitivities and reconstruction quality we decided to limit the final parameter selection to (1) vertical temperature difference to the grid cell above, (2) salinity, (3) oxygen and (4) chlorophyll-a concentration.

Model training

The root mean squared error (RMSE) after the first epoch ranged roughly between 0.7 and 1.0. Each training epoch took 10 - 15 seconds using a graphic card with 768 gpu-cores and we trained each model for 15 epochs until a plateau was reached (Supplementary Figure S 1). The final training and validation RMSE of our selected model were $RMSE_{Tr} \sim 0.33$ and $RMSE_{Val} \sim 0.35$ (Figure 3).

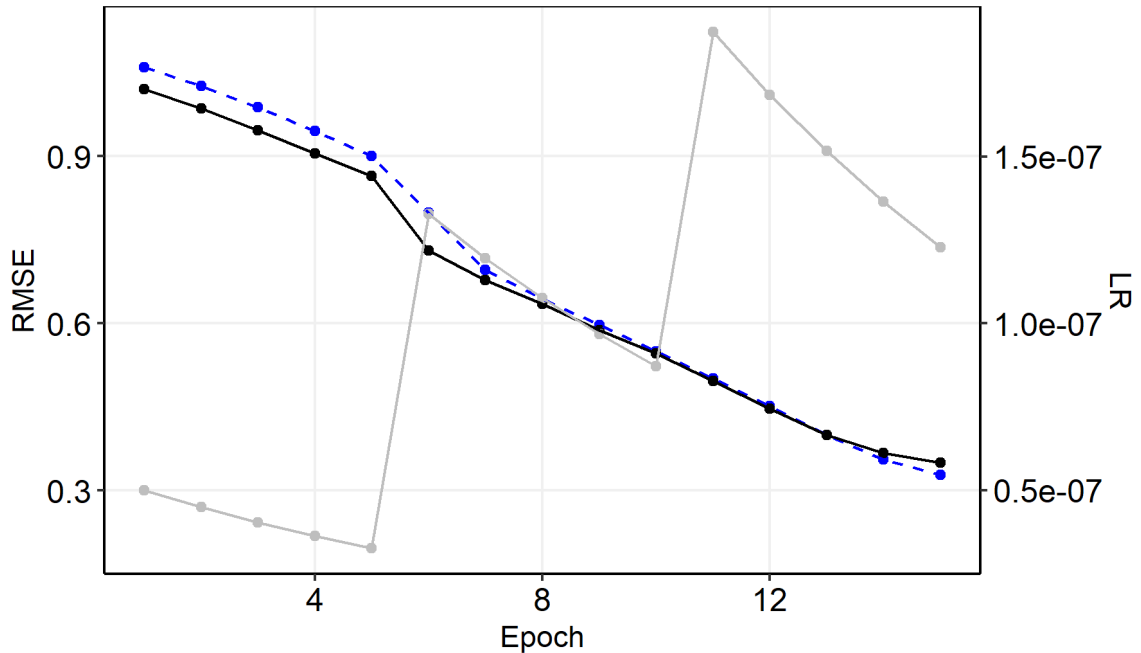


Figure 3: History of model training. black: validation, blue: training, grey: learning rate (sec. axis).

Clustering

Depending on the HDBSCAN parameter selection, micro-habitats were grouped into 2 - 20 macro-habitats that ranged in size from < 0.1 % to 97 % of all micro-habitats in a transect. We present exemplary the results for the segregation of T3 into different numbers of macro-habitats. Different parameter combinations could lead to an identical number of segregations. We chose an inverse size-cluster-relationship in the figure since more macro-habitats were usually ecologically less plausible (Figure 4). Mostly, 'epsilon' had a great impact on the segregation with specific combinations of 'min_cluster_size' and 'min_samples' but less influence with other tested combinations of those two parameters, indicating that segregations changed discontinuously with slopes and plateaus (Supplementary Figure S 2).

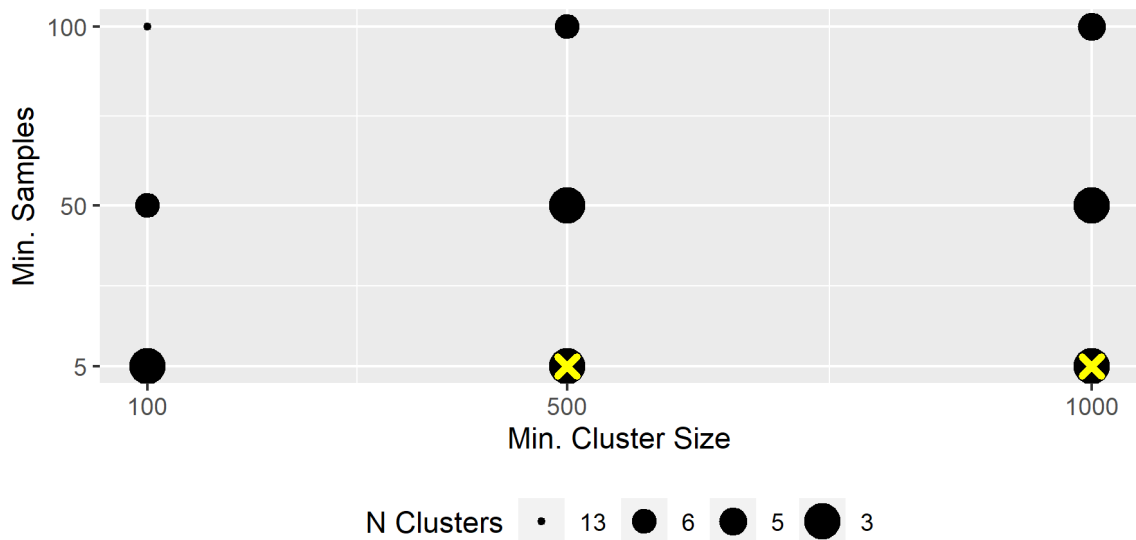


Figure 4: Number of segregated macro-habitats. The Figure was produced with $\epsilon = 0.32$. yellow cross: selected segregation.

Even though the label '-1' is used by HDBSCAN to indicate the lack of belonging to a specific cluster, we observed a close relationship between micro-habitats labelled '-1' and exceptionally strong chlorophyll-peaks throughout all transects. Therefore, we decided to treat '-1' as a macro-habitat of its own instead of unclassified micro-habitats. Micro-habitats labelled as '-1' were also frequently located between the BL and the SL.

Projections

While cluster-labelling was not consistent in that cluster '0' always referred to e.g. the 'surface mixed layer', the projections of the 'surface mixed layer' micro-habitats were always located in a similar position throughout all projection plots. Thus, while the cluster denotations related to a macro-habitat were not consistent, the position indicated the affiliation to a specific macro-habitat (Figure 5).

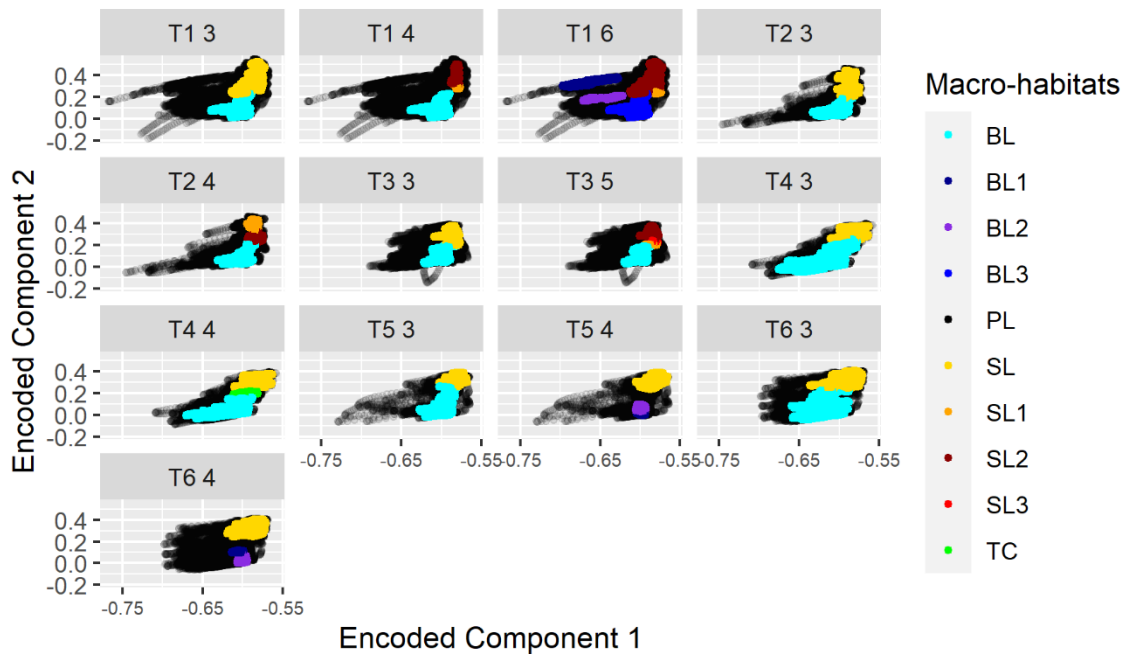


Figure 5: Projections for different segregations of all transects from HE429. The number of segregated macro-habitats is indicated by the last digit in the panel header. BL: Bottom layer (blue); SL: Surface mixed layer (yellow-red); TC: Thermocline Layer; PL: Productive layer (white). Numbers were used to indicate that e.g. more than 1 “Bottom layer” was segregated (e.g. B1/ B2 instead of BL).

Silhouette method

The segregation into three macro-habitats gave the highest average silhouette-scores in most cases: notably high chlorophyll-peaks were merged into one macro-habitat (1) and two further macro-habitats were separated at around 17°C in an upper surface mixed layer (2) and a lower bottom layer (3). There was only one exception from this rule in T1 where in the northern, deeper area the bottom layer (3) was replaced with the layer including the chlorophyll-peaks (1). Another anomaly occurred in T2, where one of the basic macro-habitats was further separated into two ‘sublayers’ so that a total of four macro-habitats were segregated. The highest silhouette-scores ranged from 0.35 (T1) to 0.59 (T5) (Table 1).

In the following we will use abbreviations for the three main layers and their sublayers, namely ‘PL’ for the productive layer with the high chlorophyll values, ‘BL’ for the bottom layer and ‘SL’ for the surface mixed layer. A segregation into more than one layer is indicated using numbers, e.g. ‘SL1’/ ‘SL2’ instead of ‘SL’.

Table 1: Silhouette method to select the best clustering. The numbers in ‘Segregation’ give the number of micro-habitats by macro-habitat, e.g. X_Y_Z indicates 3 macro-habitats with X, Y and Z micro-habitats, respectively.

Transect	Silhouette score	Segregation
HE429 T1	0.35	14583_21220_22015
HE429 T1	0.11	24657_21940_1949_9272
HE429 T1	0.07	12983_596_682_589_21028_21940
HE429 T2	0.54	5221_24912_12213
HE429 T2	0.56	6548_25232_5138_5428
HE429 T3	0.44	18748_21014_56730
HE429 T3	0.34	21660_56167_1371_15859_1435
HE429 T4	0.46	2054_5928_8107
HE429 T4	0.39	3202_5928_605_6354
HE429 T5	0.59	1729_4823_9331
HE429 T5	0.39	4582_5148_990_5163
HE429 T6	0.45	5558_14792_12965
HE429 T6	0.18	14589_12789_1229_4708
HE534 T1	0.40	7488_15344_53469
HE534 T2	0.26	21737_15535_123541
HE534 T2	-0.08	22248_4062_2099_132404
HE534 T2	-0.04	24788_15535_4115_1977_113215_1183
HE534 T3	-0.08	18944_2382_1156_1597_7140_9128

Habitat maps

We present T2 exemplarily for all transects of HE429 (Figure 6). Segregating the output of the Encoder into 3 macro-habitats, we got the typical scheme of a SL with temperatures above 17°C, a macro-habitat which was strongly associated with extraordinary high chlorophyll-peaks (PL) and a BL as a third macro-habitat. The average silhouette-score for the entire transect was 0.54 (Table 1). However, this clustering did not account e.g. for the intrusion of marine snow particles into the SL in

the eastern half of the transect. When accepting 4 different macro-habitats, the BL and PL macro-habitats were mostly unaffected, while the SL was further separated into 2 different macro-habitats. One corresponded to the area where marine snow particles were predominant while the second macro-habitat corresponded to the area where pluteus larvae were observed in high densities. Segregating characteristics of the two macro-habitats were a salinity difference of 0.2 and a shallowing of the thermocline from 10 m to 5 m water depth. Notably, this change around section distance 18-20 km was located at the entry point of the transect into the Offshore Wind Farm BARD. This segregation increased the average silhouette-score for the entire transect to 0.56 (Table 1).

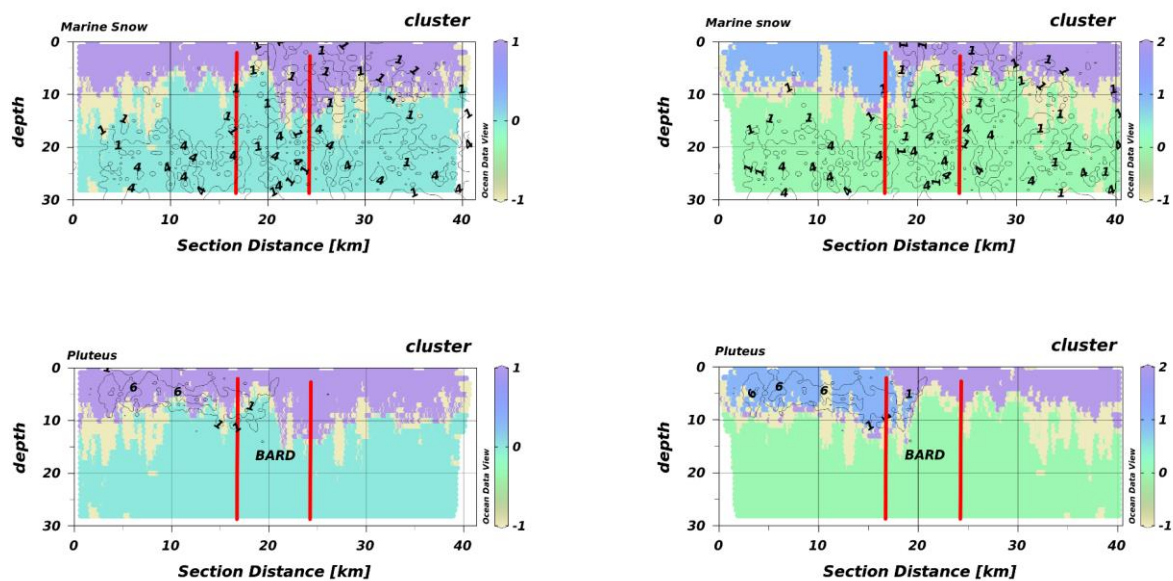


Figure 6: Habitatmap of T2 (HE429) with marine snow density (A, B) and pluteus density (C, D) as isolines. A, C: 3 segregated macro-habitats, B, D: 4 segregated macro-habitats.

Species Abundance Plots

The segregation into four macro-habitats was further supported by the modified SAPs. The relative amounts of PL and BL did not change much between 3 and 4 macro-habitats. However, SL1 included all micro-habitats with copepod densities $> 8 \text{ N}\cdot\text{l}^{-1}$ and basically all micro-habitats where pluteus larvae occurred. SL2 instead included micro-habitats with copepod densities $< 8 \text{ N}\cdot\text{l}^{-1}$ and generally less

chlorophyll, but most micro-habitats where appendicularia occurred. Thus, the SL1 and SL2 plankton communities were clearly distinct (Figure 7).

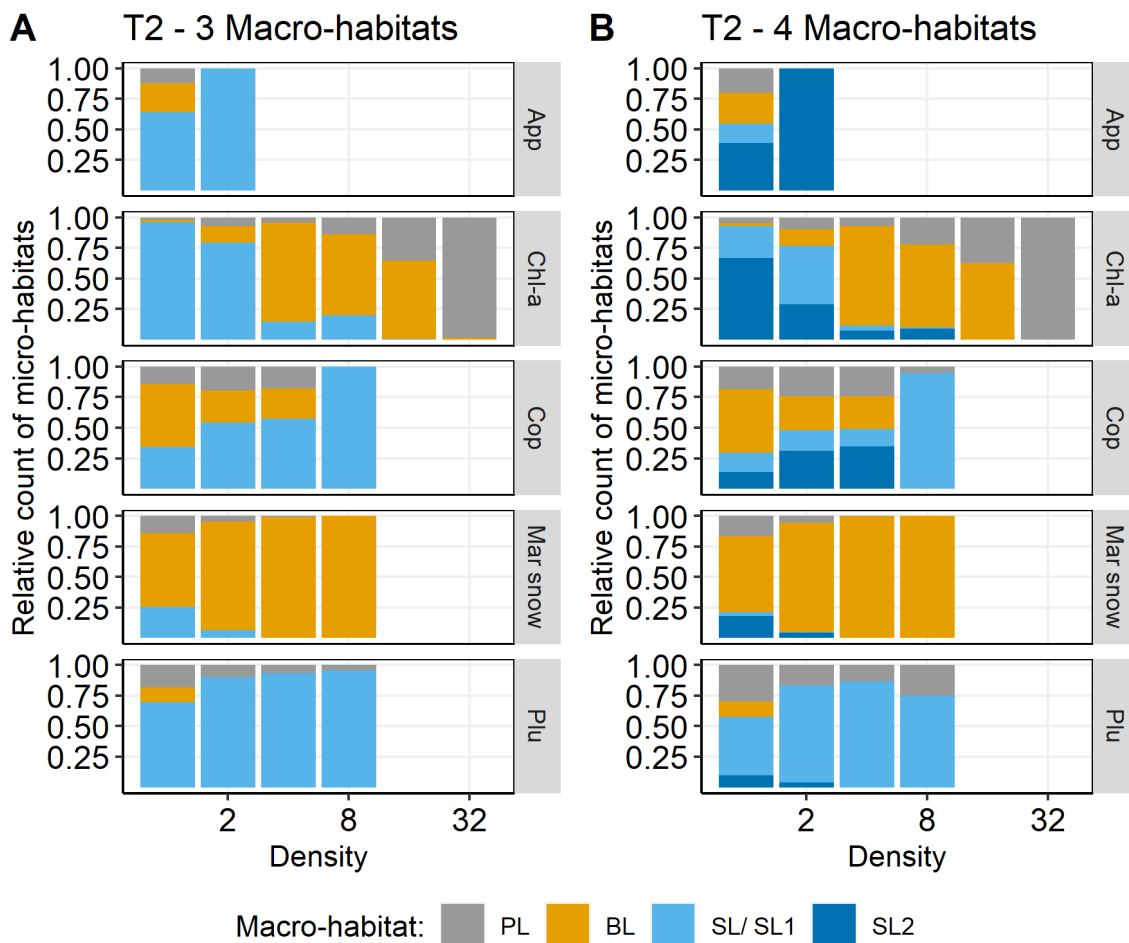


Figure 7: Relative count of density by macro-habitat for selected species on transect T2 (HE429). A: 3 macro-habitats, B: 4 macro-habitats. App: appendicularia ($N \cdot l^{-1}$); Chl-a: chlorophyll-a (RFU); Cop: copepoda ($N \cdot l^{-1}$); Mar snow: marine snow ($N \cdot l^{-1}$); Plu: pluteus ($N \cdot l^{-1}$).

Lloyd

Lloyd's mean crowding underpinned the SAP results. Patchiness in PL and BL did not change for 'marine snow' and 'pluteus' but differed clearly between SL1 and SL2, indicating a higher pluteus aggregation in SL1 and a higher aggregation of marine snow in SL2 (Table 2).

Table 2: Lloyd's mean crowding and Bez's Index of aggregation for 3 and 4 segregated macro-habitats for transect T2 (HE429). Presence gives the relative number of micro-habitats with density > 0. ms: marine snow; plu: pluteus.

Number of macro-habitats	macro-habitat	Plankton group	Lloyd's MC	Presence	Index of Aggregation
3	PL	ms	0.79	0.37	5.61e-04
3	BL	ms	2.35	0.78	6.72e-05
3	SL	ms	0.43	0.25	2.67e-04
4	PL	ms	0.65	0.34	4.50e-04
4	BL	ms	2.35	0.77	6.66e-05
4	SL1	ms	0.23	0.07	1.02e-03
4	SL2	ms	0.43	0.41	4.14e-04
3	PL	plu	2.19	0.14	2.10e-03
3	BL	plu	0.00	0.01	6.75e-04
3	SL	plu	3.56	0.36	3.78e-04
4	PL	plu	3.39	0.20	1.51e-03
4	BL	plu	0.00	0.01	6.58e-04
4	SL1	plu	3.56	0.67	4.63e-04
4	SL2	plu	0.36	0.06	1.57e-03

Test dataset HE534

The temperature maximum during HE534 was around 15 °C, i.e. 2 °C lower than the threshold that separated SL and BL in HE429. Consequently, no thermal stratification was detected by the Encoder trained with HE429 measurements. However, this model segregated an oxygen-rich layer that, based on the projections of the Encoder, resembled a similar habitat as the SL in HE429. This oxygen driven stratifications were not consistent over the entire range of a transect and in some areas the macro-habitat with projections similar to BL in HE429 comprised the entire water column, indicating a mixed water column closer towards the coast. Notably, plankton aggregations were commonly located at the border between oxygen-stratified and

mixed water columns (Figure 8). The highest average silhouette-scores were reached with three segregated macro-habitats. However, the scores were much lower compared to HE429 with a maximum between 0.26 and 0.40.

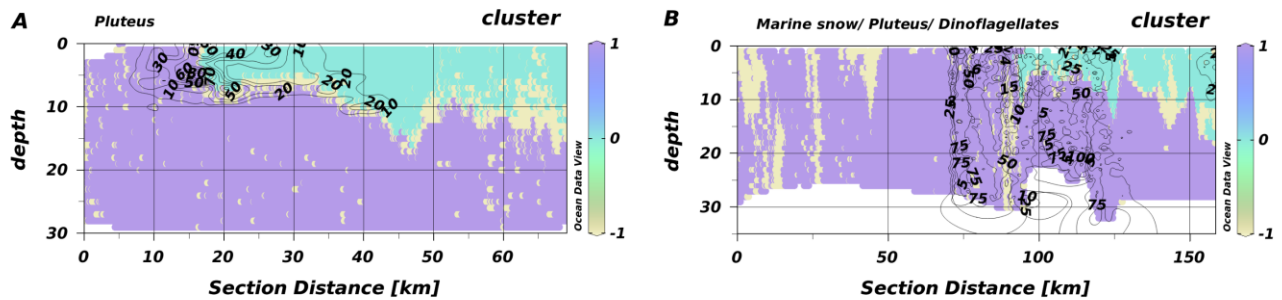


Figure 8: Habitat maps for T1 and T2 from HE534. The isolines give the densities ($N \cdot l^{-1}$) of pluteus for T1. Isolines in the figure of T2 include pluteus, marine snow and dinoflagellates (all in $N \cdot l^{-1}$). A: T1, B: T2.

Discussion

Selection of parameters

When training the model, we got the best results with a limited selection of parameters compared to the entire set of available data. The selected parameters are however in accordance with previous findings that physical properties contribute most to differences in habitat utilization by plankton organisms (Schulz *et al.*, 2012; Friedland *et al.*, 2020). In contrast to Alvarez-Berastegui *et al.* (2014), we did not benefit from the combination of gradients with the original measurements. However, a prior wavelet analysis as in North *et al.* (2016) could help to identify relevant spatial scales for the derivation of gradients. It is also possible, that the architecture of the model limited the amount of compressed information accessible to the clustering algorithm. In convolutional AEs, the size of the bottleneck (intermediate output) limits the generalization of the model (Manakov *et al.*, 2019). This is also true for the fully connected AE architecture of this model and might limit the potential of including more variables like species densities and environmental gradients.

Reconstruction loss

The loss for the optimization of an AE is based on the difference between the reconstruction and the original input. However, driving forces behind habitat partitioning vary with study region and season and specific parameters have a higher contribution than others (Schulz *et al.*, 2012; Espinasse *et al.*, 2014; Friedland *et al.*, 2020). Thus, we deemed it more important to accurately reconstruct specific features (parameters) compared to entire vectors (micro-habitats). Accordingly, we calculated the sum of the batchwise RMSE between the specific feature-values (e.g. temperature) of each input and the corresponding feature-values of the reconstructions and not the RMSE of an entire feature-vector (micro-habitat) and its reconstruction. This forced the AE to learn all parameters individually and furthermore made it possible to give specific parameters a higher priority if appropriate.

Aggregation

Lloyd's IP is an area-related quality measure for Lloyd's MC and thus sensitive to zeros. As can be seen in our example (Table 2), the 'spillover' from a crowded to an empty macro-habitat in the area of the decision boundary leads to misleadingly high Lloyd's IP and, to a lesser degree, misleadingly high Bez's IoA, even though this Index is supposedly insensitive to zeros. In accordance with the recommendation by Bez (2000) we therefore suggest Lloyd's MC as a measure of aggregation within a macro-habitat. Lloyd's IP and Bez's IoA might still be informative if the overall colonialization of the macro-habitat is considered.

Pelagic habitats

The model segregated three (four) distinct pelagic habitats in HE429: (1) a SL (SL1/SL2) mainly characterized by temperatures > 17 °C, (2) a BL on the other side of that threshold and (3) a PL dominated by high chlorophyll concentrations. In contrast to SL and BL, PL was not a true cluster by the definition of HDBSCAN, which indicates a great variability within the micro-habitats belonging to PL. That makes them 'special' or at least 'different' from common micro-habitats in SL and BL. Micro-habitats of PL were usually located around the 17 °C isoline and at the occurrence of exceptionally strong chlorophyll-peaks.

In the North Sea, peaks of primary production following the spring bloom were observed in subsurface layers (Richardson *et al.*, 1998, 2000). The PL most likely resembles such areas of subsurface productivity.

Furthermore, the model detected an upwards doming of the thermocline within an OWF, probably caused by enhanced vertical mixing (Segtnan and Christakos, 2015; Floeter *et al.*, 2017; Schultze *et al.*, 2020). The upwards doming and the resulting temperature differences are comparable to those observed within cyclonic eddies (Dong and McWilliams, 2007; Marmorino *et al.*, 2018), indicating that OWF's can influence the pelagic realm in the same order of magnitude as natural (sub-) mesoscale processes like eddies. The doming of colder, nutrient-rich water can produce chlorophyll peaks (Munk *et al.*, 1999), indicating the potential for an enhanced primary production in this area.

Cumulative effects of single foundations might lead to a blocking effect around OWF's, similar as observed for islands (Simpson *et al.*, 1982), which has the potential to produce submesoscale eddies (Dong and McWilliams, 2007) in addition to local upwelling fronts (Floeter *et al.*, 2017). Common properties that are used to describe hydrographic eddies and fronts include water velocity, vorticity and the Rossby number (e.g. Lévy *et al.*, 2012; Marmorino *et al.*, 2018), all of which were not available to us, which makes it less likely to detect such features.

