Beyond blooms: A novel time series analysis framework predicts seasonal keystone species and sheds light on Arctic pelagic ecosystem stability

Ellen Oldenburg [1,2,3,◇,*], **Raphael M. Kronberg** [3,4,◇], **Katja Metfies** [5,6], **Wilken-Jon von Appen** [7], **Matthias Wietz** [3,8,9], **Christina Bienhold** [3,8], **Ovidiu Popa** [1,△] and **Oliver Ebenhöh** [1,2,△]

[1] *Institute of Theoretical and Quantitative Biology, Heinrich Heine University, Düsseldorf, 40225, North Rhine-Westphalia, Germany*

[2] *Cluster of Excellence on Plant Sciences, Heinrich Heine University, Düsseldorf, 40225, North Rhine-Westphalia, Germany*

[3] *Deep-Sea Ecology and Technology, Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, Bremerhaven, Germany*

[4] *Mathematical Modelling of Biological Systems, Heinrich Heine University, Düsseldorf, 40225, North Rhine-Westphalia, Germany*

[5] *Polar Biological Oceanography, Alfred-Wegener-Institut Helmholtz Center for Polar and Marine Research*

[6] *Helmholtz Institute for Functional Marine Biodiversity at the University of Oldenburg, Oldenburg, Germany*

[7] *Physical Oceanography of the Polar Seas, Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, Bremerhaven, Germany*

[8] *Max Planck Institute for Marine Microbiology, Bremen, Germany*

[9] *Institute for Chemistry and Biology of the Marine Environment,University of Oldenburg, Oldenburg, Germany*

Correspondence*:
Ellen Oldenburg - Universitätstraße 1, 40225 Düsseldorf, Germany
Ellen.Oldenburg@hhu.de

◇ Equal contribution and first authorship: These authors contributed equally to this work and share first authorship.
△ Last authorship: These authors share last authorship.

ABSTRACT

A thorough understanding of ecosystem functioning in the Arctic Ocean, a region under severe threat by climate change, requires detailed studies on linkages between biodiversity and ecosystem stability. The identification of keystone species with special relevance for ecosystem stability is of great importance, yet difficult to achieve with established community assessments. In the case of microbes, metabarcoding and metagenomics offer fundamental insights into community structure and function, yet remain limited regarding conclusions about the role of individual species within the ecosystem. To overcome this limitation, we have developed an analytical approach based on three different methods: Co-Occurrence Networks, Convergent Cross Mapping, and Energy Landscape Analysis. These methods enable the identification of seasonal communities in microbial ecosystems, elucidate their interactions, and predict potential stable community configurations under varying environmental conditions. Combining the outcomes of these three methods allowed us to define 19 keystone species that are representative for the different trophic modes sustaining the local food web. They may serve as indicator species for ecosystem functionality under the impact of environmental change. Our research reveals a clear seasonal pattern in the composition of the phytoplankton community, with distinct assemblages characterizing the carbon fixation (light) and consumption (dark). Species interactions exhibit strong seasonality, with significant influence of summer communities on winter communities but not vice versa. During spring, two distinct groups are present: consumers (heterotrophs), strongly linked to polar night, and photoautotrophs (mainly Bacillariophyta), initiating growth (photoautotrophic Bacillariophyta). These groups are not causally related, suggesting a ”winter reset” with selective effects that facilitates a new blooming period, allowing survivors of the dark phase to emerge. Investigating the fragility of these ecological systems using Energy Landscape Analysis we demonstrate that winter communities are more stable than summer communities. In summary, the ecological landscape of the Fram Strait,a passage, that acts as a gateway between the Arctic Ocean and the Atlantic Ocean, can be categorized by two distinct phases: a production phase governed by specialized organisms that are highly responsive to environmental variability, and a heterotrophic phase dominated by generalist species with enhanced resilience.

Keywords: Cross Convergence Mapping, Co-Occurrence Network, Energy Landscape Analysis, Keystone species, Eukaryotes

1 INTRODUCTION

The Arctic Ocean and climate change

The Arctic Ocean is a unique ecosystem, undergoing a major transition due to the ongoing repercussions of climate change. Over the past two decades, Arctic surface temperatures have risen at a rate more than twice the global average (Meredith et al., 2019). This temperature amplification is linked to feedback mechanisms triggered by reductions in sea ice and snow cover, which exacerbate warming trends. In particular, the Arctic sea ice extent has been declining throughout the year, with the most pronounced decline occurring in September (Meredith et al., 2019). These environmental changes have a wide range of consequences, including profound shifts in biodiversity (Sala et al., 2000), and thus have a fundamental impact on ecosystems of the Arctic Ocean.