The situation during HE534 was fundamentally different from HE429, most likely due to the weather conditions prior to the cruise that dispersed a thermal stratification. However, the projections indicated that similar SL and BL as in HE429 were detected. In case of HE534, segregations occurred along an oxygen isoline ($> 235 \mu\text{mol}\cdot\text{l}^{-1}$) instead of temperature ($> 17 \text{ }^\circ\text{C}$) as during HE429. This is in accordance with findings of Friedland *et al.* (2020) and references within that the predictive power of variables might change. This variable nature inherently present in pelagic data (Hinchey *et al.*, 2008; Thompson *et al.*, 2016) makes it so challenging to accurately predict pelagic habitats. The temperature isolines in T1 and T2 (Supplementary Figure S 3) clearly indicate the presence of a tidal mixing front (see Hill *et al.*, 1993). A convergence slick, which is typically associated with such tidal mixing fronts (Hill *et al.*, 1993), would also explain the observed aggregation of plankton particles at the intersection of the two macro-habitats (Figure 8).

There exists plenty of evidence that physical properties also structure the marine plankton communities (e.g. Leeuwen *et al.*, 2015; Swalethorp *et al.*, 2015; Lindegren *et al.*, 2020). However, the generated habitat maps have only limited explanatory power considering the observed plankton communities. This is not unexpected since physical properties are merely incomplete predictors for the community structure which is most likely further shaped by niche-based processes and interactions (Houliez *et al.*, 2021).

Top predators aggregate in areas with the highest prey-patch densities (not to be confused with the area of highest prey densities!) (Benoit-Bird *et al.*, 2013) and peak abundances of zooplankton and fish larvae are frequently observed in the direct vicinity of frontal convergence zones (Munk *et al.*, 1995, 2002; Höffle *et al.*, 2013; Munk, 2014; Swalethorp *et al.*, 2015). In addition to the horizontal agglomerations, thermo- and haloclines can produce further vertical structuring (Höffle *et al.*, 2013; Lindegren *et al.*, 2020). Thereby, more pronounced differences lead to a stronger niche separation and less overlap between different species (Lindegren *et al.*, 2020). Changes in nitrate (Scharfe and Wiltshire, 2019) and silicate (Wiltshire *et al.*, 2015) availability produce a temporal succession of different dominant taxa in the tidal advected phytoplankton community. Especially the plankton community is thus shaped by complex spatio-temporal dynamics and local prey patches have the potential to shape the distribution of higher trophic levels (Pope *et al.*, 1994; Burkhard *et al.*, 2011; Benoit-Bird *et al.*, 2013; Defriez *et al.*, 2016), even though this might be of less importance for ecosystem services in a highly diverse and partly functionally redundant plankton communities like that of the North Sea (Atkinson *et al.*, 2015).

Comments and recommendations

Future work should aim to include species densities and water current related measurements in order to accurately predict not only physical habitats but also realized ecological niches and hopefully improve our understanding of the complex dynamics shaping the pelagic realm.

Our approach offers beneficial properties to solve this challenge: the AE is a highly non-linear tool to reduce the dimensionality of a nearly unlimited amount of data that can be extended as needed. Additionally, HDBSCAN is a cluster algorithm that

makes as few assumptions as possible, i.e. regarding number or shape of clusters. HDBSCAN can also handle outliers on its own in opposite to e.g. k-means, and even enables to treat them in our case as an own macro-habitat. While machine learning might not give insight into the underlying mechanistic, it can give a starting point from which to begin future investigations (Friedland *et al.*, 2020).

Acknowledgments

The RV Heincke research cruises were supported by grant numbers AWI_HE429_00 and AWI_HE534_00.

References

- Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z. and Citro, C. *et al.* 2015 *TensorFlow: Large-scale machine learning on heterogeneous distributed systems*. Available at: <https://www.tensorflow.org/>.
- Alvarez-Berastegui, D., Ciannelli, L., Aparicio-Gonzalez, A., Reglero, P., Hidalgo, M. and López-Jurado, J. L. *et al.* 2014 'Spatial scale, means and gradients of hydrographic variables define pelagic seascapes of bluefin and bullet tuna spawning distribution', *PLOS ONE*, 9(10), p. e109338. doi: [10.1371/journal.pone.0109338](https://doi.org/10.1371/journal.pone.0109338).
- Amorim, E., Ramos, S., Elliott, M. and Bordalo, A. A. 2018 'Dynamic habitat use of an estuarine nursery seascape: Ontogenetic shifts in habitat suitability of the european flounder (*platichtys flesus*)', *Journal of Experimental Marine Biology and Ecology*, 506, pp. 49–60. doi: [10.1016/j.jembe.2018.05.011](https://doi.org/10.1016/j.jembe.2018.05.011).
- Arbelaitz, O., Gurrutxaga, I., Muguerza, J., Pérez, J. M. and Perona, I. 2013 'An extensive comparative study of cluster validity indices', *Pattern Recognition*, 46(1), pp. 243–256. doi: [10.1016/J.PATCOG.2012.07.021](https://doi.org/10.1016/J.PATCOG.2012.07.021).
- Atkinson, A., Harmer, R. A., Widdicombe, C. E., McEvoy, A. J., Smyth, T. J. and Cummings, D. G. *et al.* 2015 'Questioning the role of phenology shifts and trophic mismatching in a planktonic food web', *Progress in Oceanography*. (The UK western channel observatory: Integrating pelagic and benthic observations in a shelf sea ecosystem), 137, pp. 498–512. doi: [10.1016/j.pocean.2015.04.023](https://doi.org/10.1016/j.pocean.2015.04.023).
- Baschek, B. and Maarten Molemaker, J. 2010 'Aerial and in situ measurements of submesoscale eddies, fronts, and filaments', 12, p. 1106. Available at: <http://adsabs.harvard.edu/abs/2010EGUGA..12.1106B> (Accessed: 26 April 2021).
- Bellido, J. M., Brown, A. M., Valavanis, V. D., Giráldez, A., Pierce, G. J. and Iglesias, M. *et al.* 2008 'Identifying essential fish habitat for small pelagic species in spanish mediterranean waters', in Valavanis, V. D. (ed.) *Essential fish habitat mapping in the mediterranean*. Dordrecht: Springer Netherlands (Developments in hydrobiology), pp. 171–184. doi: [10.1007/978-1-4020-9141-4_13](https://doi.org/10.1007/978-1-4020-9141-4_13).
- Bengio, Y., Lamblin, P., Popovici, D. and Larochelle, H. 2006 'Greedy layer-wise training of deep networks', in *Proceedings of the 19th international conference on neural information processing systems*. Cambridge, MA, USA: MIT Press (NIPS'06), pp. 153–160.
- Benoit-Bird, K. J., Battaile, B. C., Heppell, S. A., Hoover, B., Irons, D. and Jones, N. *et al.* 2013 'Prey patch patterns predict habitat use by top marine predators with diverse foraging strategies', *PLOS ONE*, 8(1), p. e53348. doi: [10.1371/journal.pone.0053348](https://doi.org/10.1371/journal.pone.0053348).
- Bertrand, A., Grados, D., Colas, F., Bertrand, S., Capet, X. and Chaigneau, A. *et al.* 2014 'Broad impacts of fine-scale dynamics on seascape structure from zooplankton to seabirds', *Nature Communications*, 5(1), p. 5239. doi: [10.1038/ncomms6239](https://doi.org/10.1038/ncomms6239).
- Bez, N. 2000 'On the use of lloyd's index of patchiness', *Fisheries Oceanography*, 9(4), pp. 372–376. doi: <https://doi.org/10.1046/j.1365-2419.2000.00148.x>.
- Bishop, C. M. 1995 *Neural networks for pattern recognition*. Oxford: Oxford University Press.

- Buckingham, C. E., Garabato, A. C. N., Thompson, A. F., Brannigan, L., Lazar, A. and Marshall, D. P. *et al.* 2016 'Seasonality of submesoscale flows in the ocean surface boundary layer', *Geophysical Research Letters*, 43(5), pp. 2118–2126. doi: [10.1002/2016GL068009](https://doi.org/10.1002/2016GL068009).
- Burkhard, B., Opitz, S., Lenhart, H., Ahrendt, K., Garthe, S. and Mendel, B. *et al.* 2011 'Ecosystem based modeling and indication of ecological integrity in the german north sea—case study offshore wind parks', *Ecological Indicators*, 11(1), pp. 168–174. doi: [10.1016/j.ecolind.2009.07.004](https://doi.org/10.1016/j.ecolind.2009.07.004).
- Chen, Z., Yeo, C. K., Lee, B. S. and Lau, C. T. 2018 'Autoencoder-based network anomaly detection', in *2018 wireless telecommunications symposium (WTS)*. *2018 wireless telecommunications symposium (WTS)*, pp. 1–5. doi: [10.1109/WTS.2018.8363930](https://doi.org/10.1109/WTS.2018.8363930).
- Chollet, F. 2015 Keras. GitHub. Available at: <https://github.com/fchollet/keras>.
- Defriez, E. J., Sheppard, L. W., Reid, P. C. and Reuman, D. C. 2016 'Climate change-related regime shifts have altered spatial synchrony of plankton dynamics in the north sea', *Global Change Biology*, 22(6), pp. 2069–2080. doi: [10.1111/gcb.13229](https://doi.org/10.1111/gcb.13229).
- Devlin, M. J., Barry, J., Mills, D. K., Gowen, R. J., Foden, J. and Sivyler, D. *et al.* 2008 'Relationships between suspended particulate material, light attenuation and secchi depth in UK marine waters', *Estuarine, Coastal and Shelf Science*, 79(3), pp. 429–439. doi: [10.1016/j.ecss.2008.04.024](https://doi.org/10.1016/j.ecss.2008.04.024).
- Dong, C. and McWilliams, J. C. 2007 'A numerical study of island wakes in the southern california bight', *Continental Shelf Research*. (Recent developments in physical oceanographic modelling: Part IV), 27(9), pp. 1233–1248. doi: [10.1016/j.csr.2007.01.016](https://doi.org/10.1016/j.csr.2007.01.016).
- Erhan, D., Manzagol, P.-A., Bengio, Y., Bengio, S. and Vincent, P. 2009 'The difficulty of training deep architectures and the effect of unsupervised pre-training', in *Artificial intelligence and statistics*. PMLR, pp. 153–160.
- Espinasse, B., Carlotti, F., Zhou, M. and Devenon, J. L. 2014 'Defining zooplankton habitats in the gulf of lion (NW mediterranean sea) using size structure and environmental conditions', *Marine Ecology Progress Series*, 506, pp. 31–46. doi: [10.3354/meps10803](https://doi.org/10.3354/meps10803).
- Floeter, J., Beusekom, J. E. E. van, Auch, D., Callies, U., Carpenter, J. and Dudeck, T. *et al.* 2017 'Pelagic effects of offshore wind farm foundations in the stratified north sea', *Progress in Oceanography*, 156, pp. 154–173. doi: [10.1016/j.pocean.2017.07.003](https://doi.org/10.1016/j.pocean.2017.07.003).
- Friedland, K. D., Bachman, M., Davies, A., Frelat, R., McManus, M. C. and Morse, R. *et al.* 2020 'Machine learning highlights the importance of primary and secondary production in determining habitat for marine fish and macroinvertebrates', *Aquatic Conservation: Marine and Freshwater Ecosystems*, n/a(n/a). doi: <https://doi.org/10.1002/aqc.3527>.
- Funk, S., Krumme, U., Temming, A. and Möllmann, C. 2020 'Gillnet fishers' knowledge reveals seasonality in depth and habitat use of cod (*gadus morhua*) in the western baltic sea', *ICES Journal of Marine Science*, 77(5), pp. 1816–1829. doi: [10.1093/icesjms/fsaa071](https://doi.org/10.1093/icesjms/fsaa071).

- Giannoulaki, M., Pyrounaki, M. M., Liorzou, B., Leonori, I., Valavanis, V. D. and Tsagarakis, K. *et al.* 2011 'Habitat suitability modelling for sardine juveniles (*sardina pilchardus*) in the mediterranean sea', *Fisheries Oceanography*, 20(5), pp. 367–382. doi: <https://doi.org/10.1111/j.1365-2419.2011.00590.x>.
- Gigot, C. 2018 *Epiphy: Analysis of plant disease epidemics*. Available at: <https://CRAN.R-project.org/package=epiphy>.
- Harris, P. T. and Baker, E. K. 2012 '1 - Why map benthic habitats?', in Harris, P. T. and Baker, E. K. (eds) *Seafloor geomorphology as benthic habitat*. London: Elsevier, pp. 3–22. doi: [10.1016/B978-0-12-385140-6.00001-3](https://doi.org/10.1016/B978-0-12-385140-6.00001-3).
- Hill, A. E., James, I., Linden, P., Matthews, J., Prandle, D. and Simpson, J. *et al.* 1993 'Dynamics of tidal mixing fronts in the north sea', *Philosophical Transactions of The Royal Society B: Biological Sciences*, 343, pp. 431–446. doi: [10.1098/rsta.1993.0057](https://doi.org/10.1098/rsta.1993.0057).
- Hinchey, E. K., Nicholson, M. C., Zajac, R. N. and Irlandi, E. A. 2008 'Preface: Marine and coastal applications in landscape ecology', *Landscape Ecology*, 23(1), pp. 1–5. doi: [10.1007/s10980-007-9141-3](https://doi.org/10.1007/s10980-007-9141-3).
- Hinton, G. E. 2006 'Reducing the dimensionality of data with neural networks', *Science*, 313(5786), pp. 504–507. doi: [10.1126/science.1127647](https://doi.org/10.1126/science.1127647).
- Houliez, E., Lefebvre, S., Dessier, A., Huret, M., Marquis, E. and Bréret, M. *et al.* 2021 'Spatio-temporal drivers of microphytoplankton community in the bay of biscay: Do species ecological niches matter?', *Progress in Oceanography*, p. 102558. doi: [10.1016/j.pocean.2021.102558](https://doi.org/10.1016/j.pocean.2021.102558).
- Höffle, H., Nash, R. D. M., Falkenhaug, T. and Munk, P. 2013 'Differences in vertical and horizontal distribution of fish larvae and zooplankton, related to hydrography', *Marine Biology Research*, 9(7), pp. 629–644. doi: [10.1080/17451000.2013.765576](https://doi.org/10.1080/17451000.2013.765576).
- Kirby, R. R., Beaugrand, G. and Lindley, J. A. 2008 'Climate-induced effects on the meroplankton and the benthic-pelagic ecology of the north sea', *Limnology and Oceanography*, 53(5), pp. 1805–1815. doi: [10.4319/lo.2008.53.5.1805](https://doi.org/10.4319/lo.2008.53.5.1805).
- Labat, J.-P., Gasparini, S., Mousseau, L., Prieur, L., Boutoute, M. and Mayzaud, P. 2009 'Mesoscale distribution of zooplankton biomass in the northeast atlantic ocean determined with an optical plankton counter: Relationships with environmental structures', *Deep Sea Research Part I Oceanographic Research Papers*, 56, pp. 1742–1756. doi: [10.1016/j.dsr.2009.05.013](https://doi.org/10.1016/j.dsr.2009.05.013).
- Laman, E. A., Rooper, C. N., Turner, K., Rooney, S., Cooper, D. W. and Zimmermann, M. 2017 'Using species distribution models to describe essential fish habitat in alaska', *Canadian Journal of Fisheries and Aquatic Sciences*. doi: [10.1139/cjfas-2017-0181](https://doi.org/10.1139/cjfas-2017-0181).
- Leeuwen, S. van, Tett, P., Mills, D. and Molen, J. van der 2015 'Stratified and nonstratified areas in the north sea: Long-term variability and biological and policy implications', *Journal of Geophysical Research: Oceans*, 120(7), pp. 4670–4686. doi: [10.1002/2014JC010485](https://doi.org/10.1002/2014JC010485).
- Levy, M. and Martin, A. P. 2013 'The influence of mesoscale and submesoscale heterogeneity on ocean biogeochemical reactions', *Global Biogeochemical Cycles*, 27(4), pp. 1139–1150. doi: [10.1002/2012GB004518](https://doi.org/10.1002/2012GB004518).

- Lévy, M., Ferrari, R., Franks, P. J. S., Martin, A. P. and Rivière, P. 2012 'Bringing physics to life at the submesoscale', *Geophysical Research Letters*, 39(14). doi: [10.1029/2012GL052756](https://doi.org/10.1029/2012GL052756).
- Lévy, M., Franks, P. J. S. and Smith, K. S. 2018 'The role of submesoscale currents in structuring marine ecosystems', *Nature Communications*, 9(1), p. 4758. doi: [10.1038/s41467-018-07059-3](https://doi.org/10.1038/s41467-018-07059-3).
- Lindegren, M., Thomas, M. K., Jónasdóttir, S. H., Nielsen, T. G. and Munk, P. 2020 'Environmental niche separation promotes coexistence among ecologically similar zooplankton species—north sea copepods as a case study', *Limnology and Oceanography*, 65(3), pp. 545–556. doi: [10.1002/lno.11322](https://doi.org/10.1002/lno.11322).
- Manakov, I., Rohm, M. and Tresp, V. 2019 'Walking the tightrope: An investigation of the convolutional autoencoder bottleneck'. Available at: <https://openreview.net/forum?id=ryguP1BFwr> (Accessed: 23 September 2020).
- Marmorino, G. O., Smith, G. B., North, R. P. and Baschek, B. 2018 'Application of airborne infrared remote sensing to the study of ocean submesoscale eddies', *Frontiers in Mechanical Engineering*, 4. doi: [10.3389/fmech.2018.00010](https://doi.org/10.3389/fmech.2018.00010).
- Mavraki, N., Degraer, S., Vanaverbeke, J. and Braeckman, U. 2020 'Organic matter assimilation by hard substrate fauna in an offshore wind farm area: A pulse-chase study', *ICES Journal of Marine Science*, 77(7), pp. 2681–2693. doi: [10.1093/icesjms/fsaa133](https://doi.org/10.1093/icesjms/fsaa133).
- McInnes, L., Healy, J. and Astels, S. 2017 'Hdbscan: Hierarchical density based clustering', *Journal of Open Source Software*, 2(11), p. 205. doi: [10.21105/joss.00205](https://doi.org/10.21105/joss.00205).
- Munk, P. 2014 'Fish larvae at fronts: Horizontal and vertical distributions of gadoid fish larvae across a frontal zone at the norwegian trench', *Deep Sea Research Part II: Topical Studies in Oceanography*. (Fronts, fish and top predators), 107, pp. 3–14. doi: [10.1016/j.dsr2.2014.01.016](https://doi.org/10.1016/j.dsr2.2014.01.016).
- Munk, P., Larsson, P., Danielsen, D. and Moksness, E. 1995 'Larval and small juvenile cod gadus morhua concentrated in the highly productive areas of a shelf break front', *Marine Ecology Progress Series*, 125, pp. 21–30. doi: [10.3354/meps125021](https://doi.org/10.3354/meps125021).
- Munk, P., Larsson, P., Danielsen, D. and Moksness, E. 1999 'Variability in frontal zone formation and distribution of gadoid fish larvae at the shelf break in the northeastern north sea', *Marine Ecology Progress Series*, 177, pp. 221–233. doi: [10.3354/meps177221](https://doi.org/10.3354/meps177221).
- Munk, P., Wright, P. J. and Pihl, N. J. 2002 'Distribution of the early larval stages of cod, plaice and lesser sandeel across haline fronts in the north sea', *Estuarine, Coastal and Shelf Science*, 55(1), pp. 139–149. doi: [10.1006/ecss.2001.0892](https://doi.org/10.1006/ecss.2001.0892).
- North, R. P., Riethmüller, R. and Baschek, B. 2016 'Detecting small-scale horizontal gradients in the upper ocean using wavelet analysis', *Estuarine, Coastal and Shelf Science*, 180, pp. 221–229. doi: [10.1016/j.ecss.2016.06.031](https://doi.org/10.1016/j.ecss.2016.06.031).
- Omand, M. M., D'Asaro, E. A., Lee, C. M., Perry, M. J., Briggs, N. and Cetinić, I. *et al.* 2015 'Eddy-driven subduction exports particulate organic carbon from the spring bloom', *Science*, 348(6231), pp. 222–225. doi: [10.1126/science.1260062](https://doi.org/10.1126/science.1260062).

- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B. and Grisel, O. *et al.* 2011 'Scikit-learn: Machine learning in python', *MACHINE LEARNING IN PYTHON*, p. 6.
- Pittman, S. J., Kneib, R. T. and Simenstad, C. A. 2011 'Practicing coastal seascape ecology', *Marine Ecology Progress Series*, 427, pp. 187–190. doi: [10.3354/meps09139](https://doi.org/10.3354/meps09139).
- Plonus, R.-M., Conradt, J., Harmer, A., Janßen, S. and Floeter, J. 2021 'Automatic plankton image classification—can capsules and filters help cope with data set shift?', *Limnology and Oceanography: Methods*, 19(3), pp. 176–195. doi: <https://doi.org/10.1002/lom3.10413>.
- Pope, J. G., Shepherd, J. G., Webb, J., Stebbing, A. R. D., Mangel, M. and Beverton, R. J. H. *et al.* 1994 'Successful surf-riding on size spectra: The secret of survival in the sea', *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 343(1303), pp. 41–49. doi: [10.1098/rstb.1994.0006](https://doi.org/10.1098/rstb.1994.0006).
- QGIS.org 2021 *QGIS geographic information system*. QGIS Association. Available at: <http://www.qgis.org>.
- RCoreTeam 2020 *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Available at: <https://www.R-project.org/>.
- Richardson, K., Nielsen, T., Pedersen, F., Heilmann, J., Løkkegaard, B. and Kaas, H. 1998 'Spatial heterogeneity in the structure of the planktonic food web in the north sea', *Marine Ecology Progress Series*, 168, pp. 197–211. doi: [10.3354/meps168197](https://doi.org/10.3354/meps168197).
- Richardson, K., Visser, A. W. and Pedersen, F. B. 2000 'Subsurface phytoplankton blooms fuel pelagic production in the north sea', *Journal of Plankton Research*, 22(9), pp. 1663–1671. doi: [10.1093/plankt/22.9.1663](https://doi.org/10.1093/plankt/22.9.1663).
- Rousseeuw, P. J. 1987 'Silhouettes: A graphical aid to the interpretation and validation of cluster analysis', *Journal of Computational and Applied Mathematics*, 20, pp. 53–65. doi: [10.1016/0377-0427\(87\)90125-7](https://doi.org/10.1016/0377-0427(87)90125-7).
- Sano, M., Maki, K., Nishibe, Y., Nagata, T. and Nishida, S. 2013 'Feeding habits of mesopelagic copepods in sagami bay: Insights from integrative analysis', *Progress in Oceanography*, 110, pp. 11–26. doi: [10.1016/j.pocean.2013.01.004](https://doi.org/10.1016/j.pocean.2013.01.004).
- Scharfe, M. and Wiltshire, K. H. 2019 'Modeling of intra-annual abundance distributions: Constancy and variation in the phenology of marine phytoplankton species over five decades at helgoland roads (north sea)', *Ecological Modelling*, 404C, pp. 46–60. doi: [10.1016/j.ecolmodel.2019.01.001](https://doi.org/10.1016/j.ecolmodel.2019.01.001).
- Schlitzer, R. 2020 *Ocean data view*. Available at: <https://odv.awi.de/> (Accessed: 30 March 2021).
- Schultze, L. K. P., Merckelbach, L. M., Horstmann, J., Raasch, S. and Carpenter, J. R. 2020 'Increased mixing and turbulence in the wake of offshore wind farm foundations', *Journal of Geophysical Research: Oceans*, 125(8), p. e2019JC015858. doi: [10.1029/2019JC015858](https://doi.org/10.1029/2019JC015858).
- Schulz, J., Peck, M. A., Barz, K., Schmidt, J. O., Hansen, F. C. and Peters, J. *et al.* 2012 'Spatial and temporal habitat partitioning by zooplankton in the bornholm basin (central baltic sea)', *Progress in Oceanography*, 107, pp. 3–30. doi: [10.1016/j.pocean.2012.07.002](https://doi.org/10.1016/j.pocean.2012.07.002).

- Segtnan, O. H. and Christakos, K. 2015 'Effect of offshore wind farm design on the vertical motion of the ocean', *Energy Procedia*. (12th deep sea offshore wind r&D conference, EERA DeepWind'2015), 80, pp. 213–222. doi: [10.1016/j.egypro.2015.11.424](https://doi.org/10.1016/j.egypro.2015.11.424).
- Shulman, I., Penta, B., Richman, J., Jacobs, G., Anderson, S. and Sakalaukus, P. 2015 'Impact of submesoscale processes on dynamics of phytoplankton filaments', *Journal of Geophysical Research: Oceans*, 120(3), pp. 2050–2062. doi: [10.1002/2014JC010326](https://doi.org/10.1002/2014JC010326).
- Simpson, J. H., Tett, P. B., Argote-Espinoza, M. L., Edwards, A., Jones, K. J. and Savidge, G. 1982 'Mixing and phytoplankton growth around an island in a stratified sea', *Continental Shelf Research*, 1(1), pp. 15–31. doi: [10.1016/0278-4343\(82\)90030-9](https://doi.org/10.1016/0278-4343(82)90030-9).
- Subbey, S. 2018 'Parameter estimation in stock assessment modelling: Caveats with gradient-based algorithms', *ICES Journal of Marine Science*, 75(4), pp. 1511–1511. doi: [10.1093/icesjms/fsy060](https://doi.org/10.1093/icesjms/fsy060).
- Swailethorp, R., Malanski, E., Dalgaard Agersted, M., Gissel Nielsen, T. and Munk, P. 2015 'Structuring of zooplankton and fish larvae assemblages in a freshwater-influenced greenlandic fjord: Influence from hydrography and prey availability', *Journal of Plankton Research*, 37(1), pp. 102–119. doi: [10.1093/plankt/fbu099](https://doi.org/10.1093/plankt/fbu099).
- Thompson, A. F., Lazar, A., Buckingham, C., Naveira Garabato, A. C., Damerell, G. M. and Heywood, K. J. 2016 'Open-ocean submesoscale motions: A full seasonal cycle of mixed layer instabilities from gliders', *Journal of Physical Oceanography*, 46(4), pp. 1285–1307. doi: [10.1175/JPO-D-15-0170.1](https://doi.org/10.1175/JPO-D-15-0170.1).
- Troupin, C., Barth, A., Sirjacobs, D., Ouberdous, M., Brankart, J. M. and Brasseur, P. *et al.* 2012 'Generation of analysis and consistent error fields using the data interpolating variational analysis (DIVA)', *Ocean Modelling*, 52-53, pp. 90–101. doi: [10.1016/j.ocemod.2012.05.002](https://doi.org/10.1016/j.ocemod.2012.05.002).
- Tugores, M. P., Giannoulaki, M., Iglesias, M., Bonanno, A., Tičina, V. and Leonori, I. *et al.* 2011 'Habitat suitability modelling for sardine *sardina pilchardus* in a highly diverse ecosystem: The mediterranean sea', *Marine Ecology Progress Series*, 443, pp. 181–205. doi: [10.3354/meps09366](https://doi.org/10.3354/meps09366).
- Van Rossum, G. and Drake, F. L. 2009 *Python 3 reference manual*. Scotts Valley, CA: CreateSpace.
- Vincent, P., Larochelle, H., Lajoie, I., Bengio, Y. and Manzagol, P.-A. 2010 'Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion', *The Journal of Machine Learning Research*, 11, pp. 3371–3408.
- Wang, J., Zou, X., Yu, W., Zhang, D. and Wang, T. 2019 'Effects of established offshore wind farms on energy flow of coastal ecosystems: A case study of the rudong offshore wind farms in china', *Ocean & Coastal Management*, 171, pp. 111–118. doi: [10.1016/j.ocecoaman.2019.01.016](https://doi.org/10.1016/j.ocecoaman.2019.01.016).
- Wang, T., Yu, W., Zou, X., Zhang, D., Li, B. and Wang, J. *et al.* 2018 'Zooplankton community responses and the relation to environmental factors from established offshore wind farms within the rudong coastal area of china', *Journal of Coastal Research*, 34(4), pp. 843–855. Available at: <https://www.jstor.org/stable/26455006> (Accessed: 5 July 2021).

- Wedding, L. M., Lepczyk, C. A., Pittman, S. J., Friedlander, A. M. and Jorgensen, S. 2011 'Quantifying seascape structure: Extending terrestrial spatial pattern metrics to the marine realm', *Marine Ecology Progress Series*, 427, pp. 219–232. doi: [10.3354/meps09119](https://doi.org/10.3354/meps09119).
- Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D. and François, R. *et al.* 2019 'Welcome to the tidyverse', *Journal of Open Source Software*, 4(43), p. 1686. doi: [10.21105/joss.01686](https://doi.org/10.21105/joss.01686).
- Wiltshire, K. H., Boersma, M., Carstens, K., Kraberg, A. C., Peters, S. and Scharfe, M. 2015 'Control of phytoplankton in a shelf sea: Determination of the main drivers based on the helgoland roads time series', *Journal of Sea Research*, 105, pp. 42–52. doi: [10.1016/j.seares.2015.06.022](https://doi.org/10.1016/j.seares.2015.06.022).
- Wollschläger, J., Wiltshire, K., Petersen, W. and Metfies, K. 2015 'Analysis of phytoplankton distribution and community structure in the german bight with respect to the different size classes', *Journal of Sea Research*, 99. doi: [10.1016/j.seares.2015.02.005](https://doi.org/10.1016/j.seares.2015.02.005).
- Zhao, Y., Deng, B., Shen, C., Liu, Y., Lu, H. and Hua, X.-S. 2017 'Spatio-temporal AutoEncoder for video anomaly detection', in *Proceedings of the 25th ACM international conference on multimedia*. New York, NY, USA: Association for Computing Machinery (MM '17), pp. 1933–1941. doi: [10.1145/3123266.3123451](https://doi.org/10.1145/3123266.3123451).
- Zuur, A. F., Ieno, E. N. and Elphick, C. S. 2010 'A protocol for data exploration to avoid common statistical problems', *Methods in Ecology and Evolution*, 1(1), pp. 3–14. doi: <https://doi.org/10.1111/j.2041-210X.2009.00001.x>.
- Zuur, A., Ieno, E. N., Walker, N., Saveliev, A. A. and Smith, G. M. 2009 *Mixed effects models and extensions in ecology with r*. New York: Springer-Verlag (Statistics for biology and health). doi: [10.1007/978-0-387-87458-6](https://doi.org/10.1007/978-0-387-87458-6).

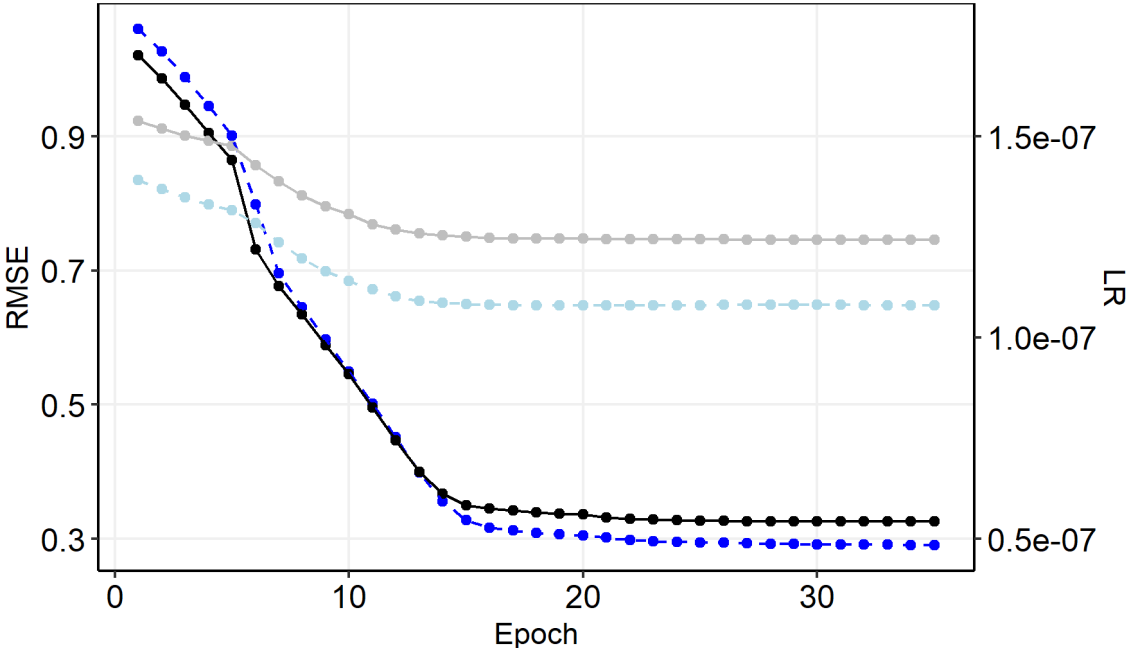
Supplementary material

Table

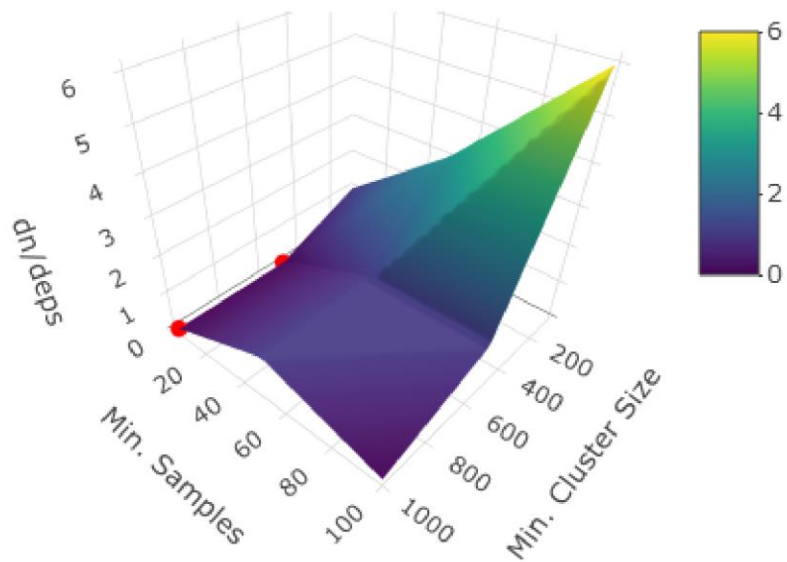
Supplementary Table S 1: Final parameter collecting and respective VIF values. NS: vertical gradient; WE: horizontal gradient.

Parameter	GVIF
dtNS	1.98
dtWE	1.59
Oxygen	2.16
doNS	1.84
doWE	1.57
Salinity	1.36
dsNS	1.09
dsWE	1.07
Chlorophyll_a	1.17
dcNS	1.12
dcWE	1.03
Appendicularia	1.04
Copepoda	1.25
Pluteus	1.13
MarineSnow	1.30
Dinoflagelattes	1.02

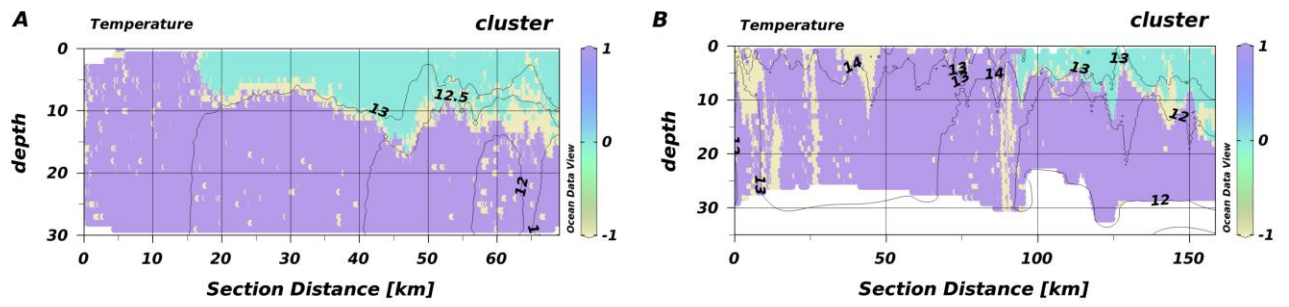
Figures



Supplementary Figure S 1: Prolonged history of model training. black: validation, blue: training.



Supplementary Figure S 2: Influence of Epsilon regarding the segregation of macro-habitats. $dn/deps$ is the difference in segregated macro-habitats at a specific range in Epsilon (0.2 - 0.32).



Supplementary Figure S 3: Habitat maps of T1 (A) and T2 (B) from HE534 overlaid by temperature isolines.

Identification of plankton habitats in the North Sea

Rene Plonus, Rolf Riethmüller, Jens Floeter

Abstract

The definition of an ecological niche makes it possible to anticipate the responses of a species to changing environmental conditions. Broad tolerance limits and a paucity of readily observable niches in the pelagic zone make it difficult to anticipate responses of the plankton community related to anthropogenic or environmental changes. Plankton distributions are closely linked to climate change and shape the seascape for higher trophic levels, so monitoring plankton distributions and defining ecological niches will help to understand and predict ecosystem responses. Here we apply a machine learning autoencoder and a density-based clustering algorithm to high-frequency datasets sampled with a ROTV Triaxus and a SCANFISH in the North Sea. The results indicate that in the highly dynamic North Sea, local hydrography prevents niche-based separation of plankton species at the sub-mesoscale, despite the availability of different habitats. Plankton patches were associated with naturally occurring frontal systems and anthropogenically induced upwelling-downwelling dipoles in the vicinity of offshore wind farms (OWFs). This highlights the potential for OWFs to induce local productivity patches and temporarily alter the seascape for higher trophic levels.

Introduction

The concept of an 'ecological niche' was first applied by Grinnell (1917) and referred to the abiotic demands of a species towards its environment and the behavioral adaptations of the species to the same (Grinnellian niche). A formal definition of the concept followed in Hutchinson (1957), describing an ecological niche as a 'n-dimensional hypervolume', where each dimension is influenced by a different environmental parameter. A further specification by Hutchinson (1957) was the distinction between fundamental and realized ecological niches. While the

fundamental niche represents the total area that allows a population to survive, the realized niche describes the area in which a species dominates over other competitors. The latter refers more to the 'Eltonian niche', which focuses on the interaction of different species (Elton, 1927), especially in the more modern differentiation of the two (Dehling and Stouffer, 2018). The original understanding of the Eltonian niche as stated by Hutchinson (1957) assumes that co-occurring species occupy different niches that do not intersect. The paucity of readily observable physical niches in the pelagic zone (Behrenfeld *et al.*, 2021) and a seemingly unstructured environment (Martin *et al.*, 2021) had led to the formulation of the 'paradox of plankton' by Hutchinson (1961): the co-existence of relatively many species in an apparently homogeneous environment, even though species richness tends to increase with habitat heterogeneity (MacArthur and MacArthur, 1961; Lapointe and Bourget, 1999). However, niche-based models do not explain redundancy (Leibold and McPeck, 2006) or the existence of functional groups or traits (Barton *et al.*, 2013; Dehling and Stouffer, 2018), both of which have been observed in plankton communities and are generally considered positive aspects of biodiversity (Leibold and McPeck, 2006). 'Neutral community models' see communities as assemblages of ecologically similar species in the absence of predation (Bell, 2001). It has been suggested that meaningful insights could be gained by the synthesis of niche-based and neutral theories (Leibold and McPeck, 2006). In addition, the traditional view of turbulence homogenized plankton communities has been challenged by recent studies, suggesting that biological and physical processes create a structured realm at scales down to few centimeters (Mitchell *et al.*, 2008; Basterretxea *et al.*, 2020).

Plankton communities in turn shape the seascape for higher trophic levels (Bertrand *et al.*, 2014) and are considered sentinels of ocean changes (Drago *et al.*, 2022), which makes it an important task to understand and monitor their spatio-temporal variation (Hays *et al.*, 2005; McGill *et al.*, 2006; McGinty *et al.*, 2018; Friedland *et al.*, 2020). However, traditional methods in plankton ecology have been time consuming and thus prevented the up-scale to pan-oceanic observations (Irisson *et al.*, 2022). This paucity of data and the inconsistency in sampling methods and scales has limited our understanding of the factors and processes determining abundance or diversity of plankton (Lombard *et al.*, 2019). New optical sampling methods have

emerged in the last decade which produce a wealth of information (Irisson *et al.*, 2022), but their scientific use was limited by the concepts and methods applicable to the huge amount of data they generate (Alvarez-Berastegui *et al.*, 2014; North *et al.*, 2016; Lombard *et al.*, 2019; Irisson *et al.*, 2022).

That's where machine learning excels. But even though machine learning has already successfully brought insights into pelagic ecology in the past, fully automated predictions can still only be trusted for the most abundant species (Plonus *et al.*, 2021a; Irisson *et al.*, 2022). Fortunately, plankton communities are usually highly diverse (Siegel, 1998) but dominated by a few, very abundant taxa (Fuhrman, 2009).

Still, the dynamic environment of plankton communities makes it challenging to determine and predict ecological niches and community structures (Beaugrand *et al.*, 2013; Brun *et al.*, 2015; Barth and Stone, 2022). In addition, partially functional redundant species make niche predictions an even more complex task (Fuhrman, 2009; Atkinson *et al.*, 2015; Brun *et al.*, 2015), especially in diverse ecosystems such as the North Sea. The study of plankton communities is a constantly changing field (Brun *et al.*, 2015) and becomes ever more precarious due to the tight coupling between plankton and climate change (Hays *et al.*, 2005; Beaugrand *et al.*, 2009; Barton *et al.*, 2013; Matus-Hernández *et al.*, 2019). However, there has been evidence that plankton community changes are predictable and can help to anticipate responses in a changing world (Beaugrand *et al.*, 2009; Fuhrman, 2009).

The basis of these predictions are habitat maps, which link bio-physically distinct areas to specific species communities (Harris and Baker, 2012). However, due to the high variability of relevant spatial and temporal scales in the pelagic environment (Hinchey *et al.*, 2008; Thompson *et al.*, 2016) it is a daunting task to accurately determine pelagic habitats and identify associated plankton communities. Here we present a fully automated method of mapping plankton communities to physically distinct habitats. Our approach takes advantage of machine learning speed to pry information from a wealth of data and make it accessible to human researchers. Using a fully connected Autoencoder (AE) and a density-based clustering algorithm we generate habitat maps from high-frequency data including physical and biological variables, sampled using a Remotely Operated Towed Vehicle (ROTV).

Material and methods

Data acquisition and preparation

Physical and biological oceanographic measurements were recorded on different North Sea surveys with the RV Heincke (Knust *et al.*, 2017) using a MacArtney TRIAXUS ROTV, complemented by a Video Plankton Recorder (VPR), and a ScanFish Mark III. A detailed description of the TRIAXUS and the associated sampling procedure can be found in Plonus *et al.* (2021a) and Plonus *et al.* (2021b). A detailed description of the SCANFISH sampling is given in Zhao *et al.* (2019). The map was generated using ggplot2 in R (Figure 1). For our analyses we selected the following parameters to determine bio-physical niches in the pelagic realm: temperature ($^{\circ}\text{C}$), salinity (psu), oxygen ($\mu\text{mol}\cdot\text{l}^{-1}$), and chlorophyll-a (RFU). Furthermore, we had sufficient VPR derived density data ($\text{N}\cdot\text{l}^{-1}$) available for the plankton groups 'Appendicularia', 'Copepoda', 'Dinoflagellates', 'Marine snow', and 'Pluteus larvae'.

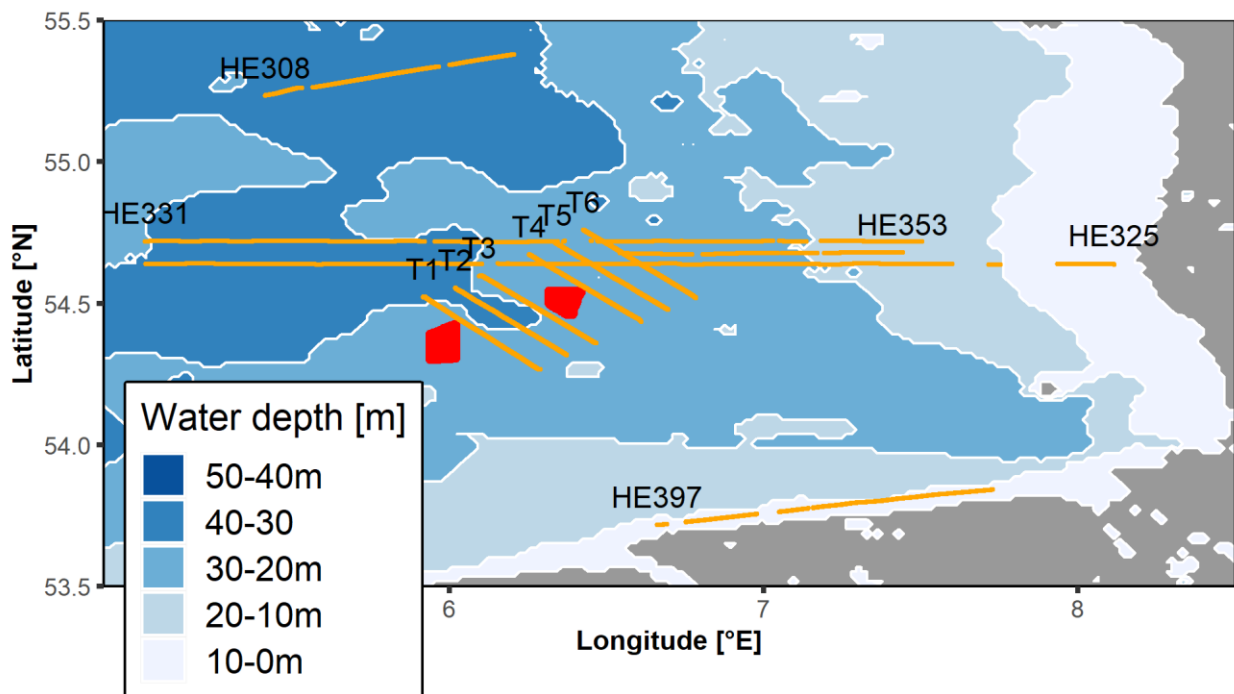


Figure 1: Sampling transects from HE466 (T1 - T6) and the Scanfish cruises. Red dots: Wind turbines. Depth ranges from 0 m (white) to 50 m (dark blue). A map of the cruises HE429 and HE534 was provided in Plonus *et al.* 2021b.

Transect diagrams were generated using Ocean Data View (ODV, Schlitzer, 2020) with the embedded spatial interpolation software DIVA (Troupin *et al.*, 2012) and exported as grids with a resolution of ~ 25 m length x 1 m depth. Abiotic measurements were normalized and rescaled to range from -1 to 1. All plankton density values and chlorophyll a measurements were transformed using the logarithm of $x + 1$ prior to normalization. The euclidean distances in the multidimensional space defined by the plankton densities and the chlorophyll a concentration between each grid cell and the top left grid cell of each transect were calculated (Harris *et al.*, 2020). Those distances were again normalized and rescaled as described above. This was necessary since deep learning models generally perform better with homogeneous, small values (Bishop, 1995).

The exported grids for each selected parameter and the above calculated distance measure were stacked and transformed into feature-vectors where each grid cell became one vector with 4 features (1 parameter = 1 feature). In our definition, a pelagic micro-habitat with a spatial extent of ~ 25 m x 1 m corresponds to one of those feature-vectors, which reflect the niche space at this point (Colwell and Rangel, 2009).

Based on these feature-vectors the AE was trained to reconstruct the original micro-habitats and thereby learn relevant abstractions that represent important patterns in the pelagic environment. We used a GPU supported Tensorflow backend (Abadi *et al.*, 2015) for Keras (Chollet, 2015) under Python 3.7 (Van Rossum and Drake, 2009) to build and train our AE.

Model description

The AE consisted of four fully connected layers in the Encoder and Decoder, respectively. The Decoder used the transposed weights of the Encoder in reversed order, e.g. the weights of the first Encoder-Layer were shared with the last Decoder-Layer. The first layer of the Encoder inflated the 4-dimensional feature-vector to a 200-dimensional feature-vector, which was reduced to a 100-, 50- and 4-dimensional feature-vector by the following layers (4 - 200 - 100 - 50 - 4). The Decoder did the same in reverse (4 - 50 - 100 - 200 - 4). The batchsize (number of input micro-habitats that are processed simultaneously) was set to 100 and the learning rate followed a sawtooth-like scheme, initialized at $5e^{-8}$. Each input feature-vector

corresponded to one micro-habitat and included 1 measurement of each parameter. The model was trained using data recorded during the research cruises HE429 (July 2014) and HE534 (June 2019). The best performance was achieved using a limited training set of only 50.000 randomly selected micro-habitats over 250 epochs while reserving the remaining ~200.000 micro-habitats for validation.

Micro-habitat segregation

The final model was tested using data generated on the research cruise HE466 (June 2016). In opposite to the human validated plankton densities from HE429 and HE534, the plankton densities for HE466 were estimated exclusively automatic using the threshold procedure proposed by Faillettaz *et al.* (2016). By applying the trained Encoder only, we created 4-dimensional representations of the original input. In the following we will refer to the processing of the micro-habitats by the Encoder as 'projection'. Micro-habitats with similar characteristics were projected closer to each other than micro-habitats with different characteristics. HDBSCAN calculates the Euclidean distance to build clusters from the, in this case, 4-dimensional inputs (McInnes *et al.*, 2017). For more information regarding the clustering see Plonus *et al.* (2021b). Further testing was conducted using data sampled during SCANFISH cruises HE308, HE331, HE353 and HE397.

Macro-habitat segregation

A drawback of HDBSCAN is the inconsistency between transects. Cluster 0 in transect A is not necessarily a similar macro-habitat to cluster 0 in transect B. Therefore, we calculated the 'center of gravity' (COG) for each cluster to identify similar macro-habitats in the entirety of all transects. Here, a COG is basically the average location of a cluster within the 4-dimensional space of the projections. We used a weighted average location based on the probabilities generated by HDBSCAN, which indicated whether a point was close to the center (1) of an assigned cluster or in the periphery ($\rightarrow 0$). In the final segregation step, the COGs were also clustered using HDBSCAN to identify global (including all transects) instead of local (each transect alone) 'groups of macro-habitats'. Thus, we make a distinction between local macro-habitats (MH) and global macro-habitats (GMH). GMHs do not necessarily occur in every transect, but at least in one (Figure 2).

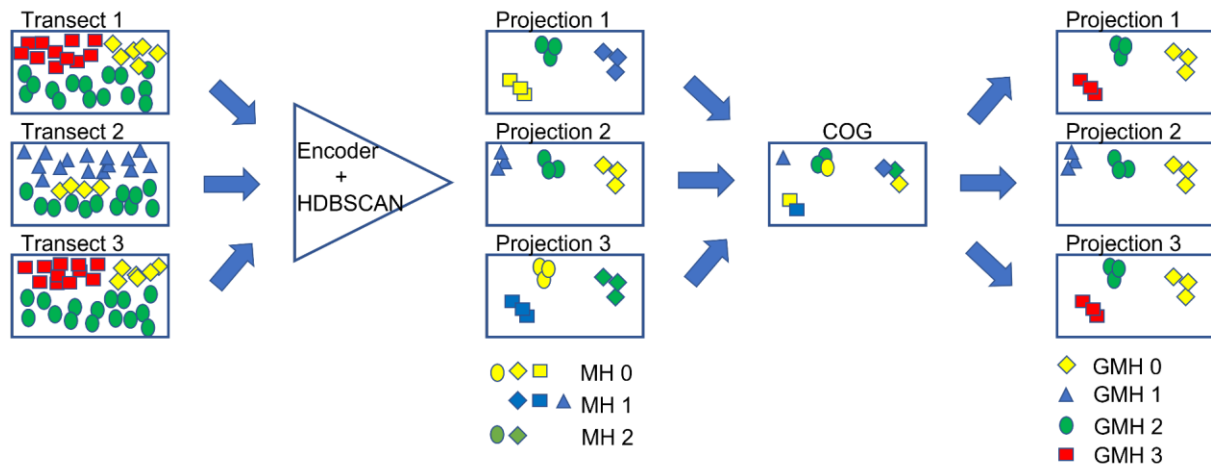


Figure 2: Schematic of the workflow. Micro-habitats from each transect are projected and clustered into macro-habitats (MH). Centres of gravity (COG) are calculated for each macro-habitat and all COGs are clustered again to identify global macro-habitats (GMH). Each symbol in the “transect” and “projection” maps represents 1 micro-habitat, while each symbol in “COG” represents the COG of one of the macro-habitats.

Macro-habitats in smaller transects (< 50 km) were grouped with macro-habitats from other transects exclusively. Macro-habitats of longer transects were sometimes also grouped with neighboring macro-habitats of the same transect, indicating a differentiation less meaningful by HDBSCAN or a suboptimal setting of the parameters. In such cases we merged all local macro-habitats within the transect of the same global macro-habitat and then updated the COG.