There are first signs that the geographical ranges of temperate species are shifting northwards (Kraft et al., 2013), while polar fish and ice-associated species are experiencing a reduction in their habitat due to changing environmental conditions in the Arctic Ocean. These ecological changes have impact on the entire ecosystem and jeopardize its stability (Meredith et al., 2019). The complex relationship between biodiversity and ecosystem stability, particularly in the Arctic Ocean, remains poorly understood. Consequently, the rapid changes in Arctic sea ice and environmental conditions urgently require an improved understanding of the mechanisms governing the resilience and stability of biological processes and ecosystem functions in the Arctic Ocean.

Marine ecosystem functioning and primary production

Within marine ecosystems, primary production is a key service supporting the food web, the survival and growth of marine organisms (Eppley and Peterson, 1979; Lin et al., 2003). The extent of primary production

and its species composition has implications on biodiversity, the abundance and community structure at higher trophic levels, but also carbon sequestration. In the Central Arctic Ocean (CAO), sea-ice algae rather than phytoplankton account for much of the primary production (Gosselin et al., 1997; Fernández-Méndez et al., 2015). This distinct ecosystem feature is supported by a highly productive microalgal community that thrives in sea ice, accompanied by a remarkably diverse heterotrophic community ranging from bacteria to metazoans (Bluhm et al., 2017). These sea ice algae have the potential to initiate pelagic blooms beneath the ice (van Leeuwe et al., 2022). Historically, it was widely accepted that phytoplankton production beneath sea ice in the CAO was negligible (Arrigo et al., 2012). Typically, phytoplankton growth starts mainly within the marginal ice zone in spring, co-occurring with increased solar radiation and meltwater-induced stratification (Clement Kinney et al., 2020). However, over the past three decades, increasing evidence has documented the occurrence of under-ice blooms in the Arctic Ocean, challenging this prevailing view (Strass and Nöthig, 1996; Fortier et al., 2002; Leu et al., 2011; Assmy et al., 2017). Furthermore, phytoplankton in the water column below the ice shows significant differences from the microalgal communities in the sea ice (Hardge et al., 2017). Recent decades have seen a remarkable increase in the abundance of pelagic phytoplankton in the Arctic Ocean, a direct consequence of ongoing environmental change, coincident with increasing primary production by more than 50% (Arrigo et al., 2015; Lewis et al., 2020; Nöthig et al., 2020). Emerging evidence highlights the potential of biodiversity to enhance the stability of ecosystem processes, particularly in the face of dynamic environmental change (Duffy, 2009; Cardinale et al., 2012; Hooper et al., 2012; Mori et al., 2013). Nevertheless, the underlying mechanisms driving this phenomenon remain a subject of debate and limited understanding (Loreau and De Mazancourt, 2013). In natural ecosystems, species interactions are subject to variation as a result of changes in environmental conditions, which can cause a transition from one stable state to another (Ives and Carpenter, 2007). The persence of nearly 2,000 phytoplankton taxa and 1,000 ice-associated protists in the Arctic (Bluhm et al., 2011) indicates the relevance of identifying keystone species for ecosystem stability. These species are crucial for primary production and ecological stability and may serve as indicator species for monitoring the consequences of environmental change in Arctic marine ecosystems (Frey, 2017; Barber et al., 2015; Richter-Menge and Farrell, 2013). Recent findings indicate that high temperatures in natural ecosystems may affect ecological stability, whereas the consequences of alterations to biodiversity remain variable (Zhao et al., 2023). Understanding biological and ecological dynamics in a region of ongoing, and accelerating, change is substantially fostered by novel statistical approaches. Techniques are now accessible to monitor and investigate ecosystem dynamics, enabling examination of the impact of ecological variables on structural stability. For instance, co-occurrence networks (CON) determine and visualise how species coexist within communities or ecosystems (Priest et al., 2023; Ma et al., 2016). Cross-convergence mapping (CCM) helps to identify causality in complex ecosystems, while energy landscape analysis (ELA) aids in building ecological models that simulate and predict how ecosystems respond to disturbances or changes of the environmental parameters (Sugihara et al., 2012; Suzuki et al., 2020, 2021).