Identification of key parameters

While very deep architectures can easily become some kind of ‘black box’, it was quite possible to trace individual inputs in our relatively shallow model which on top consisted only of fully connected Dense-Layers and did not rely on convolutional filters. Fully connected layers take a n-dimensional input which is multiplied by a pre-defined number ‘X’ of weight sets of length n, resulting in a weight matrix of n x X weights. Basically, the output of a Dense-Layer is the dot product of the input and the weight matrix fed into some nonlinear activation function and thus the weights ultimately define the relative influence of a given input upon the final output. Weights close to zero result in minor changes in the output even if the input variable varies a lot while increasing weights (positive as well as negative) can facilitate major

changes in the output. Thus, analyzing the final weights after completion of the training phase revealed the relative influence each input variable had upon a specific output dimension of the Encoder, regardless of the analysed transect. A similar approach has been used in Drago *et al.* (2022) to evaluate the importance of single nodes in a random forest algorithm.

Sensitivity analysis

Sensitivity analysis (SA) estimates the importance of an input variable for the model output. We applied a global SA following Sobol (2001) to estimate the importance of each output dimension (D1 - D4) of our AE. In a global sensitivity analysis, all parameters are varied simultaneously, allowing not only the contributions of individual parameters to be assessed, but also the contribution of their interactions to the variability of the model output.

HDBSCANs prediction method performed poorly on the randomly generated inputs for the SA, likely due to the ‘curse of dimensionality’ (Bellman and Dreyfus, 2015). Randomly generated points were unlikely to be close to HDBSCAN’s pre-estimated ‘core points’ and were therefore usually classified as outliers. Thus, we trained a Support Vector Machine (SVM) to separate the clusters identified by HDBSCAN and to predict the randomly alternating data for the SA. SVMs define boundaries between existing clusters by maximizing the distance between the boundary and each adjacent cluster. Because HDBSCAN forms clusters from spatially distinct groups, SVMs are uniquely qualified to assign randomly generated data points to the closest cluster despite the ‘curse of dimensionality’. The SVM was trained in R (Meyer *et al.*, 2023) and the SA was performed in python (Iwanaga *et al.*, 2022).

Data handling was done with R (RCoreTeam, 2020) and the tidyverse packages purrr, tibble, dplyr, ggplot2, rstatix, and tidyr (Wickham *et al.*, 2019).

Results

Model training

Each training epoch took ~ 5 seconds using a graphic card with 768 gpu-cores and we trained the model for 250 epochs. The final training and validation RMSE of our model were $RMSE_{Tr} \sim 0.69$ and $RMSE_{Val} \sim 0.48$ (Figure 3).

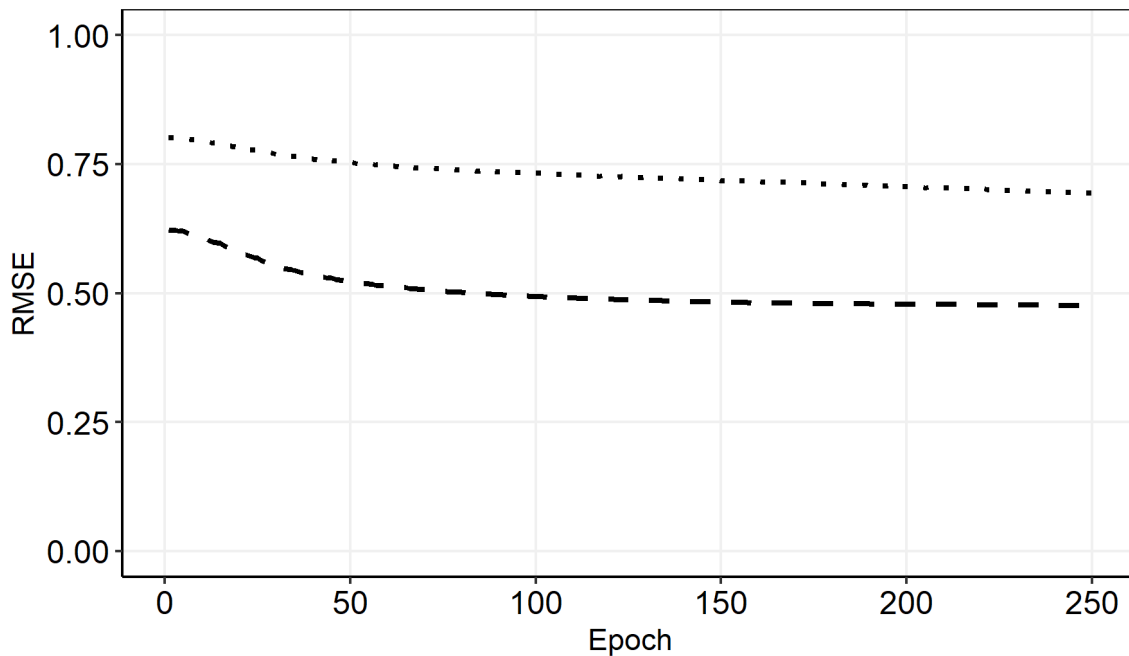


Figure 3: History of the model training. Root-mean-squared error (RMSE - dotted) and validation RMSE (dashed).

Identification of key parameters

Temperature had the highest impact upon the output dimension 1 (D1) with a ~ 46 % share on the final outcome, followed by salinity with ~ 36 % and the euclidean distance (~ 17 %). Variations in oxygen had likely no effect on D1 since the oxygen components were weighted down to a share of only ~ 1 %. D2 was dominated by the euclidean distance (~ 45 %). Temperature and salinity had an equal share of ~ 23 % and oxygen had again the least influence but with ~ 9 % more than on D1. D3 was affected by all input variables similarly with shares between 19 - 29 %. D4 was most sensible towards salinity (~ 39 %) followed by temperature (~ 27 %), oxygen (~ 20 %) and finally the euclidean distance (~ 14 %). Temperature was twice as influential on D1 than on any other output dimension. Variations in oxygen had nearly no effect in D1 but increasingly so on D2, D4 and D3. Overall, salinity was the most influential variable but was never as dominant as temperature or euclidean distance on a single output dimension. Similar to temperature the euclidean distance was twice as influential on D2 as on any other output dimension.

Table 1: Influence of each input variable upon an output variable. Output variables (dimensions) in rows and input variables in columns.

Dimension/Variable	Temperature	Oxygen	Salinity	Euclidean
D1	0.462	0.007	0.359	0.171
D2	0.234	0.091	0.225	0.449
D3	0.194	0.288	0.291	0.228
D4	0.273	0.197	0.386	0.144

Sensitivity

The most sensitive dimension was D3 in more than half of the cases (Figure 4). Conversely, D4 was the least sensitive dimension in almost 50% of all cases. D1 and D2 were similarly important, with both rank 1 and rank 4 more often on the side of D1, leaving D2 as the second or third most important dimension in most cases (~ 73 %).

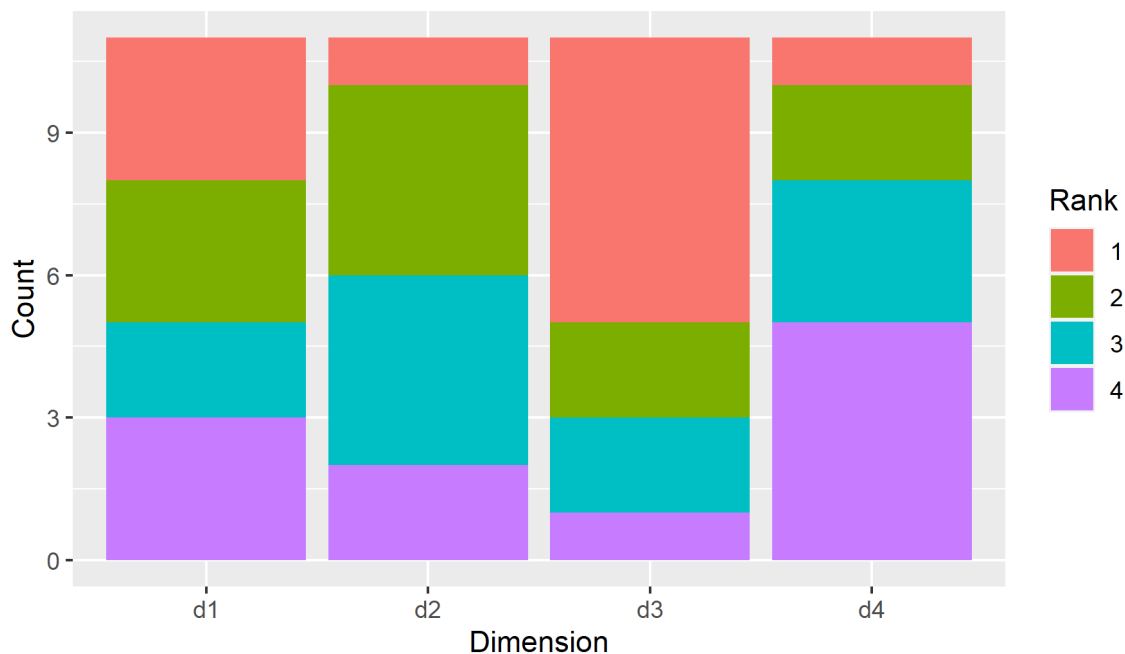


Figure 4: Count of Sobol sensitivity rankings for all transects by dimension. Fixing the dimension with the highest rank (1) would result in the greatest average reduction in variability in output.

Both transects of April cruises (HE353 and HE397), which exhibited fully mixed water columns, were most sensitive towards D1 and thus temperature and to a slightly

lower extend also salinity. The Sobol sensitivity was $ST = 0.86$ for HE353 and $ST = 0.95$ for HE397. Without interactions, D1 would have been the only variable with any significant effect. The temperature range for HE353 was $5.37 - 5.63^{\circ}\text{C}$ (55 km transect) and that of HE397 was $2.69 - 3.34^{\circ}\text{C}$ (100 km transect).

Characterisation of macro-habitats

The habitat map of HE466 T5 (Figure 5) features everything one could expect from a tidal mixing front as it was described in Hill *et al.* (1993). At the beginning of the transect is a stratified water column. The surface layer (GMH3) is outlined by the $240 \mu\text{mol}\cdot\text{l}^{-1}$ isoline for oxygen, with higher concentrations within the GMH. The bottom layer (GMH8) is characterized by temperatures below 12.25°C . GMH16 probably resembles the area of density-driven circulation and is framed by the 13°C isoline for temperature. Towards the end of the transect (GMH5) there is a fully mixed water column with salinity below 33.95 psu. The area left out by those isolines belongs to GMH1, which shows the typical characteristics of an along front jet, namely reduced temperatures and upward doming of bottom front isolines.

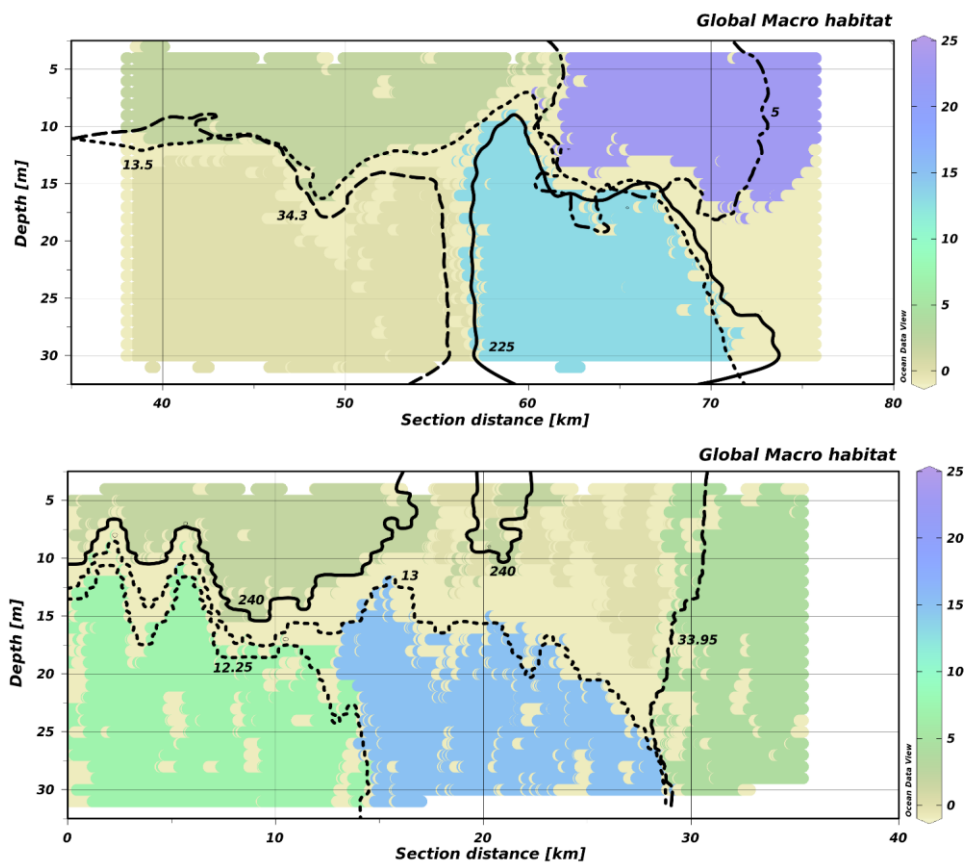


Figure 5: Habitat maps for T1 and T5 of haul 5 from HE466. Dotted: temperature isolines [$^{\circ}\text{C}$], dashed: salinity isolines [psu], solid: oxygen isolines [$\mu\text{mol}\cdot\text{l}^{-1}$], dot-dashed: pluteus larvae abundance [ind / micro-habitat]. Each global macro-habitat is represented by a different colour. Yellow areas denote transition areas (originally labeled -1).

At the beginning of transect T1 there is a stratified water column with the surface layer this time limited to temperatures above 13.5°C (GMH3). The bottom layer (GMH1) is outlined by the 34.3 psu salinity isoline, with higher salinities within the GMH. In this case, the area of the density-driven circulation (GMH14) is characterized by oxygen concentrations below $225\ \mu\text{mol}\cdot\text{l}^{-1}$. There is no fully mixed water column towards the end of the transect, however what is clustered into GMH24 has mostly pluteus larvae densities above 5 individuals per micro-habitat ($\text{ind}\cdot\text{mh}^{-1}$). While one could assume that GMH24 is the along front jet as in T5, the temperatures do not fully support this (Figure 6). However, it is possible that the transect did not fully cross the tidal mixing front and this could affect the predictive ability of the model. As with GMH1 in T5, GMH24 in T1 shows the highest densities of copepods and pluteus larvae, indicating an accumulation of particles as would be expected by an along front jet. However, T1 also shows strong evidence of an upwelling-downwelling dipole (Floeter *et al.*, 2022), with the upwelling pole aligned closely with the surface temperature minimum.

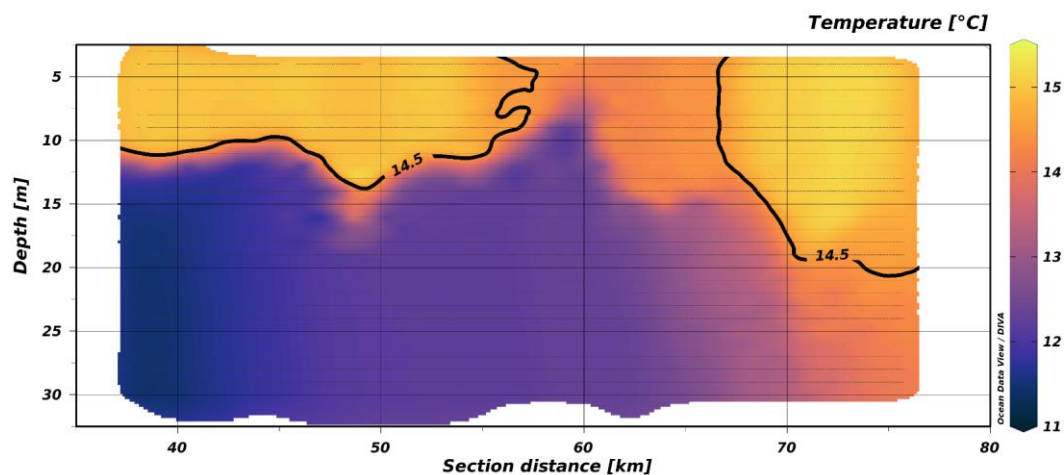


Figure 6: Map of temperature for HE466 haul 5 T1.

As can be seen in Figure 5, the boundaries of global macro-habitats tend to fit the observed parameter isolines very well. It is noticeable that different parameters contribute to the respective isolines, in T1 there is even a GMH that is most closely related to the increased abundance of pluteus larvae. Despite this fairly accurate segregation of potential different habitats, copepods, appendicularians and pluteus larvae tend to share the same MH where they occur at maximum abundance within a transect. However, only in one case the MH are from the same GMH, even if the GMH occurs in several transects, such as GMH3 in T1, T2, T3, and T5.

Discussion

Model inputs

We used values normalized within a transect as model inputs. Thus, peak temperatures in April were treated equally to peak temperatures in August, even though they were naturally much lower. A possible solution would be to use an overall normalisation. In this case, however, we lose the within-transect contrast because the April cruises would have had too narrow an input range, as the initial temperature variance was much smaller than later cruises.

The model had difficulty interpreting plankton abundance, most likely due to the skewness of the data. The patchiness of plankton produces many zeros and only relatively few high abundances. Using the Euclidean distance had two main advantages. First, the distance was no longer biased towards zero. Secondly, the euclidean distance combined the data on the resident plankton communities, using chlorophyll a as a proxy for phytoplankton abundance, into one variable. Thus, the input for the model did not vary between different regions which had different communities as was the case for our SCANFISH data for example. During the SCANFISH cruises, no plankton abundances were available which reduced the 'community' information to just chlorophyll a. The mahalanobis distance (Mahalanobis, 1936) might be an even better distance measure for this purpose (Beaugrand *et al.*, 2011). Unfortunately, it requires at least two dimensional inputs which was not the case for the SCANFISH data.

Global macro-habitats

A drawback of HDBSCAN was the inconsistency of macro-habitat labels across transects. However, following the logic that similar micro-habitats form a single macro-habitat, the COGs of similar macro-habitats should be close to each other even across transects. Thus, we used clustered the COGs to find types of similar macro-habitats across transects. Since the macro-habitats of a single transect were selected under the prejudice of forming the best possible clustering among the tested variations, they should ideally only be grouped with clusters from other transects. The fact that some COGs were assigned the same type despite originating from the same transect indicated an inferior solution by the previous clustering step which formed the local macro-habitats. However, even though only three parameters, minimum sample size, minimum cluster size, and epsilon, were involved in finding different sets of local macro-habitats, it was quite time-consuming to calculate the silhouette scores even within the parameter space we routinely sampled. Therefore, it was quite possible that we found at a local optimum instead of a global one. Logically, the search for global macro-habitats reduced the parameter space from three to one. First, the minimum number of samples within a cluster was set to 2, since it was likely that some types of macro-habitats were less common than others and could thus also occur only once or twice. Local macro-habitats that occurred only once were declared as outlier during the COG clustering. Second, if minimum cluster size was two the number of required neighbors had to be 1 by consequence. Therefore, only Epsilon (size of the neighborhood) was left to vary. Additionally, clustering ~ 100 COGs was computationally much less expensive compared to clustering multiple thousand micro-habitats. We therefore accepted global macro-habitats as the global optimum and rejected local macro-habitats as the inferior local optimum in our loss-surface in cases of conflict, e.g. when COGs of local macro-habitats on the same transect were grouped within the same global macro-habitat.

Key factors behind habitat segregations

A fully mixed water column along an entire transect was only present in the two April cruises HE353 and HE397 (Supplementary Figure S 1, Supplementary Figure S 2). For both cruises the local macro-habitats exhibited a high sensitivity to AE output dimension 1 which had temperature as key parameter. Most of the other transects with stratified water columns were instead most sensitive to AE output dimension 3 (6

out of 9 cases), with salinity as the main key parameter. Temperature had a strong influence on a vertical segregation, which was also observed in Plonus *et al.* (2021b). Salinity was likely stronger involved in a horizontal segregation. Neither oxygen nor chlorophyll displayed any general patterns. However, in the mixed water columns, the temperature had mainly a horizontal gradient. Therefore, the most influential dimensions tended to be those that best reflected the horizontal gradients. This is consistent with previous findings that including horizontal gradients improved the predictive power of a distribution model (Alvarez-Berastegui *et al.*, 2014). Important traits defining the realized niche of a species will probably strongly react to physical gradients (McGill *et al.*, 2006). Therefore, it was to be expected that the dimension that best reflected the horizontal gradients should be the most important one in the prediction of the potential habitats.

Bio-physical characteristics of macro-habitats

The habitat maps showed the results of a variety of different processes. The transects in HE308 (Supplementary Figure S 1) and HE325 (Supplementary Figure S 2) crossed a salinity or plume front. The habitat map of HE308 featured distinct surface and bottom habitats towards the deeper parts of the transect, whereas the habitat map of HE325 generally showed a fully mixed water column on both sides of the front. The most obvious explanation for this difference was the time offset between the two cruises. While HE308 took place in August (2009), HE325 took place already in May (2010). The GMHs of the two transects appeared to be confusing at first sight, but despite the fact that both maps showed a salinity front, the situations that generated these fronts were very different. The haline front in HE325 was based on land-based run-off since it was much closer to the shore than the one in HE308 (Krause *et al.*, 1986). GMH15 represented the area of surface convergence which is usually associated with a tidal intrusion front (Simpson and Nunes, 1981). The front itself (GMH2) consisted of cold, oxygen-rich water. The haline front in HE308 was further offshore and depicted the boundary between North Sea and Atlantic water. In this case GMH15 represented the bottom water immediately to the front, but as in HE325 it was an area of reduced oxygen concentration. The relatively high chlorophyll a concentrations in nearby GMH1 indicated an ongoing subsurface bloom. Such subsurface blooms occur in the vicinity of bottom fronts in the North Sea at depth > 25m and fuel the pelagic production during summer (Richardson *et al.*,

1998; Richardson *et al.*, 2000; Jónasdóttir and Koski, 2011). Dead particles were probably transported towards the front and caused the oxygen minimum by decomposition processes.

Transects T1 to T6 of HE466 (Supplementary Figure S 5 - Supplementary Figure S 10) were not only in the immediate vicinity of an offshore wind farm (OWF), but partly also crossed a tidal mixing front (Hill *et al.*, 1993). While in T1 - T3 all GMHs had mean salinities >34 psu, there was a decreasing trend in salinity in the transects T4 - T6, indicating a shift from North Sea to coastal water (Lee, 1980). Only T5 had a MH that spanned the entire water column with a mean salinity <34 psu, indicating a fully mixed habitat with characteristics of coastal water. T4 and T6 did not extend into the mixed coastal water. However, all three transects had a similar MH (GMHs 7, 16 & 17) with increased temperature (~ .8 °C) and decreased salinity (~ .2 psu) compared to the adjacent bottom habitat. The MH directly above those three had always the highest plankton abundances (GMHs 1, 4 & 15), indicating an aggregation of particles in this specific area. Tidal mixing fronts are basically slowly rotating, near-geostrophic currents (Hill *et al.*, 1993). The rotation creates an upwelling effect on the mixed coastal side of the front, which leads to surface convergence on the stratified offshore side and hence particle aggregation along the front (Hill *et al.*, 1993). It is therefore reasonable that the surface habitats resembled the area of convergence with increased plankton abundance and the bottom habitats the area of the frontal jet.

Transects T1 - T3 showed an upwelling-downwelling dipole (Floeter *et al.*, 2022). Plankton aggregations similar to those in T4 - T6 were located adjacent to the upwelling pole on the coastal side. The upwelling pole, which was perpendicular to the tidal currents, acted as a barrier, accumulating particles like the convergence zone of the tidal mixing front. While the salinity was not different between the MHs in T1 - T3, there was a distinct oxygen pattern with reduced oxygen concentrations in GMH 14 and 16 in the bottom oriented MHs. While the bottom habitats in T1 and T3 were correlated with the upwelling and downwelling poles respectively, the area of reduced oxygen concentrations in T2 encompassed the entire dipole. It is therefore unlikely that the presence of the dipole caused the observed differences between the MHs, and its main effect was to change the mixed layer depth.

Although very similar to T1 and T2 in terms of temperature and salinity, T3 had lower oxygen concentrations leading to the closer relationship with T5. T4 was closest to OWF Global Tech I and had little similarity to the other transects. Apart from T4, the surface habitat was always GMH3, while there was a shift at the bottom from GMH1/14 to GMH8/16. Thus, while the surface habitat remained the same across all transects, there was a change in the bottom waters in a northwesterly direction parallel to our transects. T6, like T4, had no GMHs in common with the other transects despite a similar parameter range. The main difference was in the oxygen concentrations, indicating an effect of the downwelling pole on T1 and T2.

Plankton niches

The low taxonomic resolution of the VPR makes it impossible to distinguish morphological similar species that might have entirely different traits and distributions (Martin *et al.*, 2021). For example, there are 31 species known to have pluteus larvae in the North Sea (Laakmann *et al.*, 2016). However, *Echinocardium cordatum* Pennant, 1777 is the most abundant one, probably due to higher winter temperatures (Kirby *et al.*, 2007), complemented by *Amphiura* spp. (Lindley *et al.*, 1995). Since the 1980s, pluteus larvae have become even more abundant than copepods at times (Lindley *et al.*, 1995; Kirby *et al.*, 2007).

An extensive overview of copepod species in the North Sea was provided by Fransz *et al.* (1991). Even though there are a number of different species, 4 of them provide 85% of the biomass, namely *Acartia clausi* Giesbrecht, 1889, *Centropages hamatus* Liljeborg, 1853, *Temora longicornis* Müller, 1785, and *Pseudocalanus elongatus* Boeck, 1865 (Hickel, 1975; Beaugrand *et al.*, 2002a). The by far most abundant one is *P. elongatus* (Fransz *et al.*, 1991). None of these species showed any correlation with hydrographic conditions, suggesting that they all coped quite well across the North Sea (Hickel, 1975). Two other copepod species are common in the North Sea, *Calanus helgolandicus* Claus, 1863 and *Calanus finmarchicus* Gunnerus, 1770, the latter being a cold-boreal species that prefers lower temperatures (Planque and Fromentin, 1996). Although restricted to cooler bottom waters, *C. finmarchicus* can contribute significantly to copepod biomass even in summer (Jónasdóttir and Koski, 2011). However, the preferred temperatures of < 9 °C (Jónasdóttir and Koski, 2011) were clearly exceeded in the present sampling area (> 12 °C) and therefore the

occurrence of *C. finmarchicus* is very unlikely (Wilson *et al.*, 2016). The warm-temperate *C. helgolandicus* is of Atlantic origin and, despite its wide temperature tolerance, generally prefers warmer waters (Planque and Fromentin, 1996). A northward shift in the distribution of *C. helgolandicus* has been repeatedly reported, suggesting that it has by now replaced *C. finmarchicus* in the North Sea (Beaugrand *et al.*, 2002b; Beaugrand *et al.*, 2009; Villarino *et al.*, 2015; Wilson *et al.*, 2016), at least during peak summer temperatures.

Second in abundance only to copepods (Landry *et al.*, 1994), appendicularians are of special importance for the vertical particle fluxes in the world's oceans (Winder *et al.*, 2017). Due to their affinity for higher temperatures and tolerance towards more acidic conditions, it is likely that their importance will increase under global warming (Winder *et al.*, 2017). The most abundant species in European coastal waters are *Oikopleura longicauda* Vogt, 1854, *Oikopleura dioica* Fol, 1872, *Oikopleura fusiformis* Fol, 1872, and *Fritillaria borealis* Lohmann, 1896 (Lopez-Urrutia *et al.*, 2005). Although all species tolerate the range of temperatures and salinities observed in this study (Lopez-Urrutia *et al.*, 2005), *O. dioica* dominates the other species at temperatures below 20 °C (Lombard *et al.*, 2010) and *F. borealis* declines in abundance as early as May (Wyatt, 1973).

Despite the great biodiversity of the North Sea, it is therefore likely that the most abundant taxa are dominated by single most common species and that their behavior and their spatial niches can potentially be estimated from VPR images. However, peak plankton abundances for all groups considered were always found in areas of surface convergence, whether or not the area was identified as a distinct habitat. Since plankton is, per definition, subjected to the currents that shape the pelagic environment, the strong velocities associated with upwelling-downwelling dipoles (Broström, 2008) and frontal jets (Hill *et al.*, 1993) have the potential to superimpose behavioral niche segregation (Hidalgo *et al.*, 2012). These results are in accordance with findings of Beaugrand and Ibañez (2002) and Beaugrand *et al.* (2001), that local hydrography is a primary driver of biodiversity. Peak copepod abundances in the central and northern North Sea were usually close to the thermocline and not at the surface as in our transects (Daro, 1988). Furthermore, niche segregation has been reported to be more pronounced under extreme conditions, probably above 35 psu or

below 10-12 °C (Martin *et al.*, 2021), neither of which were met in transects of HE466. Prey selection is another relevant factor for niche diversity (Fransz *et al.*, 1991; Cleary *et al.*, 2016), which is not revealed using VPR images. Thus, niche separation of plankton species probably takes place at much smaller (Mitchell *et al.*, 2008) or bigger (Iglesias-Rodríguez *et al.*, 2002; Barton *et al.*, 2013; Brun *et al.*, 2015; McGinty *et al.*, 2018) scales than observed in this study.

Other studies found a strong link between water masses and zooplankton communities at bigger scales (Beaugrand *et al.*, 2002a; Swalethorp *et al.*, 2015). At this wider scale, our model can help to monitor the shifts in associated species. This information is crucial for the successful management of commercially important species such as cod (Beaugrand and Kirby, 2010; Edwards *et al.*, 2013). In general, the inclusion of plankton species improves predictions of distributions at higher trophic levels using species distribution models (Beaugrand and Kirby, 2010; Friedland *et al.*, 2020). Close monitoring of plankton distributions under global warming is therefore essential. Changes in mixed layer depth (MLD) were reliably detected, indicating the potential to monitor anthropogenic effects as well as naturally occurring frontal systems. This will be of particular interest for the expansion of OWFs in the deeper North Sea or in the Baltic Sea (Christiansen *et al.*, 2022). Although clearly limited at the scales investigated in this study, we have demonstrated the ability of our model to objectively process large amounts of data and produce reasonable habitat maps despite the changing importance of different variables. The method does not require any manual pre-processing and we have shown that, despite the current black box perception, machine learning models allow the influence of different inputs on outputs to be assessed. Fast, autonomous processing is a prerequisite for extracting information from the rapidly growing data available to marine biologists, and will advance our understanding of complex ecosystems.

Acknowledgments

The RV Heincke research cruises were supported by grant numbers AWI_HE308_00, AWI_HE325_00, AWI_HE331_00, AWI_HE353_00, AWI_HE397_00, AWI_HE429_00, AWI_HE466_00, and AWI_HE534_00.

References

- Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., *et al.* 2015 *TensorFlow: Large-scale machine learning on heterogeneous distributed systems*. Available at: <https://www.tensorflow.org/>.
- Alvarez-Berastegui, D., Ciannelli, L., Aparicio-Gonzalez, A., Reglero, P., Hidalgo, M., López-Jurado, J. L., Tintoré, J., *et al.* 2014 'Spatial scale, means and gradients of hydrographic variables define pelagic seascapes of bluefin and bullet tuna spawning distribution', *PLOS ONE*, 9(10), p. e109338. doi: [10.1371/journal.pone.0109338](https://doi.org/10.1371/journal.pone.0109338).
- Atkinson, A., Harmer, R. A., Widdicombe, C. E., McEvoy, A. J., Smyth, T. J., Cummings, D. G., Somerfield, P. J., *et al.* 2015 'Questioning the role of phenology shifts and trophic mismatching in a planktonic food web', *Progress in Oceanography*. (The UK western channel observatory: Integrating pelagic and benthic observations in a shelf sea ecosystem), 137, pp. 498–512. doi: [10.1016/j.pocean.2015.04.023](https://doi.org/10.1016/j.pocean.2015.04.023).
- Barth, A. and Stone, J. 2022 'Comparison of an in situ imaging device and net-based method to study mesozooplankton communities in an oligotrophic system', *Frontiers in Marine Science*, 0. doi: [10.3389/fmars.2022.898057](https://doi.org/10.3389/fmars.2022.898057).
- Barton, A. D., Pershing, A. J., Litchman, E., Record, N. R., Edwards, K. F., Finkel, Z. V., Kjørboe, T., *et al.* 2013 'The biogeography of marine plankton traits', *Ecology Letters*, 16(4), pp. 522–534. doi: [10.1111/ele.12063](https://doi.org/10.1111/ele.12063).
- Basterretxea, G., Font-Muñoz, J. S. and Tuval, I. 2020 'Phytoplankton orientation in a turbulent ocean: A microscale perspective', *Frontiers in Marine Science*, 7. Available at: <https://www.frontiersin.org/articles/10.3389/fmars.2020.00185> (Accessed: 8 November 2022).
- Beaugrand, G. and Ibañez, F. 2002 'Spatial dependence of calanoid copepod diversity in the north atlantic ocean', *Marine Ecology Progress Series*, 232, pp. 197–211. doi: [10.3354/meps232197](https://doi.org/10.3354/meps232197).
- Beaugrand, G., Ibañez, F. and Lindley, J. 2001 'Geographical distribution and seasonal and diel changes in the diversity of calanoid copepods in the north atlantic and north sea', *Marine Ecology-progress Series - MAR ECOL-PROGR SER*, 219, pp. 189–203. doi: [10.3354/meps219189](https://doi.org/10.3354/meps219189).
- Beaugrand, G., Ibañez, F., Lindley, J. A. and Reid, P. C. 2002a 'Diversity of calanoid copepods in the north atlantic and adjacent seas: Species associations and biogeography', *Marine Ecology Progress Series*, 232, pp. 179–195. doi: [10.3354/meps232179](https://doi.org/10.3354/meps232179).
- Beaugrand, G. and Kirby, R. R. 2010 'Climate, plankton and cod', *Global Change Biology*, 16(4), pp. 1268–1280. doi: [10.1111/j.1365-2486.2009.02063.x](https://doi.org/10.1111/j.1365-2486.2009.02063.x).
- Beaugrand, G., Lenoir, S., Ibañez, F. and Manté, C. 2011 'A new model to assess the probability of occurrence of a species, based on presence-only data', *Marine Ecology Progress Series*, 424, pp. 175–190. doi: [10.3354/meps08939](https://doi.org/10.3354/meps08939).
- Beaugrand, G., Luczak, C. and Edwards, M. 2009 'Rapid biogeographical plankton shifts in the north atlantic ocean', *Global Change Biology*, 15(7), pp. 1790–1803. doi: [10.1111/j.1365-2486.2009.01848.x](https://doi.org/10.1111/j.1365-2486.2009.01848.x).

- Beaugrand, G., Mackas, D. and Goberville, E. 2013 'Applying the concept of the ecological niche and a macroecological approach to understand how climate influences zooplankton: Advantages, assumptions, limitations and requirements', *Progress in Oceanography*, 111, pp. 75–90. doi: [10.1016/j.pocean.2012.11.002](https://doi.org/10.1016/j.pocean.2012.11.002).
- Beaugrand, G., Reid, P. C., Ibañez, F., Lindley, J. A. and Edwards, M. 2002b 'Reorganization of north atlantic marine copepod biodiversity and climate', *Science*, 296(5573), pp. 1692–1694. doi: [10.1126/science.1071329](https://doi.org/10.1126/science.1071329).
- Behrenfeld, M. J., O'Malley, R., Boss, E., Karp-Boss, L. and Mundt, C. 2021 'Phytoplankton biodiversity and the inverted paradox', *ISME Communications*, 1(1), pp. 1–9. doi: [10.1038/s43705-021-00056-6](https://doi.org/10.1038/s43705-021-00056-6).
- Bell, G. 2001 'Neutral macroecology', *Science*, 293(5539), pp. 2413–2418. doi: [10.1126/science.293.5539.2413](https://doi.org/10.1126/science.293.5539.2413).
- Bellman, R. E. and Dreyfus, S. E. 2015 *Applied dynamic programming*. Princeton University Press.
- Bertrand, A., Grados, D., Colas, F., Bertrand, S., Capet, X., Chaigneau, A., Vargas, G., *et al.* 2014 'Broad impacts of fine-scale dynamics on seascape structure from zooplankton to seabirds', *Nature Communications*, 5(1), p. 5239. doi: [10.1038/ncomms6239](https://doi.org/10.1038/ncomms6239).
- Bishop, C. M. 1995 *Neural networks for pattern recognition*. Oxford: Oxford University Press.
- Broström, G. 2008 'On the influence of large wind farms on the upper ocean circulation', *Journal of Marine Systems*, 74(1), pp. 585–591. doi: [10.1016/j.jmarsys.2008.05.001](https://doi.org/10.1016/j.jmarsys.2008.05.001).
- Brun, P., Vogt, M., Payne, M. R., Gruber, N., O'Brien, C. J., Buitenhuis, E. T., Le Quéré, C., *et al.* 2015 'Ecological niches of open ocean phytoplankton taxa', *Limnology and Oceanography*, 60(3), pp. 1020–1038. doi: [10.1002/lno.10074](https://doi.org/10.1002/lno.10074).
- Chollet, F. 2015 *Keras*. GitHub. Available at: <https://github.com/fchollet/keras>.
- Christiansen, N., Daewel, U. and Schrum, C. 2022 'Tidal mitigation of offshore wind wake effects in coastal seas', *Frontiers in Marine Science*, 9. doi: [10.3389/fmars.2022.1006647](https://doi.org/10.3389/fmars.2022.1006647).
- Cleary, A. C., Durbin, E. G., Rynearson, T. A. and Bailey, J. 2016 'Feeding by pseudocalanus copepods in the bering sea: Trophic linkages and a potential mechanism of niche partitioning', *Deep Sea Research Part II: Topical Studies in Oceanography*. (Understanding ecosystem processes in the eastern bering sea IV), 134, pp. 181–189. doi: [10.1016/j.dsr2.2015.04.001](https://doi.org/10.1016/j.dsr2.2015.04.001).
- Colwell, R. K. and Rangel, T. F. 2009 'Hutchinson's duality: The once and future niche', *Proceedings of the National Academy of Sciences*, 106, pp. 19651–19658. doi: [10.1073/pnas.0901650106](https://doi.org/10.1073/pnas.0901650106).
- Daro, M. H. 1988 'Migratory and grazing behavior of copepods and vertical distribution of phytoplankton', *Bulletin of Marine Science*, 43(3), pp. 710–729.
- Dehling, D. M. and Stouffer, D. B. 2018 'Bringing the eltonian niche into functional diversity', *Oikos*, 127(12), pp. 1711–1723. doi: [10.1111/oik.05415](https://doi.org/10.1111/oik.05415).
- Drago, L., Panaïotis, T., Irisson, J.-O., Babin, M., Biard, T., Carlotti, F., Coppola, L., *et al.* 2022 'Global distribution of zooplankton biomass estimated by in situ imaging and machine learning', *Frontiers in Marine Science*,

9. Available at: <https://www.frontiersin.org/articles/10.3389/fmars.2022.894372> (Accessed: 26 October 2022).
- Edwards, M., Bresnan, E., Cook, K., Heath, M., Helaouet, P., Lynam, C., Raine, R., *et al.* 2013 'Impacts of climate change on plankton', *MCCIP Science Review*, 4.
- Elton, C. S. 1927 *Animal ecology*. New York: Macmillan Co.
- Faillottaz, R., Picheral, M., Luo, J. Y., Guigand, C., Cowen, R. K. and Irisson, J.-O. 2016 'Imperfect automatic image classification successfully describes plankton distribution patterns', *Methods in Oceanography*, 15-16, pp. 60–77. doi: [10.1016/j.mio.2016.04.003](https://doi.org/10.1016/j.mio.2016.04.003).
- Floeter, J., Pohlmann, T., Harmer, A. and Möllmann, C. 2022 'Chasing the offshore wind farm wind-wake-induced upwelling/downwelling dipole', *Frontiers in Marine Science*, 9. Available at: <https://www.frontiersin.org/articles/10.3389/fmars.2022.884943> (Accessed: 1 March 2023).
- Franz, H. G., Colebrook, J. M., Gamble, J. C. and Krause, M. 1991 'The zooplankton of the north sea', *Netherlands Journal of Sea Research*, 28(1), pp. 1–52. doi: [10.1016/0077-7579\(91\)90003-J](https://doi.org/10.1016/0077-7579(91)90003-J).
- Friedland, K. D., Bachman, M., Davies, A., Frelat, R., McManus, M. C., Morse, R., Pickens, B. A., *et al.* 2020 'Machine learning highlights the importance of primary and secondary production in determining habitat for marine fish and macroinvertebrates', *Aquatic Conservation: Marine and Freshwater Ecosystems*, n/a. doi: <https://doi.org/10.1002/aqc.3527>.
- Fuhrman, J. A. 2009 'Microbial community structure and its functional implications', *Nature*, 459(7244), pp. 193–199. doi: [10.1038/nature08058](https://doi.org/10.1038/nature08058).
- Grinnell, J. 1917 'The niche-relationships of the california thrasher', *The Auk*, 34(4), pp. 427–433. doi: [10.2307/4072271](https://doi.org/10.2307/4072271).
- Harris, C. R., Millman, K. J., Walt, S. J. van der, Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., *et al.* 2020 'Array programming with NumPy', *Nature*, 585(7825), pp. 357–362. doi: [10.1038/s41586-020-2649-2](https://doi.org/10.1038/s41586-020-2649-2).
- Harris, P. T. and Baker, E. K. 2012 '1 - why map benthic habitats?', in Harris, P. T. and Baker, E. K. (eds) *Seafloor geomorphology as benthic habitat*. London: Elsevier, pp. 3–22. doi: [10.1016/B978-0-12-385140-6.00001-3](https://doi.org/10.1016/B978-0-12-385140-6.00001-3).
- Hays, G., Richardson, A. and Robinson, C. 2005 'Climate change and marine plankton', *Trends in ecology & evolution*, 20, pp. 337–44. doi: [10.1016/j.tree.2005.03.004](https://doi.org/10.1016/j.tree.2005.03.004).
- Hickel, W. 1975 'The mesozooplankton in the wadden sea of sylt (north sea)', *Helgoländer wissenschaftliche Meeresuntersuchungen*, 27(3), pp. 254–262. doi: [10.1007/BF01611694](https://doi.org/10.1007/BF01611694).
- Hidalgo, M., Gusdal, Y., Dingsør, G. E., Hjermmann, D., Ottersen, G., Stige, L. C., Melsom, A., *et al.* 2012 'A combination of hydrodynamical and statistical modelling reveals non-stationary climate effects on fish larvae distributions', *Proceedings of the Royal Society B: Biological Sciences*, 279(1727), pp. 275–283. doi: [10.1098/rspb.2011.0750](https://doi.org/10.1098/rspb.2011.0750).

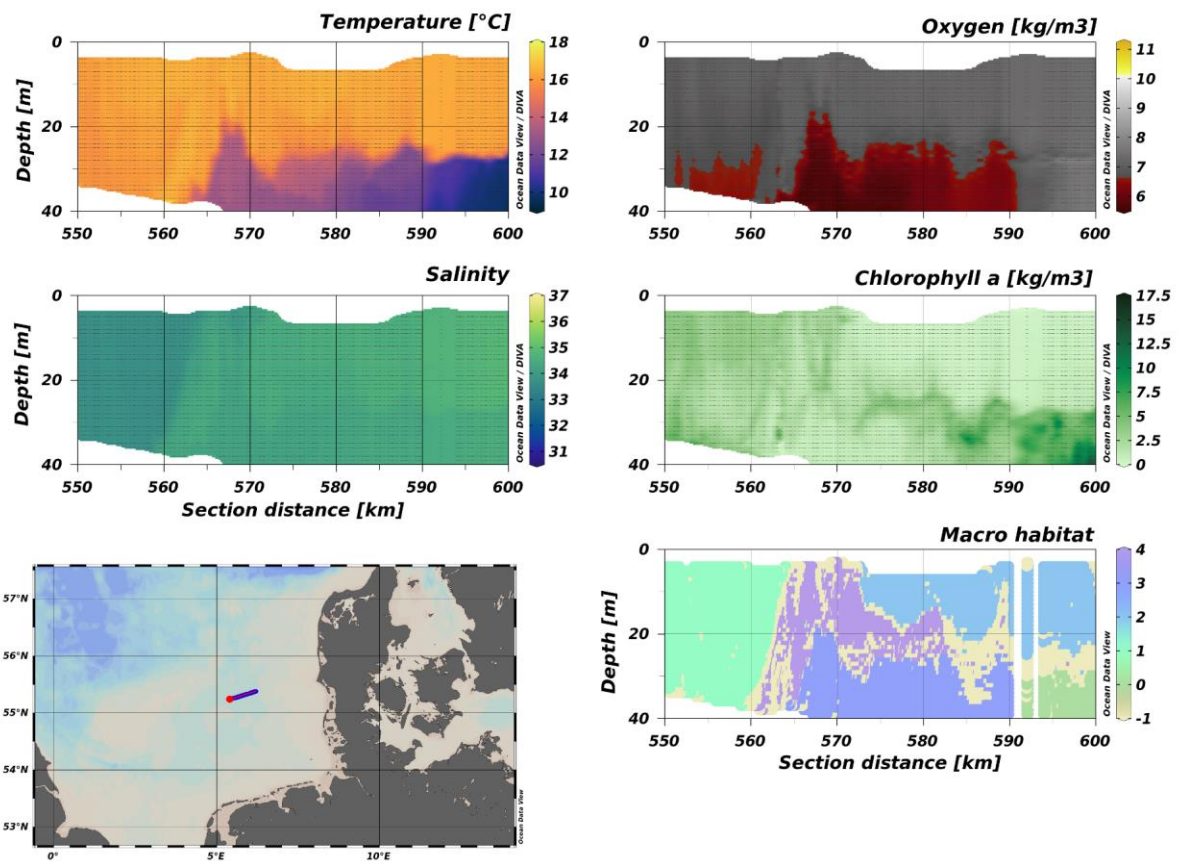
- Hill, A. E., James, I., Linden, P., Matthews, J., Prandle, D., Simpson, J., Gmitrowicz, E., *et al.* 1993 'Dynamics of tidal mixing fronts in the north sea', *Philosophical Transactions of The Royal Society B: Biological Sciences*, 343, pp. 431–446. doi: [10.1098/rsta.1993.0057](https://doi.org/10.1098/rsta.1993.0057).
- Hinchey, E. K., Nicholson, M. C., Zajac, R. N. and Irlandi, E. A. 2008 'Preface: Marine and coastal applications in landscape ecology', *Landscape Ecology*, 23(1), pp. 1–5. doi: [10.1007/s10980-007-9141-3](https://doi.org/10.1007/s10980-007-9141-3).
- Hutchinson, G. E. 1957 'Concluding remarks', *Cold Spring Harbor Symposia on Quantitative Biology*, 22(0), pp. 415–427. doi: [10.1101/SQB.1957.022.01.039](https://doi.org/10.1101/SQB.1957.022.01.039).
- Hutchinson, G. E. 1961 'The paradox of the plankton', *THE AMERICAN NATURALIST*, p. 10.
- Iglesias-Rodríguez, M. D., Brown, C. W., Doney, S. C., Kleypas, J., Kolber, D., Kolber, Z., Hayes, P. K., *et al.* 2002 'Representing key phytoplankton functional groups in ocean carbon cycle models: coccolithophorids', *Global Biogeochemical Cycles*, 16(4), pp. 47-1-47-20. doi: [10.1029/2001GB001454](https://doi.org/10.1029/2001GB001454).
- Irisson, J.-O., Ayata, S.-D., Lindsay, D. J., Karp-Boss, L. and Stemmann, L. 2022 'Machine learning for the study of plankton and marine snow from images', *Annual Review of Marine Science*, 14(1), pp. 277–301. doi: [10.1146/annurev-marine-041921-013023](https://doi.org/10.1146/annurev-marine-041921-013023).
- Iwanaga, T., Usher, W. and Herman, J. 2022 'Toward SALib 2.0: Advancing the accessibility and interpretability of global sensitivity analyses', *Socio-Environmental Systems Modelling*, 4, pp. 18155–18155. doi: [10.18174/sesmo.18155](https://doi.org/10.18174/sesmo.18155).
- Jónasdóttir, S. H. and Koski, M. 2011 'Biological processes in the north sea: Comparison of calanus helgolandicus and calanus finmarchicus vertical distribution and production', *Journal of Plankton Research*, 33(1), pp. 85–103. doi: [10.1093/plankt/fbq085](https://doi.org/10.1093/plankt/fbq085).
- Kirby, R. R., Beaugrand, G., Lindley, J. A., Richardson, A. J., Edwards, M. and Reid, P. C. 2007 'Climate effects and benthic–pelagic coupling in the north sea', *Marine Ecology Progress Series*, 330, pp. 31–38. Available at: <https://www.jstor.org/stable/24871087> (Accessed: 2 March 2023).
- Knust, R., Nixdorf, U. and Hirsekorn, M. 2017 'Research vessel HEINCKE operated by the alfred-wegener-institute', *Journal of large-scale research facilities JLSRF*, 3, pp. A120–A120. doi: [10.17815/jlsrf-3-164](https://doi.org/10.17815/jlsrf-3-164).
- Krause, G., Budeus, G., Gerdes, D., Schaumann, K. and Hesse, K. 1986 'Frontal systems in the german bight and their physical and biological effects', in Nihoul, J. C. J. (ed.) *Elsevier oceanography series*. Elsevier (Marine interfaces ecohydrodynamics), pp. 119–140. doi: [10.1016/S0422-9894\(08\)71042-0](https://doi.org/10.1016/S0422-9894(08)71042-0).
- Laakmann, S., Boos, K., Kneibelsberger, T., Raupach, M. J. and Neumann, H. 2016 'Species identification of echinoderms from the north sea by combining morphology and molecular data', *Helgoland Marine Research*, 70(1), p. 18. doi: [10.1186/s10152-016-0468-5](https://doi.org/10.1186/s10152-016-0468-5).
- Landry, M., Peterson, W. and Fagerness, V. 1994 'Mesozooplankton grazing in the southern california bight. I. Population abundances and gut pigment contents', *Marine Ecology Progress Series*, 115, pp. 55–71. doi: [10.3354/meps115055](https://doi.org/10.3354/meps115055).

- Lapointe, L. and Bourget, E. 1999 'Influence of substratum heterogeneity scales and complexity on a temperate epibenthic marine community', *Marine Ecology Progress Series*, 189, pp. 159–170. Available at: <https://www.jstor.org/stable/24854574> (Accessed: 2 February 2023).
- Lee, A. J. 1980 'Chapter 14 north sea: Physical oceanography', in Banner, F. T., Collins, M. B., and Massie, K. S. (eds) *Elsevier oceanography series*. Elsevier (The north-west european shelf seas: The sea bed and the sea in motion II. Physical and chemical oceanography, and physical resources), pp. 467–493. doi: [10.1016/S0422-9894\(08\)71359-X](https://doi.org/10.1016/S0422-9894(08)71359-X).
- Leibold, M. A. and McPeck, M. A. 2006 'COEXISTENCE OF THE NICHE AND NEUTRAL PERSPECTIVES IN COMMUNITY ECOLOGY', *Ecology*, 87(6), pp. 1399–1410. doi: [10.1890/0012-9658\(2006\)87\[1399:COTNAN\]2.0.CO;2](https://doi.org/10.1890/0012-9658(2006)87[1399:COTNAN]2.0.CO;2).
- Lindley, J. A., Gamble, J. C. and Hunt, H. G. 1995 'A change in the zooplankton of the central north sea (55° to 58° n): A possible consequence of changes in the benthos', *Marine Ecology Progress Series*, 119(1), pp. 299–303. Available at: <https://www.jstor.org/stable/24849819> (Accessed: 1 March 2023).
- Lombard, F., Boss, E., Waite, A. M., Vogt, M., Uitz, J., Stemmann, L., Sosik, H. M., *et al.* 2019 'Globally consistent quantitative observations of planktonic ecosystems', *Frontiers in Marine Science*, 6. Available at: <https://www.frontiersin.org/articles/10.3389/fmars.2019.00196> (Accessed: 8 November 2022).
- Lombard, F., Legendre, L., Picheral, M., Sciandra, A. and Gorsky, G. 2010 'Prediction of ecological niches and carbon export by appendicularians using a new multispecies ecophysiological model', *Marine Ecology Progress Series*, 398, pp. 109–125. doi: [10.3354/meps08273](https://doi.org/10.3354/meps08273).
- Lopez-Urrutia, A., Harris, R., Acuña, J., Båmstedt, Flood, P., Fyhn, Gasser, *et al.* 2005 'A comparison of appendicularian seasonal cycles in four contrasting european coastal environments', in, pp. 255–276.
- MacArthur, R. H. and MacArthur, J. W. 1961 'On bird species diversity', *Ecology*, 42(3), pp. 594–598. doi: [10.2307/1932254](https://doi.org/10.2307/1932254).
- Mahalanobis, P. 1936 'On the generalized distance in statistics', *Journal of Genetics*, 41, pp. 159–193. Available at: http://bayes.acs.unt.edu:8083/BayesContent/class/Jon/MiscDocs/1936_Mahalanobis.pdf (Accessed: 15 November 2022).
- Martin, K., Schmidt, K., Toseland, A., Boulton, C. A., Barry, K., Beszteri, B., Brussaard, C. P. D., *et al.* 2021 'The biogeographic differentiation of algal microbiomes in the upper ocean from pole to pole', *Nature Communications*, 12(1), p. 5483. doi: [10.1038/s41467-021-25646-9](https://doi.org/10.1038/s41467-021-25646-9).
- Matus-Hernández, M. Á., Martínez-Rincón, R. O., Aviña-Hernández, R. J. and Hernández-Saavedra, N. Y. 2019 'Landsat-derived environmental factors to describe habitat preferences and spatiotemporal distribution of phytoplankton', *Ecological Modelling*, 408, p. 108759. doi: [10.1016/j.ecolmodel.2019.108759](https://doi.org/10.1016/j.ecolmodel.2019.108759).
- McGill, B. J., Enquist, B. J., Weiher, E. and Westoby, M. 2006 'Rebuilding community ecology from functional traits', *Trends in Ecology & Evolution*, 21(4), pp. 178–185. doi: [10.1016/j.tree.2006.02.002](https://doi.org/10.1016/j.tree.2006.02.002).
- McGinty, N., Barton, A. D., Record, N. R., Finkel, Z. V. and Irwin, A. J. 2018 'Traits structure copepod niches in the north atlantic and southern ocean', *Marine Ecology Progress Series*, 601, pp. 109–126. doi: [10.3354/meps12660](https://doi.org/10.3354/meps12660).