In this study we establish a mathematical methodology to reveal ecological relationships, identify key species, and develop scenarios of their long-term resilience using time series data. Additionally, we aim at identifying species that play a crucial role in maintaining stable communities within the Arctic eukaryotic planktonic food web. Furthermore, we seek to define keystone species that can serve as indicators for monitoring the consequences of environmental change in Arctic marine ecosystem stability. Using network analysis techniques like co-occurrence networks and cross convergence mapping, along with energy landscape analysis, our objective is to understand which factors determine the stability of Arctic marine ecosystems. This approach will improve our understanding of the effects of climate change on this ecosystem.

2 METHODS

2.1 Data

Sampling

Samples were collected with Remote Access Samplers (RAS; McLane) deployed in conjunction with oceanographic sensors over four annual cycles (01.08.2016 to 31.07.2020) at $\tilde{3}0$ m depth at the F4 mooring (79.0118N 6.9648E) of LTER HAUSGARTEN and FRAM in the Fram Strait (Soltwedel et al., 2005). The RAS samplers contained 48 sterile bags, each collecting water samples of 500 mL at programmed sampling events every two weeks. The samples were preserved by adding 700 µl of mercuric chloride (7.5% w/v) to the bags prior to sampling. The samples collected from 2016 to 2018 reflects the pool of up to two samples collected one hour apart in two individual bags. Since 2018, we have been pooling samples taken 7 to 8 days apart from two consecutive weeks. Following RAS recovery, water samples were filtered onto Sterivex filter cartridges with a pore size of 0.22 µm (Millipore, USA). Filters were stored at -20°C for later. DNA extraction and sequence analysis (see Supplementary Material - Methods), resulting in amplicon sequence variants (ASVs). We normalised raw ASV counts for CON and CCM using the Hellinger transformation but did not for the energy landscape analysis; hence a different normalisation is introduced in the framework (Suzuki et al., 2021).

Mooring and satellite data

Temperature, salinity and oxygen concentration were measured with a CTD-O_2 attached to the RAS. Physical oceanography sensors were manufacturer-calibrated and processed as described under (von Appen et al., 2021). Raw and processed mooring data are available at PANGAEA `https://doi.org/10.1594/PANGAEA.904565`, `https://doi.org/10.1594/PANGAEA.940744`, `https://doi.pangaea.de/10.1594/PANGAEA.941125` and `https://doi.org/10.1594/PANGAEA.946447`. For chemical sensors, the raw sensor readouts are reported. The fraction of Atlantic and Polar Water were computed following (von Appen et al., 2018) for each sampling event and reported along with distance below the surface (due to mooring blowdown). Sea ice concentration derived from the Advanced Microwave Scanning Radiometer sensor AMSR-2 (Spreen et al., 2008) were downloaded from the Institute of Environmental Physics, University of Bremen (`https://seaice.uni-bremen.de/sea-ice-concentration-amsr-eamsr2`). Sentinel 3A OLCI chlorophyll surface concentrations were downloaded from `https://earth.esa.int/web/sentinel/sentinel-data-access`. For all satellite-derived data, we considered grid points within a radius of 15km around the moorings. Surface water Photosynthetically Active Radiation (PAR) data, with a 4 km grid resolution, was obtained from AQUA-MODIS (Level-3 mapped; SeaWiFS, NASA) and extracted in QGIS v3.14.16 (`http://www.qgis.org`).

Similar to (von Appen et al., 2021), we considered eigth environmental variables : mixed layer depth (MLD in m), water temperature (temp °C), polar-water fraction (PW_frac %), chlorophyll concentration from in situ sensor (chl_sens $\sim \mu$g l^{-1}), PAR (μ mol photons $m^{-}2d^{-}1$), Salinity (PSU), oxygen concentration (O2_conc μmol l^{-1}) and sampling depth (depth m).

2.2 Co-Occurrence Network

The abundance of species over the full observation period were converted into temporal profiles by employing Fourier transformation techniques to time-series signals. These temporal profiles rely on the 14 Fourier coefficients. We chose 14 coefficients because they reflect the majority of observed species abundance peaks within the four years. To investigate the similarity of temporal profiles between species pairs, we performed pairwise correlations between the individual temporal profiles, where pairs with higher Pearson correlation value show also a similar temporal profile. Pairs with at least 0.7 ($p<0.05$) Pearson correlation were then visualised in an undirected graph. Only positive correlations were retained to later focus on co-operative relationships. To identify strongly connected components that reflect the existing communities of co-occurring species, we applied the Louvain community detection algorithm (Blondel et al., 2008) on the entire graph. The entire process was implemented using the `CCM` and `networkx` packages in Python; visualization was performed using Cytoscape with the Edge-weighted Spring-Embedded Layout (Shannon et al., 2003). The whole co-occurrence network construction is described in the Supplements Supplementary Material - Methods Co-Occurrence Network.