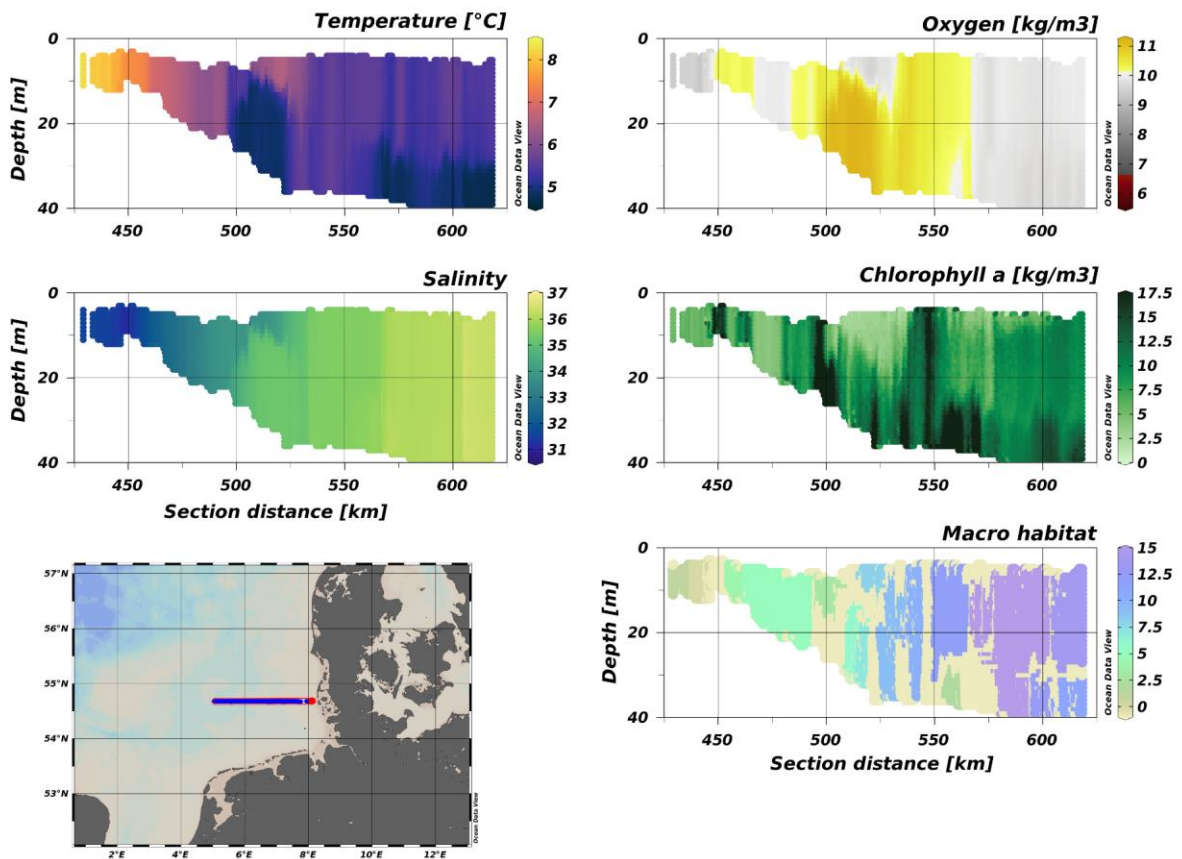
- McInnes, L., Healy, J. and Astels, S. 2017 'Hdbscan: Hierarchical density based clustering', *Journal of Open Source Software*, 2(11), p. 205. doi: [10.21105/joss.00205](https://doi.org/10.21105/joss.00205).
- Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A. and Leisch, F. 2023 'Misc functions of the department of statistics, Probability Theory group (formerly: E1071), TU wien'. Available at: <https://www.semanticscholar.org/paper/Misc-Functions-of-the-Department-of-Statistics%2C-TU-Meyer-Dimitriadou/a0816a1a3a91d22602cdcbfb4d3bb4abb922a5> (Accessed: 17 February 2023).
- Mitchell, J. G., Yamazaki, H., Seuront, L., Wolk, F. and Li, H. 2008 'Phytoplankton patch patterns: Seascape anatomy in a turbulent ocean', *Journal of Marine Systems*, 69(3), pp. 247–253. doi: [10.1016/j.jmarsys.2006.01.019](https://doi.org/10.1016/j.jmarsys.2006.01.019).
- North, R. P., Riethmüller, R. and Baschek, B. 2016 'Detecting small-scale horizontal gradients in the upper ocean using wavelet analysis', *Estuarine, Coastal and Shelf Science*, 180, pp. 221–229. doi: [10.1016/j.ecss.2016.06.031](https://doi.org/10.1016/j.ecss.2016.06.031).
- Planque, B. and Fromentin, J.-M. 1996 'Calanus and environment in the eastern north atlantic. I. Spatial and temporal patterns of c. Finmarchicus and c. helgolandicus', *Marine Ecology Progress Series*, 134(1), pp. 101–109. Available at: <https://www.jstor.org/stable/24856130> (Accessed: 2 March 2023).
- Plonus, R.-M., Conradt, J., Harmer, A., Janßen, S. and Floeter, J. 2021a 'Automatic plankton image classification—can capsules and filters help cope with data set shift?', *Limnology and Oceanography: Methods*, 19(3), pp. 176–195. doi: <https://doi.org/10.1002/lom3.10413>.
- Plonus, R.-M., Vogl, S. and Floeter, J. 2021b 'Automatic segregation of pelagic habitats', *Frontiers in Marine Science*, 8. Available at: <https://www.frontiersin.org/articles/10.3389/fmars.2021.754375> (Accessed: 10 October 2022).
- RCoreTeam 2020 *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Available at: <https://www.R-project.org/>.
- Richardson, K., Nielsen, T., Pedersen, F., Heilmann, J., Løkkegaard, B. and Kaas, H. 1998 'Spatial heterogeneity in the structure of the planktonic food web in the north sea', *Marine Ecology Progress Series*, 168, pp. 197–211. doi: [10.3354/meps168197](https://doi.org/10.3354/meps168197).
- Richardson, K., Visser, A. W. and Pedersen, F. B. 2000 'Subsurface phytoplankton blooms fuel pelagic production in the north sea', *Journal of Plankton Research*, 22(9), pp. 1663–1671. doi: [10.1093/plankt/22.9.1663](https://doi.org/10.1093/plankt/22.9.1663).
- Schlitzer, R. 2020 *Ocean data view*. Available at: <https://odv.awi.de/> (Accessed: 30 March 2021).
- Siegel, D. A. 1998 'Resource competition in a discrete environment: Why are plankton distributions paradoxical?', *Limnology and Oceanography*, 43(6), pp. 1133–1146. doi: [10.4319/lo.1998.43.6.1133](https://doi.org/10.4319/lo.1998.43.6.1133).
- Simpson, J. H. and Nunes, R. A. 1981 'The tidal intrusion front: An estuarine convergence zone', *Estuarine, Coastal and Shelf Science*, 13(3), pp. 257–IN1. doi: [10.1016/S0302-3524\(81\)80024-2](https://doi.org/10.1016/S0302-3524(81)80024-2).

- Sobol, I. M. 2001 'Global sensitivity indices for nonlinear mathematical models and their monte carlo estimates', *Mathematics and Computers in Simulation*. (The second IMACS seminar on monte carlo methods), 55(1), pp. 271–280. doi: [10.1016/S0378-4754\(00\)00270-6](https://doi.org/10.1016/S0378-4754(00)00270-6).
- Swalethorp, R., Malanski, E., Dalgaard Agersted, M., Gissel Nielsen, T. and Munk, P. 2015 'Structuring of zooplankton and fish larvae assemblages in a freshwater-influenced greenlandic fjord: Influence from hydrography and prey availability', *Journal of Plankton Research*, 37(1), pp. 102–119. doi: [10.1093/plankt/fbu099](https://doi.org/10.1093/plankt/fbu099).
- Thompson, A. F., Lazar, A., Buckingham, C., Naveira Garabato, A. C., Damerell, G. M. and Heywood, K. J. 2016 'Open-ocean submesoscale motions: A full seasonal cycle of mixed layer instabilities from gliders', *Journal of Physical Oceanography*, 46(4), pp. 1285–1307. doi: [10.1175/JPO-D-15-0170.1](https://doi.org/10.1175/JPO-D-15-0170.1).
- Troupin, C., Barth, A., Sirjacobs, D., Ouberdous, M., Brankart, J.-M., Brasseur, P., Rixen, M., *et al.* 2012 'Generation of analysis and consistent error fields using the data interpolating variational analysis (DIVA)', *Ocean Modelling*, 52-53, pp. 90–101. doi: [10.1016/j.ocemod.2012.05.002](https://doi.org/10.1016/j.ocemod.2012.05.002).
- Van Rossum, G. and Drake, F. L. 2009 *Python 3 reference manual*. Scotts Valley, CA: CreateSpace.
- Villarino, E., Chust, G., Licandro, P., Butenschön, M., Ibaibarriaga, L., Larrañaga, A. and Irigoien, X. 2015 'Modelling the future biogeography of north atlantic zooplankton communities in response to climate change', *Marine Ecology Progress Series*, 531, pp. 121–142. doi: [10.3354/meps11299](https://doi.org/10.3354/meps11299).
- Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., *et al.* 2019 'Welcome to the tidyverse', *Journal of Open Source Software*, 4(43), p. 1686. doi: [10.21105/joss.01686](https://doi.org/10.21105/joss.01686).
- Wilson, R. J., Heath, M. R. and Speirs, D. C. 2016 'Spatial modeling of calanus finmarchicus and calanus helgolandicus: Parameter differences explain differences in biogeography', *Frontiers in Marine Science*, 3. Available at: <https://www.frontiersin.org/articles/10.3389/fmars.2016.00157> (Accessed: 27 March 2023).
- Winder, M., Bouquet, J., Rafael Bermúdez, J., Berger, S. A., Hansen, T., Brandes, J., Sazhin, A. F., *et al.* 2017 'Increased appendicularian zooplankton alter carbon cycling under warmer more acidified ocean conditions', *Limnology and Oceanography*, 62(4), pp. 1541–1551. doi: [10.1002/lno.10516](https://doi.org/10.1002/lno.10516).
- Wyatt, T. 1973 'The biology of oikopleura dioica and fritillaria borealis in the southern bight', *Marine Biology*, 22(2), pp. 137–158. doi: [10.1007/BF00391778](https://doi.org/10.1007/BF00391778).
- Zhao, C., Maerz, J., Hofmeister, R., Röttgers, R., Wirtz, K., Riethmüller, R. and Schrum, C. 2019 'Characterizing the vertical distribution of chlorophyll a in the german bight', *Continental Shelf Research*, 175, pp. 127–146. doi: [10.1016/j.csr.2019.01.012](https://doi.org/10.1016/j.csr.2019.01.012).

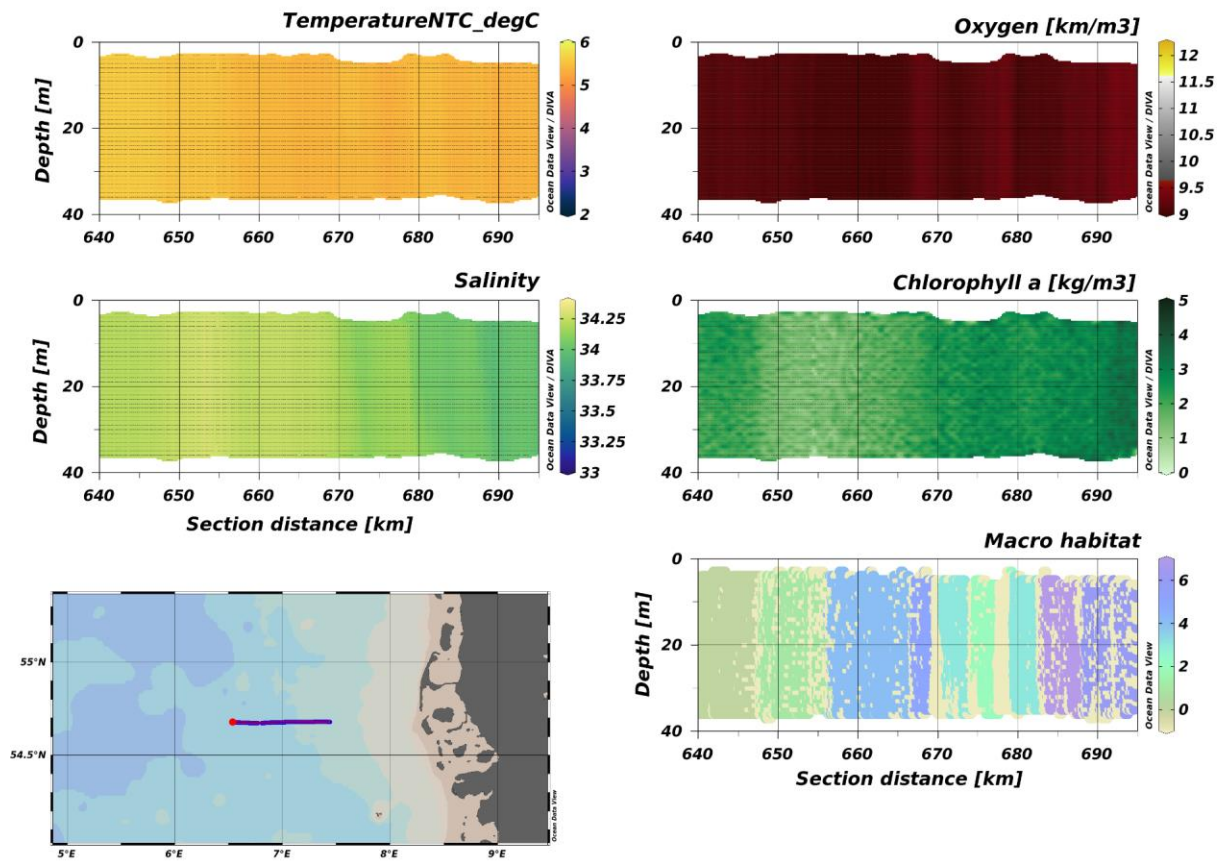
Supplementary Material



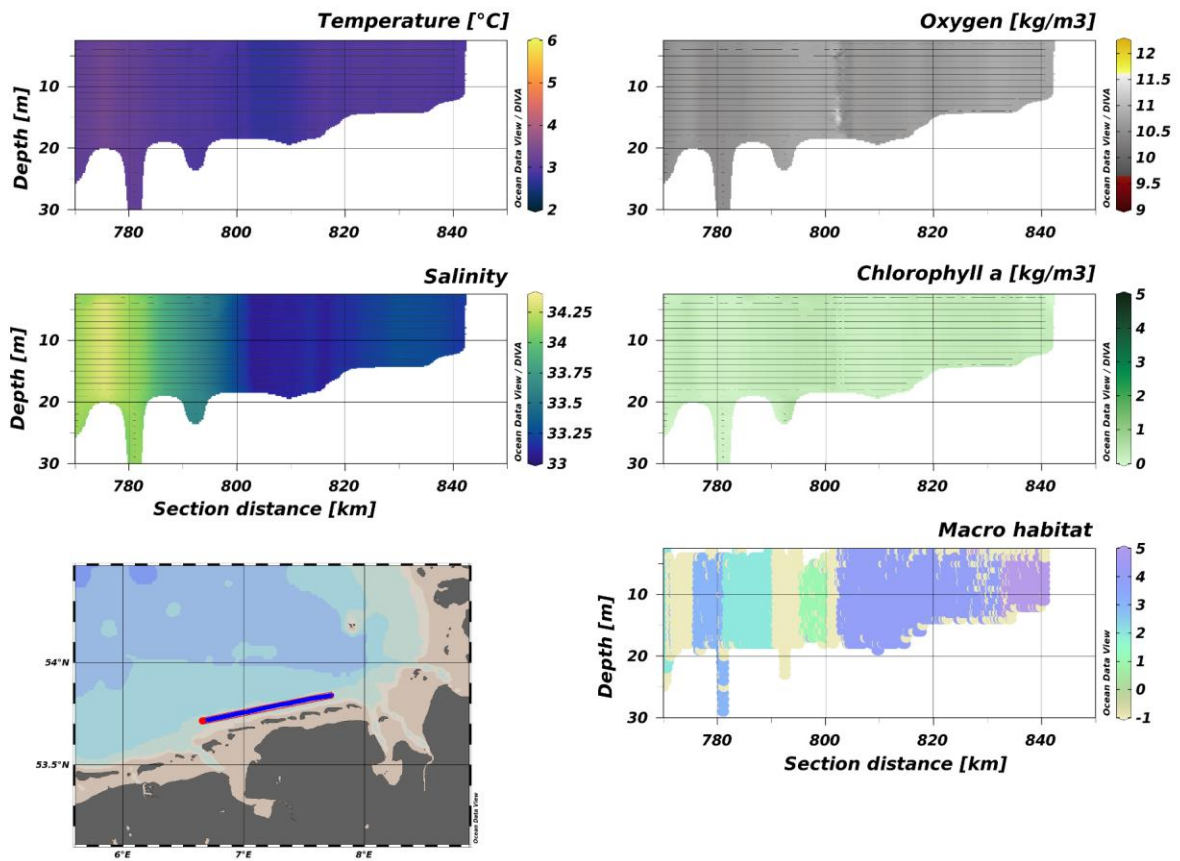
Supplementary Figure S 1: Gridded parameters and habitat map for HE308. The lower left panel shows the location of the transect. For comparison, the parameter ranges are the same as for HE325.



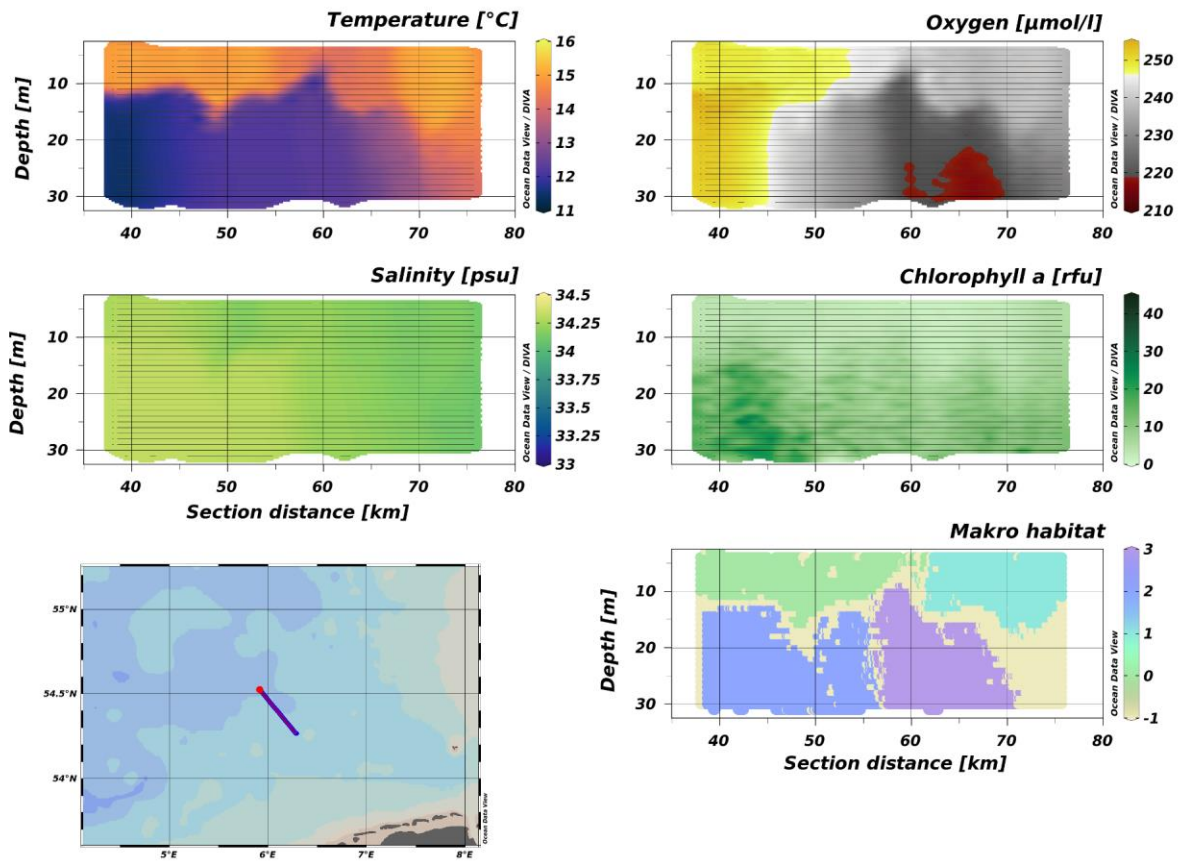
Supplementary Figure S 2: Gridded parameters and habitat map for HE325. The lower left panel shows the location of the transect. For comparison, the parameter ranges are the same as for HE308.



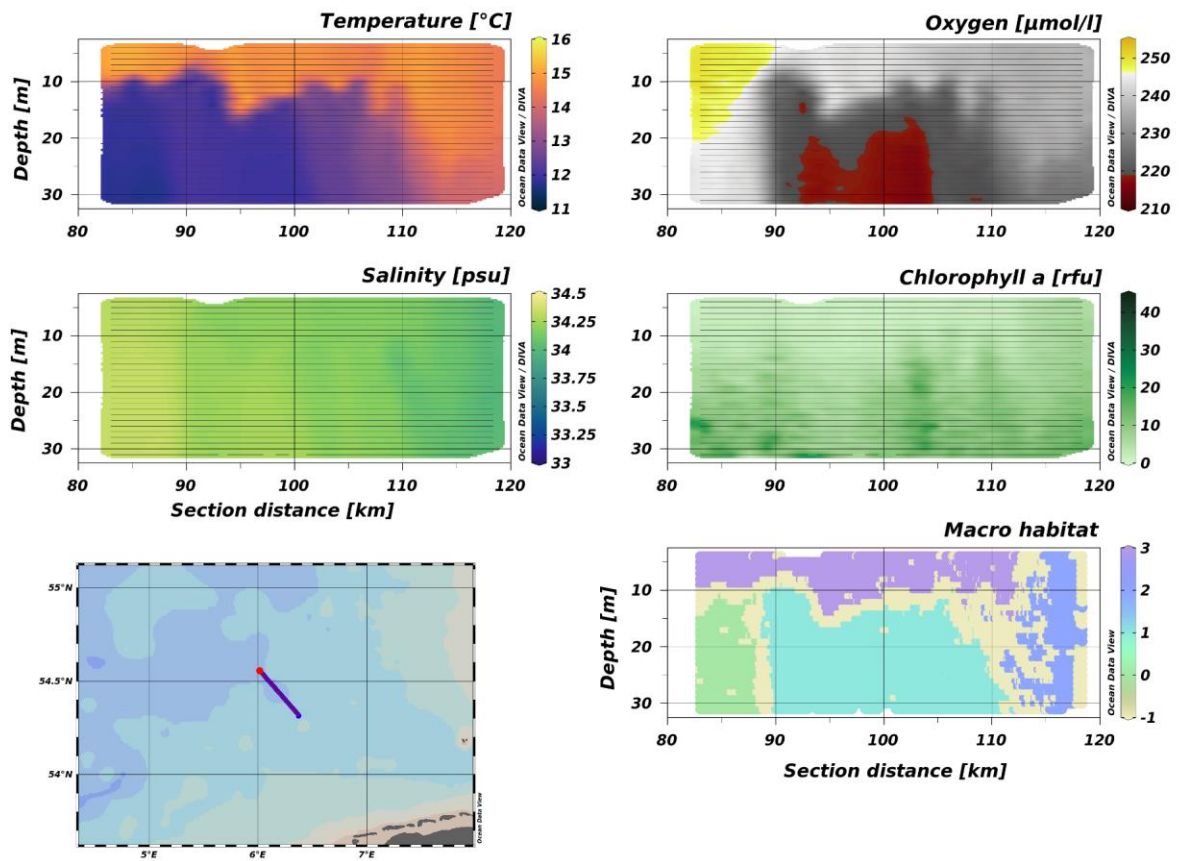
Supplementary Figure S 3: Gridded parameters and habitat map for HE353. The lower left panel shows the location of the transect. For comparison, the parameter ranges are the same as for HE397.



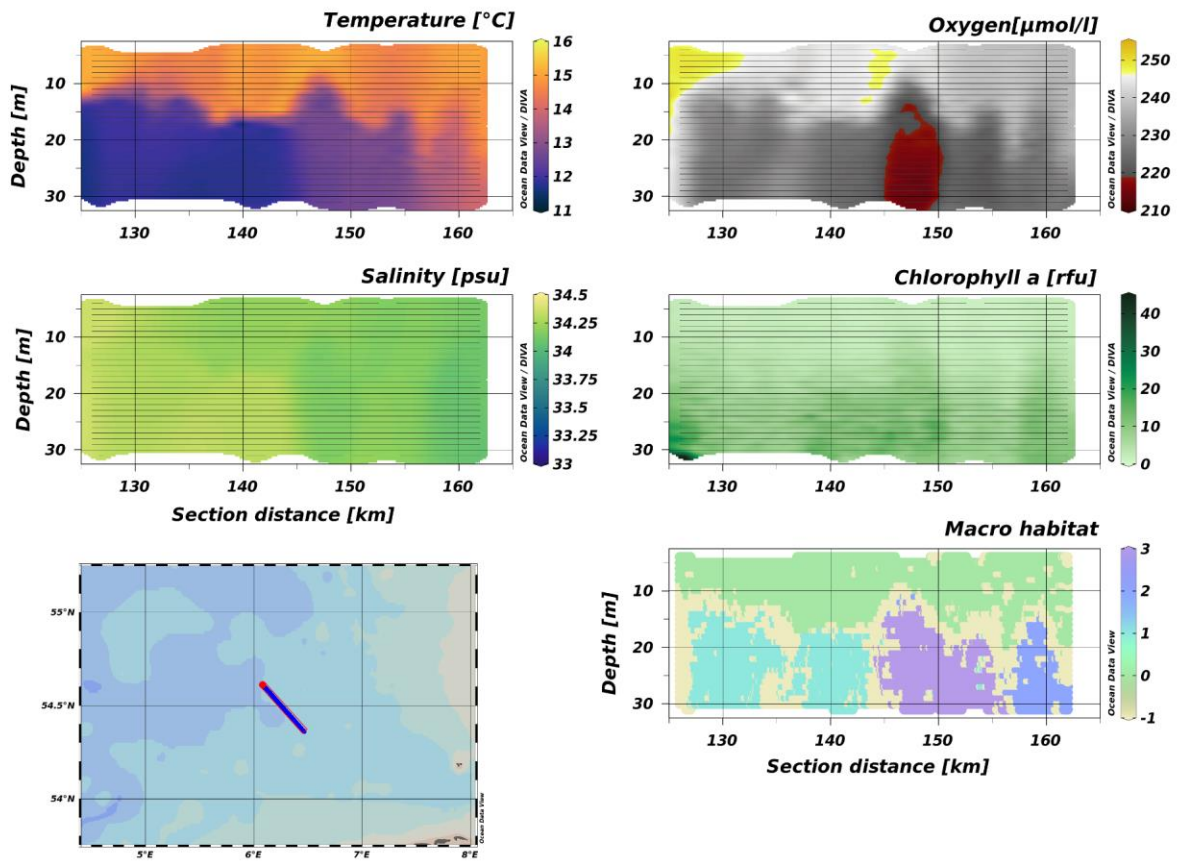
Supplementary Figure S 4: Gridded parameters and habitat map for HE397. The lower left panel shows the location of the transect. For comparison, the parameter ranges are the same as for HE353.



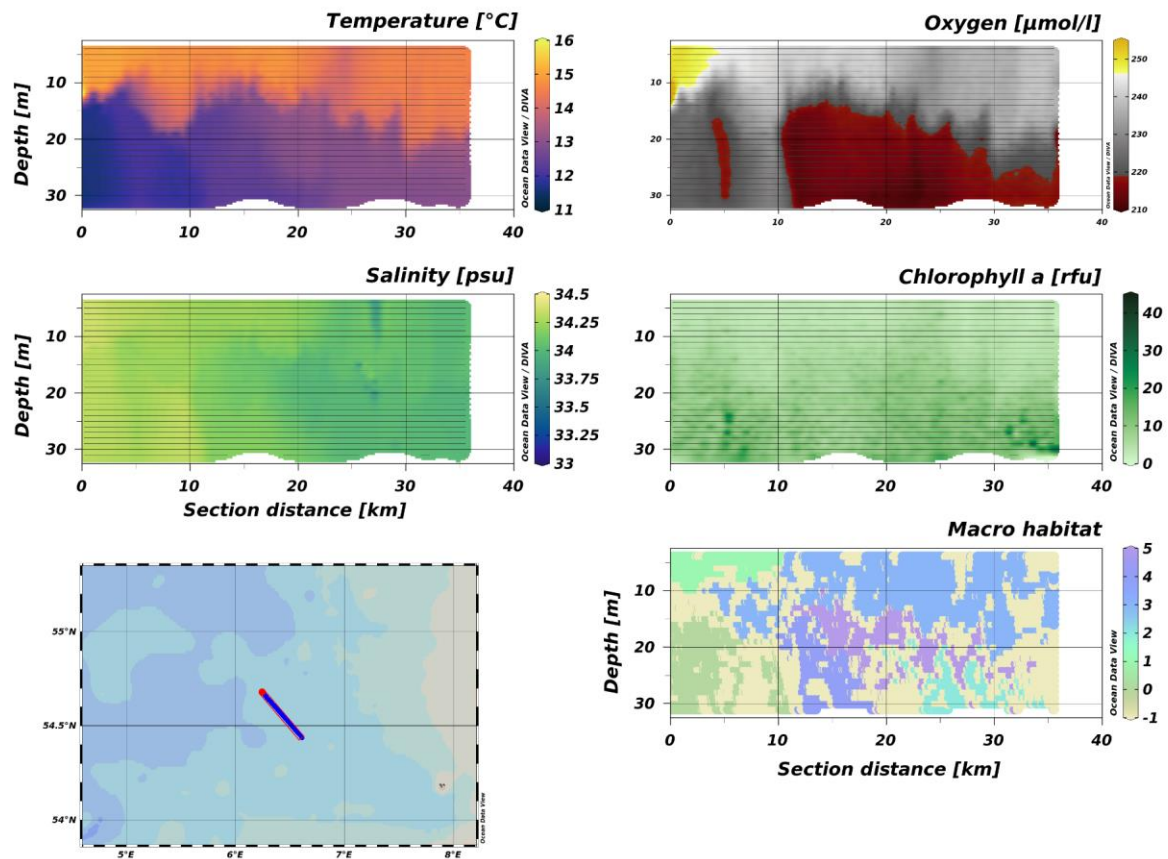
Supplementary Figure S 5: Gridded parameters and habitat map for HE466 T1. The lower left panel shows the location of the transect. The parameter ranges are the same for all transects of HE466.



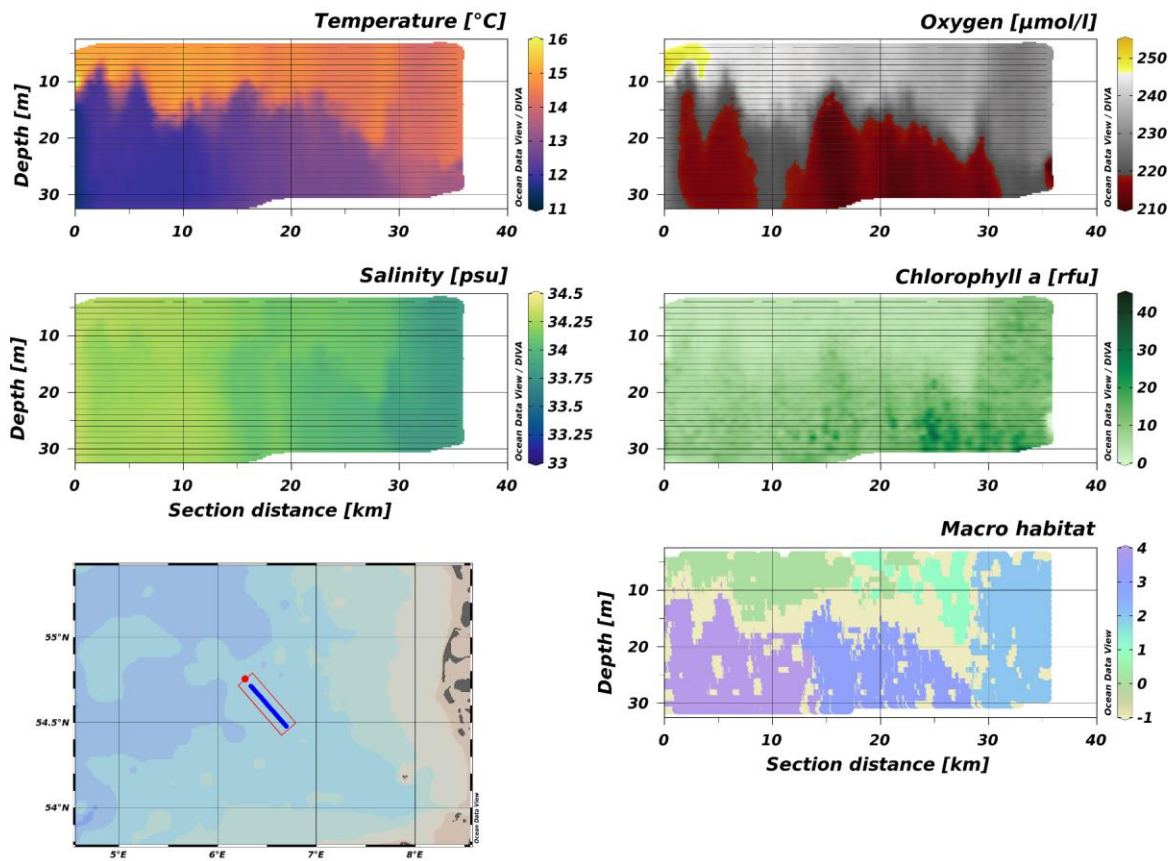
Supplementary Figure S 6: Gridded parameters and habitat map for HE466 T2. The lower left panel shows the location of the transect. The parameter ranges are the same for all transects of HE466.



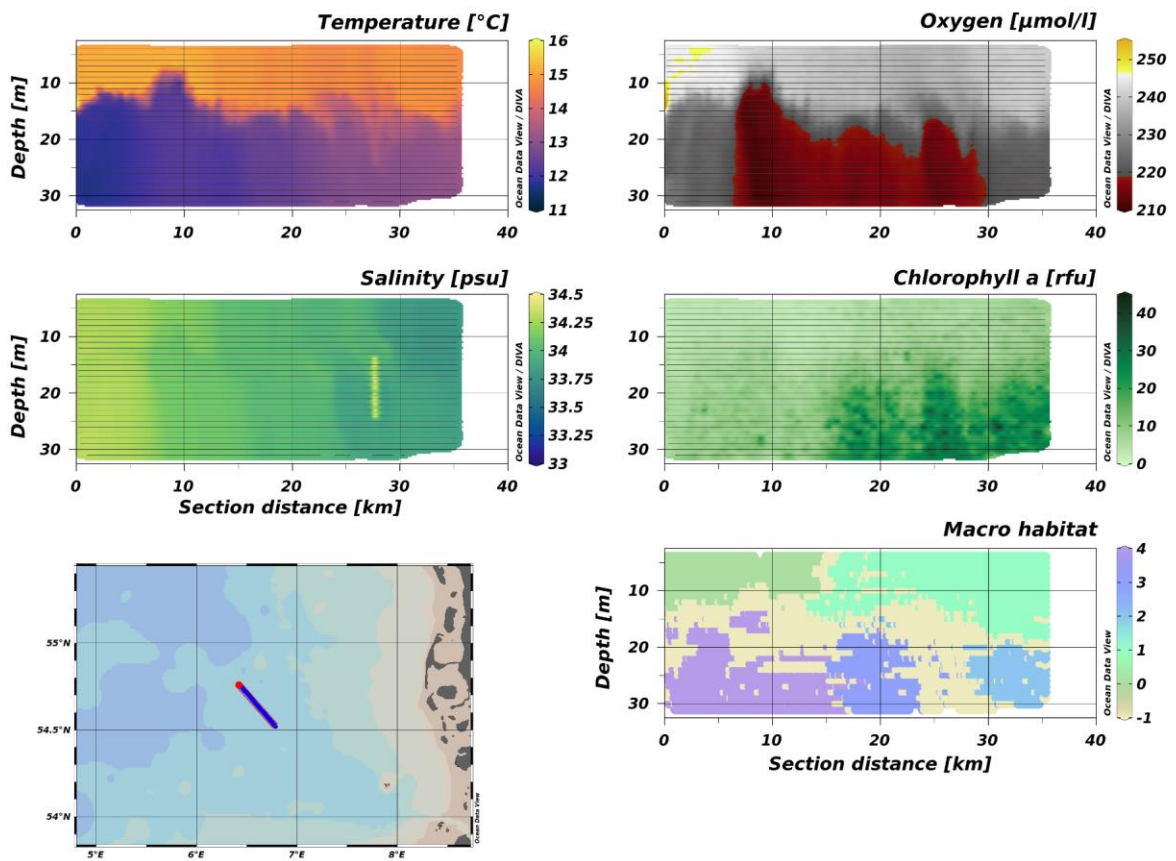
Supplementary Figure S 7: Gridded parameters and habitat map for HE466 T3. The lower left panel shows the location of the transect. The parameter ranges are the same for all transects of HE466.



Supplementary Figure S 8: Gridded parameters and habitat map for HE466 T4. The lower left panel shows the location of the transect. The parameter ranges are the same for all transects of HE466.



Supplementary Figure S 9: Gridded parameters and habitat map for HE466 T5. The lower left panel shows the location of the transect. The parameter ranges are the same for all transects of HE466.



Supplementary Figure S 10: Gridded parameters and habitat map for HE466 T6. The lower left panel shows the location of the transect. The parameter ranges are the same for all transects of HE466.

General Discussion

The role of data set shift

The issue of data set shift (DSS, Moreno-Torres *et al.*, 2012; Webb *et al.*, 2018) arises from the variability of plankton communities and renders classifications of trained models for newly generated field data insufficient. In **chapter I I** investigated the potential benefit of using a new type of model called ‘Capsule Network’ (CapsNet, Hinton *et al.*, 2011; Sabour *et al.*, 2017) to cope with this problem. I could show that the CapsNet was indeed better at maintaining a similar performance level under the prejudice of DDS, but at the drawback of having a generally lower recall. While recall is not a big issue in the more common classes, it is essential for rare classes like fish eggs or larvae (Walker and Orenstein, 2021). For some, if not most, applications, the rare classes are of more interest than the common ones, making the reduced recall of the CapsNet problematic. Our CapsNet might not have reached the full potential since the feedback mechanism of image reconstruction was not fully functional, which could explain the limited capacity of the model. Still, I reached comparable results using a common CNN which also kept a stable precision (> 90 %) under moderate levels of DSS. Although higher than that of the CapsNet to begin with, the recall of the CNN suffered under DSS (- 30 %). This should be kept in mind when comparing different data sets, since a difference in the recall produces the illusion of a different abundance which might not resemble the truth.

How to handle rare species?

The blind spot of machine learning models, and to some extent optical sampling methods, is the detection of rare species (Langenkämper *et al.*, 2019). Those however are of particular importance at times. For example fish eggs and larvae are routinely surveyed to improve stock assessments for some species (Kallasvuo *et al.*, 2017; Ratcliffe *et al.*, 2021). While certainly not perfect, I offer a possible solution to this problem in **chapter I**. What I refer to as the ‘Top-3-Accuracy’-method is designed to significantly reduce the haystack in which to look for the proverbial needle. Rather than a fully automated process, it relies on human knowledge for accuracy, providing only a pre-selected sub-sample of the total data for manual identification of the target classes (Pizarro *et al.*, 2008; Walker and Orenstein, 2021). In this case, I take advantage of the naturally skewed distribution of plankton particles, which is usually

dominated by marine snow particles and a few dominant species like copepods. Those are quite readily classified and can therefore be reliably eliminated. The method reduces handling time by ~90%, and up to ~99% for rare species. While not fully automatic, this would allow to routinely update and monitor a specific species and make use of the data in annual reports and the like.

Is automatic classification sufficient?

The big advantage of optical sampling methods is the reduced handling time when used in concert with machine learning (Pizarro *et al.*, 2008; Schmid *et al.*, 2020). However, the advantage of speed at the expense of accuracy is not progressive. While clearly not of sufficient accuracy to predict the abundance of plankton, our CNN managed to reflect the true distribution of plankton organisms in the field correctly. Therefore, automatic classification is sufficient to analyse the spatial distribution of common taxa and identify agglomerations e.g. during blooming conditions or at physical fronts, but does not allow for biomass estimations or other quantitative analyses (Faillettaz *et al.*, 2016; Luo *et al.*, 2018; Pitois *et al.*, 2018). At least not without time consuming manual validation. This however fits the strengths of the optical sampling, which has a much higher spatial resolution compared to the traditional net-based sampling (Davis *et al.*, 1992; Wiebe and Benfield, 2003; Benfield *et al.*, 2007). Furthermore, abundance estimates derived from different sampling gears are not easily comparable (Pitois *et al.*, 2018) and the value of abundance estimates increases when they are supported by a long term time-series like the 'Helgoland Roads' (Lindegren *et al.*, 2012). Until the scientific community can establish equally long standardized time-series using optical devices there is time to improve the accuracy of machine learning models to the desired level. Additionally, preparing such time series will provide scientists with more and better training data which will help to improve the models as well (Lombard *et al.*, 2019; Conradt *et al.*, 2022).

Another question that remains is: is the taxonomic resolution of the VPR sufficient to capture species-specific behavior? Most optical samplers, and the VPR is no exception, have only limited taxonomic resolution (Batten *et al.*, 2019) and the North Sea is a highly diverse ecosystem (e.g. Fransz *et al.*, 1991; Laakmann *et al.*, 2016). However, previous studies have successfully detected behavioral patterns in

zooplankton (Möller *et al.*, 2015; Faillettaz *et al.*, 2016). As I argued in the introduction, the majority of the biomass is contributed by only a few key species, and therefore species-specific distributions should be evident from the VPR images, at least for those abundant taxa.

What drives plankton spatial distributions?

Since plankton shapes the seascape even for top-predators (Bertrand *et al.*, 2014), it is essential to identify the processes that produce the patchiness of plankton itself (McGinty *et al.*, 2018). With multiple parameters at hand that can affect the distribution of plankton, like salinity or temperature, it can be difficult to disentangle the effect of single variables. Habitat maps link species communities to physically distinct areas (Harris and Baker, 2012) and can help to identify niches in the multi-dimensional space that makes up the environment. The AE in **chapter II** transformed 4-dimensional feature vectors (micro-habitats) into 2-dimensional projections, creating areas of varying density (macro-habitats). Projections of similar macro-habitats were always located in the same general area of the coordinate system, making it possible to infer the physical properties of a macro-habitat from the location of the projections of the micro-habitats. For example, micro-habitats originating from the surface mixed layer were usually projected in the upper half of the coordinate system. In the case of several different macro-habitats in the surface mixed layer, they were all projected in the upper part of the coordinate system, but still formed distinct areas separated by regions with fewer projections. This became particularly relevant in **chapter III**, where I applied HDBSCAN to the ‘centers of gravity’ of each macro-habitat to merge similar macro-habitats across different transects into what I called a ‘global macro-habitat’ (GMH). This allowed us to track local macro-habitats with similar characteristics across all transects and identify the GMH most commonly associated with a particular species or community. I investigated the distribution of copepods, appendicularia and pluteus larvae. Copepods and appendicularia are the most abundant members of the oceanic zooplankton (Landry *et al.*, 1994) and pluteus larvae seasonally dominate the zooplankton in the North Sea since the 1980s (Lindley *et al.*, 1995; Kirby *et al.*, 2007). In the North Sea the three groups are numerically dominated by *P. elongatus* (Fransz *et al.*, 1991), *E. cordatum* (Kirby *et al.*, 2007) and *Oikopleura dioica* Fol, 1872 (Lombard *et al.*, 2010) respectively. It is therefore reasonable to assume that the distribution of the groups reflects the

habitats preferred by the three species. Still, peak abundances of all three groups regularly occurred within the same macro-habitat in each transect. Moreover, the respective macro-habitats did not belong to the same GMH, indicating no clear preference for a particular hydrographic environment. The plankton peaks observed in **chapter III** were associated with either a dipole (Floeter *et al.*, 2022) or a tidal mixing front (Hill *et al.*, 1993). Plankton, by definition not or hardly capable to move against a current (Lombard *et al.*, 2019), is passively aggregated along fronts and eddies (Munk *et al.*, 1995; Hidalgo *et al.*, 2012; Swalethorp *et al.*, 2015; Schmid *et al.*, 2020), such that hydrography has been suggested as a driver of local biodiversity in previous studies (Beaugrand *et al.*, 2001; Beaugrand and Ibañez, 2002; Hidalgo *et al.*, 2012). Thus, the local flow field superimposed behavior and potentially increased competition between the three groups at the sub-mesoscale. Considering the geographical (e.g. along-front) and seasonal distribution of plankton, especially meroplankton such as pluteus or fish larvae, the spawning activity of adult individuals is certainly an important factor (Belgrano *et al.*, 1990; Marteinsdottir *et al.*, 2000; Knutsen *et al.*, 2007; Van Ginderdeuren *et al.*, 2014; Sandø *et al.*, 2020). Another factor influencing the geographical distribution is mortality, either due to unfavorable conditions (Hidalgo *et al.*, 2012) or predation (Behrenfeld *et al.*, 2021). However, the data from all six transects were collected during one day in a geographically small area and therefore do not allow any conclusions to be drawn about these factors.

Climate change is another factor behind plankton spatial distributions which will act differently on individual species (Beaugrand and Ibanez, 2004; Hays *et al.*, 2005; Lindegren *et al.*, 2020). A prominent example in the North Sea are *C. finmarchicus* and *C. helgolandicus* (Planque and Fromentin, 1996; Falkenhaug *et al.*, 2022), but similar trends have been observed around the globe (Beaugrand and Ibanez, 2004). Apart from compositional changes, rising temperatures favor smaller species (Landry *et al.*, 1994; McGinty *et al.*, 2018). Such changes also affect higher trophic levels (Möllmann *et al.*, 2003; Hays *et al.*, 2005; Thorpe *et al.*, 2022). Not only temperature but also acidification will make a difference (Hays *et al.*, 2005). Appendicularia thrive under warmer and more acidic conditions (Winder *et al.*, 2017) and have the potential to replace calanoid copepods in the diet of fish larvae (Alldredge, 1976; Paffenhöfer, 1976; Gorsky and Fenaux, 1998; Winder *et al.*, 2017). However, a change from a copepod-dominated system to one dominated by appendicularia would have

significant implications for the entire ecosystem. Appendicularia provide a shortcut in the microbial loop (Alldredge, 1976; Flood, 1978; Gorsky and Fenaux, 1998), thereby increasing the energy available to higher trophic levels. Short generation times of appendicularia (Gorsky and Fenaux, 1998) and increasing filtering rates at higher temperatures (Alldredge, 1981) allow them to respond quickly to increasing food densities and thus efficiently exploit temporary limited phytoplankton aggregations, e.g. by dipoles, further increasing the impact of OWFs. In addition, discarded houses and fecal pellets are a major source of marine snow and contribute significantly to vertical carbon fluxes (Alldredge *et al.*, 2005; Winder *et al.*, 2017). This would increase the amount of food available to benthic communities and significantly alter the benthic-pelagic coupling. A regime shift in the North Sea plankton community in the late 1980s was associated with increased abundance of the macrobenthos, and species such as *Amphiura filiformis* O. F. Müller, 1776 respond to organic enrichment with increased abundance (Lindley *et al.*, 1995). *E. cordatum* also benefits from rising temperatures (Kirby and Lindley, 2005; Kirby *et al.*, 2007). Seasonal spawning events of meroplanktic pluteus larvae further increase competitive pressure on holoplankton species such as copepods (Fransz *et al.*, 1991; Kirby *et al.*, 2007). Such peaks could become common in the vicinity of OWFs (Inger *et al.*, 2009; Floeter *et al.*, 2017).

High abundances of appendicularia alone have the potential to deplete local phytoplankton contingents (Alldredge, 1981) and thereby increase the competition for food, even though production in itself is supposed to increase with increasing temperatures as well (Thorpe *et al.*, 2022). The current regime in the North Sea is dominated by calanoid copepods, which have shown high sensitivity to climate change (Beaugrand and Ibanez, 2004). Copepods, forced into competition by local hydrography, may decrease in abundance as temperatures increase. This will be even more the case if anthropogenic hydrographic changes caused by OWFs increase the areas of higher competitive stress. Regime shifts are characterized by non-linear, discontinuous dynamics and tend to cause unexpected changes in ecosystem services (Sguotti *et al.*, 2022; Blöcker *et al.*, 2023). Especially in the context of climate change, such changes may become irreversible, with catastrophic ecological, economic and social consequences (Möllmann *et al.*, 2021; Sguotti *et al.*, 2022; Blöcker *et al.*, 2023). Predicting responses to global warming has been

attempted (Thorpe *et al.*, 2022), but such attempts are challenging and precarious, especially considering potential regime shifts (Beaugrand and Ibanez, 2004). One possible precaution to anticipate regime shifts is long-term spatially resolved time series of key organisms (Lindegren *et al.*, 2012). While general considered sentinels to climate change (Hays *et al.*, 2005; Möller *et al.*, 2015; Batten *et al.*, 2019; Drago *et al.*, 2022), the utilities to inform about forthcoming regime shifts differs between species and ecosystems (Lindegren *et al.*, 2012). This highlights the need to develop and establish near real-time monitoring for multiple classes, preferably without human support (Pizarro *et al.*, 2008; Lombard *et al.*, 2019; Walker and Orenstein, 2021). Improved habitat characterization and monitoring can help identify representative areas and thus focus limited resources (Leeuwen *et al.*, 2015). Remote sensing is also a great opportunity for trait-based ecology (McGill *et al.*, 2006). Functional traits are strongly linked to a species' realized niche (McGinty *et al.*, 2018) and potentially improve our understanding of co-occurrences and spatial distributions (McGill *et al.*, 2006; Lindegren *et al.*, 2020). Species with similar traits do not necessarily occupy the same realized niche (Lindegren *et al.*, 2020) and thus the maintenance of vital ecosystem services continues under changing conditions. Therefore, a trait-based approach may be better suited to predict change (McGinty *et al.*, 2018) and identify impending regime shifts. VPR images are not detailed enough to allow trait-based analysis. Nevertheless, the method proposed in **Chapter III** fulfills some of the key requirements for efficient ecosystem monitoring and thus sustainable management of marine resources. The method automatically separates marine pelagic habitats and detects areas of increased plankton abundance. In combination with a VPR, it has the potential to monitor plankton groups in the North Sea and at least detect changes between dominant groups, as suggested in the discussion above, e.g. between copepods and appendicularia. Provided with more details on plankton composition or even traits, it is a promising tool for future ecosystem assessment.

Conclusion and outlook

In summary, this thesis has presented new methods for dealing with high-frequency, multi-dimensional datasets that provide information at high spatial resolution in near real time. I have demonstrated the ability of a machine learning CNN to reliably extract spatial distributions from *in-situ* plankton images of common (and abundant)

groups. Combined with habitat information extracted from a fully unsupervised trained Autoencoder, I used this knowledge to investigate plankton distribution patterns in a highly dynamic ecosystem, the North Sea. Our results add to the evidence that at the sub-mesoscale, plankton spatial distributions in the North Sea are directed by local hydrography. Anthropogenic impacts from offshore wind farms are of limited importance for the current plankton community. However, in combination with global warming, their influence may increase considering a possible regime shift from a copepod dominated system to one dominated by appendicularia. In combination with wind-induced upwelling-downwelling dipoles, which have been shown to aggregate plankton to a similar extent as naturally occurring frontal systems, OWFs have the potential to alter the seascape of higher trophic levels. This, together with the proposed change in the plankton community, is likely to trigger pronounced changes in the whole ecosystem and food web. Such changes could initiate another regime shift in the North Sea, as was previously observed in the late 1980s, the consequences of which are difficult to predict. The proposed method is a valuable tool for monitoring plankton communities as a precautionary approach to detect early warning signals of a regime shift, a prerequisite for sustainable management of commercially important fish stocks. In addition, I have demonstrated how machine learning can help to provide information on the spatial distribution of rare plankton groups, such as fish larvae, in a reasonably short time. Such information can be used to further improve the assessment of commercially exploited stocks and helps to define areas of high vulnerability.

References

- Allredge, A. L. 1976 'Discarded appendicularian houses as sources of food, surface habitats, and particulate organic matter in planktonic environments: Discarded appendicularian houses', *Limnology and Oceanography*, 21(1), pp. 14–24. doi: [10.4319/lo.1976.21.1.0014](https://doi.org/10.4319/lo.1976.21.1.0014).
- Allredge, A. L. 1981 'The impact of appendicularian grazing on natural food concentrations in situ', *Limnology and Oceanography*, 26(2), pp. 247–257. doi: [10.4319/lo.1981.26.2.0247](https://doi.org/10.4319/lo.1981.26.2.0247).
- Allredge, A. L., Gorsky, G., Youngbluth, M. and Deibel, D. 2005 'The contribution of discarded appendicularian houses to the flux of particulate organic carbon from oceanic surface waters', *Response of marine ecosystems to global change: Ecological impact of appendicularians*, pp. 309–326.
- Batten, S. D., Abu-Alhaija, R., Chiba, S., Edwards, M., Graham, G., Jyothibabu, R., Kitchener, J. A., *et al.* 2019 'A global plankton diversity monitoring program', *Frontiers in Marine Science*, 6. doi: [10.3389/fmars.2019.00321](https://doi.org/10.3389/fmars.2019.00321).
- Beaugrand, G. and Ibanez, F. 2004 'Monitoring marine plankton ecosystems. II: Long-term changes in north sea calanoid copepods in relation to hydro-climatic variability', *Marine Ecology Progress Series*, 284, pp. 35–47. doi: [10.3354/meps284035](https://doi.org/10.3354/meps284035).
- Beaugrand, G. and Ibañez, F. 2002 'Spatial dependence of calanoid copepod diversity in the north atlantic ocean', *Marine Ecology Progress Series*, 232, pp. 197–211. doi: [10.3354/meps232197](https://doi.org/10.3354/meps232197).
- Beaugrand, G., Ibañez, F. and Lindley, J. 2001 'Geographical distribution and seasonal and diel changes in the diversity of calanoid copepods in the north atlantic and north sea', *Marine Ecology-progress Series - MAR ECOL-PROGR SER*, 219, pp. 189–203. doi: [10.3354/meps219189](https://doi.org/10.3354/meps219189).
- Behrenfeld, M. J., O'Malley, R., Boss, E., Karp-Boss, L. and Mundt, C. 2021 'Phytoplankton biodiversity and the inverted paradox', *ISME Communications*, 1(1), pp. 1–9. doi: [10.1038/s43705-021-00056-6](https://doi.org/10.1038/s43705-021-00056-6).
- Belgrano, A., Vincx, M., Dewarumez, J.-M., Richard, A., Craeymeersch, J. A. and Heip, C. H. R. 1990 'Recruitment of meroplanktonic larvae in the southern bight of the north sea', *Océanis (Paris)*, (3). Available at: <https://www.vliz.be/en/search-persons?module=ref&refid=7114> (Accessed: 11 April 2023).
- Benfield, M., Grosjean, P., Culverhouse, P., Irigolen, X., Sieracki, M., Lopez-Urrutia, A., Dam, H., *et al.* 2007 'RAPID: Research on automated plankton identification', *Oceanography*, 20(2), pp. 172–187. doi: [10.5670/oceanog.2007.63](https://doi.org/10.5670/oceanog.2007.63).
- Bertrand, A., Grados, D., Colas, F., Bertrand, S., Capet, X., Chaigneau, A., Vargas, G., *et al.* 2014 'Broad impacts of fine-scale dynamics on seascape structure from zooplankton to seabirds', *Nature Communications*, 5(1), p. 5239. doi: [10.1038/ncomms6239](https://doi.org/10.1038/ncomms6239).
- Blöcker, A. M., Gutte, H. M., Bender, R. L., Otto, S. A., Sguotti, C. and Möllmann, C. 2023 'Regime shift dynamics, tipping points and the success of fisheries management', *Scientific Reports*, 13(1), p. 289. doi: [10.1038/s41598-022-27104-y](https://doi.org/10.1038/s41598-022-27104-y).

- Conradt, J., Boerner, G., Lopez-Urrutia, A., Möllmann, C. and Moyano, M. 2022 'Automated plankton classification with a dynamic optimization and adaptation cycle', *Frontiers in Marine Science*, 9, p. 868420. doi: [10.3389/fmars.2022.868420](https://doi.org/10.3389/fmars.2022.868420).
- Davis, C. S., Gallager, S. M., Bermann, M. S., Haury, L. R. and Strickler, J. R. 1992 'The video plankton recorder (VPR): Design and initial results', *Archiv für Hydrobiologie–Beiheft Ergebnisse der Limnologie*, 36, pp. 67–81. Available at: https://www.researchgate.net/publication/284686405_The_Video_Plankton_Recorder_VPR_Design_and_initial_results (Accessed: 5 February 2020).
- Drago, L., Panaiotis, T., Irisson, J.-O., Babin, M., Biard, T., Carlotti, F., Coppola, L., *et al.* 2022 'Global distribution of zooplankton biomass estimated by in situ imaging and machine learning', *Frontiers in Marine Science*, 9. Available at: <https://www.frontiersin.org/articles/10.3389/fmars.2022.894372> (Accessed: 26 October 2022).
- Faillietaz, R., Picheral, M., Luo, J. Y., Guigand, C., Cowen, R. K. and Irisson, J.-O. 2016 'Imperfect automatic image classification successfully describes plankton distribution patterns', *Methods in Oceanography*, 15-16, pp. 60–77. doi: [10.1016/j.mio.2016.04.003](https://doi.org/10.1016/j.mio.2016.04.003).
- Falkenhaus, T., Broms, C., Bagøien, E. and Nikolioudakis, N. 2022 'Temporal variability of co-occurring calanus finmarchicus and c. Helgolandicus in skagerrak', *Frontiers in Marine Science*, 9. Available at: <https://www.frontiersin.org/articles/10.3389/fmars.2022.779335> (Accessed: 10 March 2023).
- Floeter, J., Beusekom, J. E. E. van, Auch, D., Callies, U., Carpenter, J., Dudeck, T., Eberle, S., *et al.* 2017 'Pelagic effects of offshore wind farm foundations in the stratified north sea', *Progress in Oceanography*, 156, pp. 154–173. doi: [10.1016/j.pocean.2017.07.003](https://doi.org/10.1016/j.pocean.2017.07.003).
- Floeter, J., Pohlmann, T., Harmer, A. and Möllmann, C. 2022 'Chasing the offshore wind farm wind-wake-induced upwelling/downwelling dipole', *Frontiers in Marine Science*, 9. Available at: <https://www.frontiersin.org/articles/10.3389/fmars.2022.884943> (Accessed: 1 March 2023).
- Flood, P. R. 1978 'Filter characteristics of appendicularian food catching nets', *Experientia*, 34(2), pp. 173–175. doi: [10.1007/BF01944659](https://doi.org/10.1007/BF01944659).
- Fransz, H. G., Colebrook, J. M., Gamble, J. C. and Krause, M. 1991 'The zooplankton of the north sea', *Netherlands Journal of Sea Research*, 28(1), pp. 1–52. doi: [10.1016/0077-7579\(91\)90003-J](https://doi.org/10.1016/0077-7579(91)90003-J).
- Gorsky, G. and Fenaux, R. 1998 'The role of appendicularia in marine food webs', in *The biology of pelagic tunicates*, pp. 161–169.
- Harris, P. T. and Baker, E. K. 2012 '1 - why map benthic habitats?', in Harris, P. T. and Baker, E. K. (eds) *Seafloor geomorphology as benthic habitat*. London: Elsevier, pp. 3–22. doi: [10.1016/B978-0-12-385140-6.00001-3](https://doi.org/10.1016/B978-0-12-385140-6.00001-3).
- Hays, G., Richardson, A. and Robinson, C. 2005 'Climate change and marine plankton', *Trends in ecology & evolution*, 20, pp. 337–44. doi: [10.1016/j.tree.2005.03.004](https://doi.org/10.1016/j.tree.2005.03.004).
- Hidalgo, M., Gusdal, Y., Dingsør, G. E., Hjermmann, D., Ottersen, G., Stige, L. C., Melsom, A., *et al.* 2012 'A combination of hydrodynamical and statistical modelling reveals non-stationary climate effects on fish

- larvae distributions', *Proceedings of the Royal Society B: Biological Sciences*, 279(1727), pp. 275–283. doi: [10.1098/rspb.2011.0750](https://doi.org/10.1098/rspb.2011.0750).
- Hill, A. E., James, I., Linden, P., Matthews, J., Prandle, D., Simpson, J., Gmitrowicz, E., *et al.* 1993 'Dynamics of tidal mixing fronts in the north sea', *Philosophical Transactions of The Royal Society B: Biological Sciences*, 343, pp. 431–446. doi: [10.1098/rsta.1993.0057](https://doi.org/10.1098/rsta.1993.0057).
- Hinton, G. E., Krizhevsky, A. and Wang, S. D. 2011 'Transforming auto-encoders', in Honkela, T., Duch, W., Girolami, M., and Kaski, S. (eds) *Artificial neural networks and machine learning – ICANN 2011*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 44–51. doi: [10.1007/978-3-642-21735-7_6](https://doi.org/10.1007/978-3-642-21735-7_6).
- Inger, R., Attrill, M. J., Bearhop, S., Broderick, A. C., James Grecian, W., Hodgson, D. J., Mills, C., *et al.* 2009 'Marine renewable energy: Potential benefits to biodiversity? An urgent call for research', *Journal of Applied Ecology*, 46(6), pp. 1145–1153. doi: [10.1111/j.1365-2664.2009.01697.x](https://doi.org/10.1111/j.1365-2664.2009.01697.x).
- Kallasvuori, M., Vanhatalo, J. and Veneranta, L. 2017 'Modeling the spatial distribution of larval fish abundance provides essential information for management', *Canadian Journal of Fisheries and Aquatic Sciences*, 74(5), pp. 636–649. doi: [10.1139/cjfas-2016-0008](https://doi.org/10.1139/cjfas-2016-0008).
- Kirby, R. R., Beaugrand, G., Lindley, J. A., Richardson, A. J., Edwards, M. and Reid, P. C. 2007 'Climate effects and benthic–pelagic coupling in the north sea', *Marine Ecology Progress Series*, 330, pp. 31–38. Available at: <https://www.jstor.org/stable/24871087> (Accessed: 2 March 2023).
- Kirby, R. R. and Lindley, J. A. 2005 'Molecular analysis of continuous plankton recorder samples, an examination of echinoderm larvae in the north sea', *Journal of the Marine Biological Association of the United Kingdom*, 85(3), pp. 451–459. doi: [10.1017/S0025315405011392](https://doi.org/10.1017/S0025315405011392).
- Knutsen, H., Olsen, E. M., Ciannelli, L., Espeland, S. H., Knutsen, J. A., Simonsen, J. H., Skreslet, S., *et al.* 2007 'Egg distribution, bottom topography and small-scale cod population structure in a coastal marine system', *Marine Ecology Progress Series*, 333, pp. 249–255. doi: [10.3354/meps333249](https://doi.org/10.3354/meps333249).
- Laakmann, S., Boos, K., Kneibelsberger, T., Raupach, M. J. and Neumann, H. 2016 'Species identification of echinoderms from the north sea by combining morphology and molecular data', *Helgoland Marine Research*, 70(1), p. 18. doi: [10.1186/s10152-016-0468-5](https://doi.org/10.1186/s10152-016-0468-5).
- Landry, M., Peterson, W. and Fagerness, V. 1994 'Mesozooplankton grazing in the southern california bight. I. Population abundances and gut pigment contents', *Marine Ecology Progress Series*, 115, pp. 55–71. doi: [10.3354/meps115055](https://doi.org/10.3354/meps115055).
- Langenkämper, D., Kevelaer, R. van and Nattkemper, T. W. 2019 'Strategies for tackling the class imbalance problem in marine image classification', in Zhang, Z., Suter, D., Tian, Y., Branzan Albu, A., Sidère, N., and Jair Escalante, H. (eds) *Pattern recognition and information forensics*. Cham: Springer International Publishing (Lecture notes in computer science), pp. 26–36. doi: [10.1007/978-3-030-05792-3_3](https://doi.org/10.1007/978-3-030-05792-3_3).
- Leeuwen, S. van, Tett, P., Mills, D. and Molen, J. van der 2015 'Stratified and nonstratified areas in the north sea: Long-term variability and biological and policy implications', *Journal of Geophysical Research: Oceans*, 120(7), pp. 4670–4686. doi: [10.1002/2014JC010485](https://doi.org/10.1002/2014JC010485).

- Lindegren, M., Dakos, V., Gröger, J. P., Gardmark, A., Kornilovs, G., Otto, S. A. and Möllmann, C. 2012 'Early detection of ecosystem regime shifts: a multiple method evaluation for management application', *PLOS ONE*, 7(7), p. e38410. doi: [10.1371/journal.pone.0038410](https://doi.org/10.1371/journal.pone.0038410).
- Lindegren, M., Thomas, M. K., Jónasdóttir, S. H., Nielsen, T. G. and Munk, P. 2020 'Environmental niche separation promotes coexistence among ecologically similar zooplankton species—north sea copepods as a case study', *Limnology and Oceanography*, 65(3), pp. 545–556. doi: [10.1002/lno.11322](https://doi.org/10.1002/lno.11322).
- Lindley, J. A., Gamble, J. C. and Hunt, H. G. 1995 'A change in the zooplankton of the central north sea (55° to 58° n): A possible consequence of changes in the benthos', *Marine Ecology Progress Series*, 119(1), pp. 299–303. Available at: <https://www.jstor.org/stable/24849819> (Accessed: 1 March 2023).
- Lombard, F., Boss, E., Waite, A. M., Vogt, M., Uitz, J., Stemmann, L., Sosik, H. M., *et al.* 2019 'Globally consistent quantitative observations of planktonic ecosystems', *Frontiers in Marine Science*, 6. Available at: <https://www.frontiersin.org/articles/10.3389/fmars.2019.00196> (Accessed: 8 November 2022).
- Lombard, F., Legendre, L., Picheral, M., Sciandra, A. and Gorsky, G. 2010 'Prediction of ecological niches and carbon export by appendicularians using a new multispecies ecophysiological model', *Marine Ecology Progress Series*, 398, pp. 109–125. doi: [10.3354/meps08273](https://doi.org/10.3354/meps08273).
- Luo, J. Y., Irisson, J.-O., Graham, B., Guigand, C., Sarafraz, A., Mader, C. and Cowen, R. K. 2018 'Automated plankton image analysis using convolutional neural networks', *Limnology and Oceanography: Methods*, 16(12), pp. 814–827. doi: [10.1002/lom3.10285](https://doi.org/10.1002/lom3.10285).
- Marteinsdóttir, G., Gudmundsdóttir, A., Thorsteinsson, V. and Stefansson, G. 2000 'Spatial variation in abundance, size composition and viable egg production of spawning cod (*Gadus morhua* L.) in Icelandic waters', *ICES Journal of Marine Science*, 57(4), pp. 824–830. doi: [10.1006/jmsc.2000.0568](https://doi.org/10.1006/jmsc.2000.0568).
- McGill, B. J., Enquist, B. J., Weiher, E. and Westoby, M. 2006 'Rebuilding community ecology from functional traits', *Trends in Ecology & Evolution*, 21(4), pp. 178–185. doi: [10.1016/j.tree.2006.02.002](https://doi.org/10.1016/j.tree.2006.02.002).
- McGinty, N., Barton, A. D., Record, N. R., Finkel, Z. V. and Irwin, A. J. 2018 'Traits structure copepod niches in the north atlantic and southern ocean', *Marine Ecology Progress Series*, 601, pp. 109–126. doi: [10.3354/meps12660](https://doi.org/10.3354/meps12660).
- Möller, K. O., Schmidt, J. O., St.John, M., Temming, A., Diekmann, R., Peters, J., Floeter, J., *et al.* 2015 'Effects of climate-induced habitat changes on a key zooplankton species', *Journal of Plankton Research*, 37(3), pp. 530–541. doi: [10.1093/plankt/fbv033](https://doi.org/10.1093/plankt/fbv033).
- Möllmann, C., Cormon, X., Funk, S., Otto, S. A., Schmidt, J. O., Schwermer, H., Sguotti, C., *et al.* 2021 'Tipping point realized in cod fishery', *Scientific Reports*, 11(1), p. 14259. doi: [10.1038/s41598-021-93843-z](https://doi.org/10.1038/s41598-021-93843-z).
- Möllmann, C., Kornilovs, G., Fetter, M., Köster, F. W. and Hinrichsen, H.-H. 2003 'The marine copepod, *pseudocalanus elongatus*, as a mediator between climate variability and fisheries in the central baltic sea: *Copepod mediates climate to fishery*', *Fisheries Oceanography*, 12(4), pp. 360–368. doi: [10.1046/j.1365-2419.2003.00257.x](https://doi.org/10.1046/j.1365-2419.2003.00257.x).

- Moreno-Torres, J. G., Raeder, T., Alaiz-Rodríguez, R., Chawla, N. V. and Herrera, F. 2012 'A unifying view on dataset shift in classification', *Pattern Recognition*, 45(1), pp. 521–530. doi: [10.1016/j.patcog.2011.06.019](https://doi.org/10.1016/j.patcog.2011.06.019).
- Munk, P., Larsson, P., Danielsen, D. and Moksness, E. 1995 'Larval and small juvenile cod gadus morhua concentrated in the highly productive areas of a shelf break front', *Marine Ecology Progress Series*, 125, pp. 21–30. doi: [10.3354/meps125021](https://doi.org/10.3354/meps125021).
- Paffenhöfer, G.-A. 1976 'On the biology of Appendicularia of the southeastern North Sea'. Available at: <https://www.vliz.be/nl/personen-opzoeken?module=ref&refid=4869&basketaction=add> (Accessed: 7 March 2023).
- Pitois, S. G., Tilbury, J., Bouch, P., Close, H., Barnett, S. and Culverhouse, P. F. 2018 'Comparison of a cost-effective integrated plankton sampling and imaging instrument with traditional systems for mesozooplankton sampling in the celtic sea', *Frontiers in Marine Science*, 5. doi: [10.3389/fmars.2018.00005](https://doi.org/10.3389/fmars.2018.00005).
- Pizarro, O., Rigby, P., Johnson-Roberson, M., Williams, S. B. and Colquhoun, J. 2008 'Towards image-based marine habitat classification', in *OCEANS 2008. OCEANS 2008*, pp. 1–7. doi: [10.1109/OCEANS.2008.5152075](https://doi.org/10.1109/OCEANS.2008.5152075).
- Planque, B. and Fromentin, J.-M. 1996 'Calanus and environment in the eastern north atlantic. I. Spatial and temporal patterns of c. Finmarchicus and c. helgolandicus', *Marine Ecology Progress Series*, 134(1), pp. 101–109. Available at: <https://www.jstor.org/stable/24856130> (Accessed: 2 March 2023).
- Ratcliffe, F. C., Uren Webster, T. M., Rodriguez-Barreto, D., O'Rourke, R., Garcia de Leaniz, C. and Consuegra, S. 2021 'Quantitative assessment of fish larvae community composition in spawning areas using metabarcoding of bulk samples', *Ecological Applications*, 31(3), p. e02284. doi: [10.1002/eap.2284](https://doi.org/10.1002/eap.2284).
- Sabour, S., Frosst, N. and Hinton, G. E. 2017 'Dynamic routing between capsules', *arXiv:1710.09829 [cs]*. Available at: <http://arxiv.org/abs/1710.09829> (Accessed: 1 April 2019).
- Sandø, A. B., Johansen, G. O., Aglen, A., Stiansen, J. E. and Renner, A. H. H. 2020 'Climate change and new potential spawning sites for northeast arctic cod', *Frontiers in Marine Science*, 7. Available at: <https://www.frontiersin.org/articles/10.3389/fmars.2020.00028> (Accessed: 11 April 2023).
- Schmid, M. S., Cowen, R. K., Robinson, K., Luo, J. Y., Briseño-Avena, C. and Sponaugle, S. 2020 'Prey and predator overlap at the edge of a mesoscale eddy: Fine-scale, in-situ distributions to inform our understanding of oceanographic processes', *Scientific Reports*, 10(1), p. 921. doi: [10.1038/s41598-020-57879-x](https://doi.org/10.1038/s41598-020-57879-x).
- Sguotti, C., Blöcker, A. M., Färber, L., Blanz, B., Cormier, R., Diekmann, R., Letschert, J., *et al.* 2022 'Irreversibility of regime shifts in the north sea', *Frontiers in Marine Science*, 9. Available at: <https://www.frontiersin.org/articles/10.3389/fmars.2022.945204> (Accessed: 10 March 2023).
- Swailethorp, R., Malanski, E., Dalgaard Agersted, M., Gissel Nielsen, T. and Munk, P. 2015 'Structuring of zooplankton and fish larvae assemblages in a freshwater-influenced greenlandic fjord: Influence from

- hydrography and prey availability', *Journal of Plankton Research*, 37(1), pp. 102–119. doi: [10.1093/plankt/fbu099](https://doi.org/10.1093/plankt/fbu099).
- Thorpe, R. B., Arroyo, N. L., Safi, G., Niquil, N., Preciado, I., Heath, M., Pace, M. C., *et al.* 2022 'The response of north sea ecosystem functional groups to warming and changes in fishing', *Frontiers in Marine Science*, 9. Available at: <https://www.frontiersin.org/articles/10.3389/fmars.2022.841909> (Accessed: 30 March 2023).
- Van Ginderdeuren, K., Van Hoey, G., Vincx, M. and Hostens, K. 2014 'The mesozooplankton community of the belgian shelf (north sea)', *Journal of Sea Research*, 85, pp. 48–58. doi: [10.1016/j.seares.2013.10.003](https://doi.org/10.1016/j.seares.2013.10.003).
- Walker, J. L. and Orenstein, E. C. 2021 'Improving rare-class recognition of marine plankton with hard negative mining', in. *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 3672–3682. Available at: https://openaccess.thecvf.com/content/ICCV2021W/OceanVision/html/Walker_Improving_Rare-Class_Recognition_of_Marine_Plankton_With_Hard_Negative_Mining_ICCVW_2021_paper.html (Accessed: 14 March 2023).
- Webb, G. I., Lee, L. K., Goethals, B. and Petitjean, F. 2018 'Analyzing concept drift and shift from sample data', *Data Mining and Knowledge Discovery*, 32(5), pp. 1179–1199. doi: [10.1007/s10618-018-0554-1](https://doi.org/10.1007/s10618-018-0554-1).
- Wiebe, P. H. and Benfield, M. C. 2003 'From the hensen net toward four-dimensional biological oceanography', *Progress in Oceanography*, 56(1), pp. 7–136. doi: [10.1016/S0079-6611\(02\)00140-4](https://doi.org/10.1016/S0079-6611(02)00140-4).
- Winder, M., Bouquet, J., Rafael Bermúdez, J., Berger, S. A., Hansen, T., Brandes, J., Sazhin, A. F., *et al.* 2017 'Increased appendicularian zooplankton alter carbon cycling under warmer more acidified ocean conditions', *Limnology and Oceanography*, 62(4), pp. 1541–1551. doi: [10.1002/lno.10516](https://doi.org/10.1002/lno.10516).

Outline of publications

Chapter I: Plonus, R.-M., Conradt, J., Harmer, A., Janßen, S. and Floeter, J. 2021 ‘Automatic plankton image classification—can capsules and filters help cope with data set shift?’, *Limnology and Oceanography: Methods*, 19(3), pp. 176–195. doi: <https://doi.org/10.1002/lom3.10413>.

RMP and JF mainly drafted the manuscript. RMP and JC developed the methodology. AH operated the Triaxus during the field sampling. SJ validated automated image classifications. All authors contributed ideas for the interpretation of the data and the critical revision of the manuscript. All authors gave final approval for publication.

Chapter II: Plonus, R.-M., Vogl, S. and Floeter, J. 2021 ‘Automatic segregation of pelagic habitats’, *Frontiers in Marine Science*, 8. doi: <https://doi.org/10.3389/fmars.2021.754375>.

RMP and JF mainly drafted the manuscript. RMP and SV developed the methodology. All authors contributed ideas for the interpretation of the data and the critical revision of the manuscript. All authors gave final approval for publication.

Chapter III: Plonus, R.-M., Riethmoeller, R. and Floeter, J. unpubl. ‘Identification of plankton habitats in the North Sea’

RMP and JF mainly drafted the manuscript. RR provided access to the SCANFISH data, helped to develop the method, and made comments improving the final manuscript. All authors gave final approval for publication.

Danksagung

An erster Stelle danken möchte ich Prof. Dr. Christian Möllmann dafür, dass er mir die Möglichkeit zu einer Dissertation eröffnet und mir dabei viel Freiraum für persönliche Entwicklung und Erfahrungen gelassen hat. Besonderer Dank gebührt meinem Betreuer Dr. Jens Floeter, der in unzähligen Diskussionen und Review-Prozessen einen entscheidenden Anteil hatte an der inhaltlichen Entwicklung dieser Arbeit und der mir immer mit Rat und Tat zur Seite stand. Bedanken möchte ich mich auch bei Prof. Dr. Stefanie Vogl dafür, dass sie aus meinem gefährlichen Halbwissen eine fundierte Grundlage geschaffen hat für alles, was da noch gekommen ist und vielleicht noch kommen wird. Außerdem gebührt mein Dank Dr. Rolf Riethmüller für seine Mitarbeit und die Bereitstellung der SCANFISH Daten. Bedanken möchte ich mich an dieser Stelle auch bei André Harmer für seinen Einsatz bei der Probennahme mit dem TRIAXUS und bei Silke Janßen, die in unermüdlicher Manier Stunde um Stunde Planktonbilder sortiert hat. Die Diskussionen mit Jan Conradt, bei denen alle anderen längst ausgestiegen sind, werden mir immer in guter Erinnerung bleiben! Dank gebührt außerdem Prof. Dr. Axel Temming für seinen Weitblick, der schon vor 20 Jahren den ersten VPR angeschafft und damit die Weichen für die aktuellen Forschungen gestellt hat. Und außerdem möchte ich mich bei Dr. Klas Ove Möller dafür bedanken, dass er für mich die Brücken geschlagen hat zu I/TAPINA.

Zu guter Letzt möchte ich mich bei meinen Freunden und Kollegen bedanken, die die Zeit am IMF zu dem besonderen Erlebnis gemacht haben, das es nun einmal war – danke!

Eidesstattliche Versicherung

„Hiermit versichere ich an Eides statt, die vorliegende Dissertation selbst verfasst und keine anderen als die angegebenen Hilfsmittel benutzt zu haben. Die eingereichte schriftliche Fassung entspricht der auf dem elektronischen Speichermedium. Ich versichere, dass diese Dissertation nicht in einem früheren Promotionsverfahren eingereicht wurde.“

12.07.2023, *Rene Plonus*

Datum, Unterschrift