2.2.1 Distance between Clusters

To measure the distance between previously defined Louvain communities (clusters), we applied UMAP on time signals obtained after Fourier decomposition of the abundance data. From this we generated a three-dimensional embedding space. Centroids for each cluster were calculated within this space (see Supplementary Material - Results Figure S4,S5,S6). The network distance between clusters was determined as the Euclidean distance between their centroids. Subsequently, a distance matrix was created and distances were rounded to integers, with only significant connections retained in the pruned matrix.

2.3 Convergent Cross Mapping

Convergent Cross Mapping (CCM) identifies potential causal relationships between variables in time series data. It quantifies how knowledge of the time series of one species allows predicting the time series of an another species. We first built a CCM network from all pairwise combinations. From this, we extracted the in- and outgoing edges between nodes that are also connected in the co-occurrence network. We used the implementation of Normalized Mutual Information (NMI) from `https://github.com/polsys/ennemi` by Petri Laarne and the Convergent Cross Mapping by Implementation from Javier, Prince `https://github.com/PrinceJavier/causal_ccm` (Javier et al., 2022) to measure the strength of the causal relationship considering also non-linear relations. We could show that the implementation of Normalized Mutual Information (NMI) results into similar findings as the original implementation based on Pearson correlation (Veilleux, 1979; Sugihara et al., 2012) (see Supplementary Material - Methods).

Using a permutation approach (Ma et al., 2016) on the connectivity of the network we calculated significance values for the edge weights, quantifying whether the respective NMI values are greater than expected for random edges (see Supplementary Material - Results). The whole CCM network construction and validation are described in Supplementary Material - Methods and Supplementary Material - Results (Convergent Cross Mapping) of the Supplementary Information.

2.3.1 Aggregation on cluster level

We simplify the network of interactions between single species into a network of interactions between clusters. For this, we assign a weight to a directed edge between two clusters by calculating the arithmetic mean of normalized mutual information (NMI) of all (directed) edges connecting species belonging to the respective clusters. This process effectively reduces the number of items in the node cloud, representing clusters through a unified composite node.

2.4 Energy Landscape Analysis

Energy Landscape Analysis is a method based on statistical physics. From data for many points in time, which contain species abundance and environmental variables, an *energy landscape* is reconstructed. This energy landscape is a function that maps species abundance and environmental variables to an energy value. In analogy to the potential energy in physics, a (local) minimum of this energy landscape indicates a stable community state. Here, we reconstruct the energy landscape function based on the complete time series of ASV abundance together with the available environmental data. We use the reconstructed function to determine the stability of observed communities, and in particular the seasonal clusters determined by the co-occurrence network, and we predict the most stable community compositions. Moreover, because this function also depends on environmental variables, we can also predict how stable a given community is under perturbed environmental conditions. Understanding the existence and the nature of stable community states and how they change in response to environmental shifts is crucial for comprehending the resilience and adaptability of ecosystems in the face of various ecological challenges. Details of our analysis, including parameters and thresholds applied, are described in in Supplementary Material - Methods (Energy Landscape Analysis). Our analysis focused on the Top 100 ASVs within each cluster. Outliers were excluded solely for the purpose of plotting.

2.5 Keystone Species Definition

After collecting attributes from a co-occurrence analysis and distinguishing between mutualistic influence and occurrence just by chance, we calculated the stable states for different clusters using ELA. This information was merged to define keystone species. We defined a keystone species as a species with i) a significant influence on other organism in the network (significant NMI value), ii) a high centrality

(closeness) value within its co-occurrence community and iii) present in at least one stable state as predicted by energy landscape analysis. A significant high centrality value was determined by comparing each centrality value of a single node to the average centrality values of all nodes from the graph using a one sided, one-sample t-test with Benjamini-Hochberg correction for multiple testing (similar to (Guimera and Nunes Amaral, 2005; Joyce et al., 2010).

2.6 Season Definition

Here we have defined the seasons as follows, based on the availability of light (PAR), the seasons and the month.

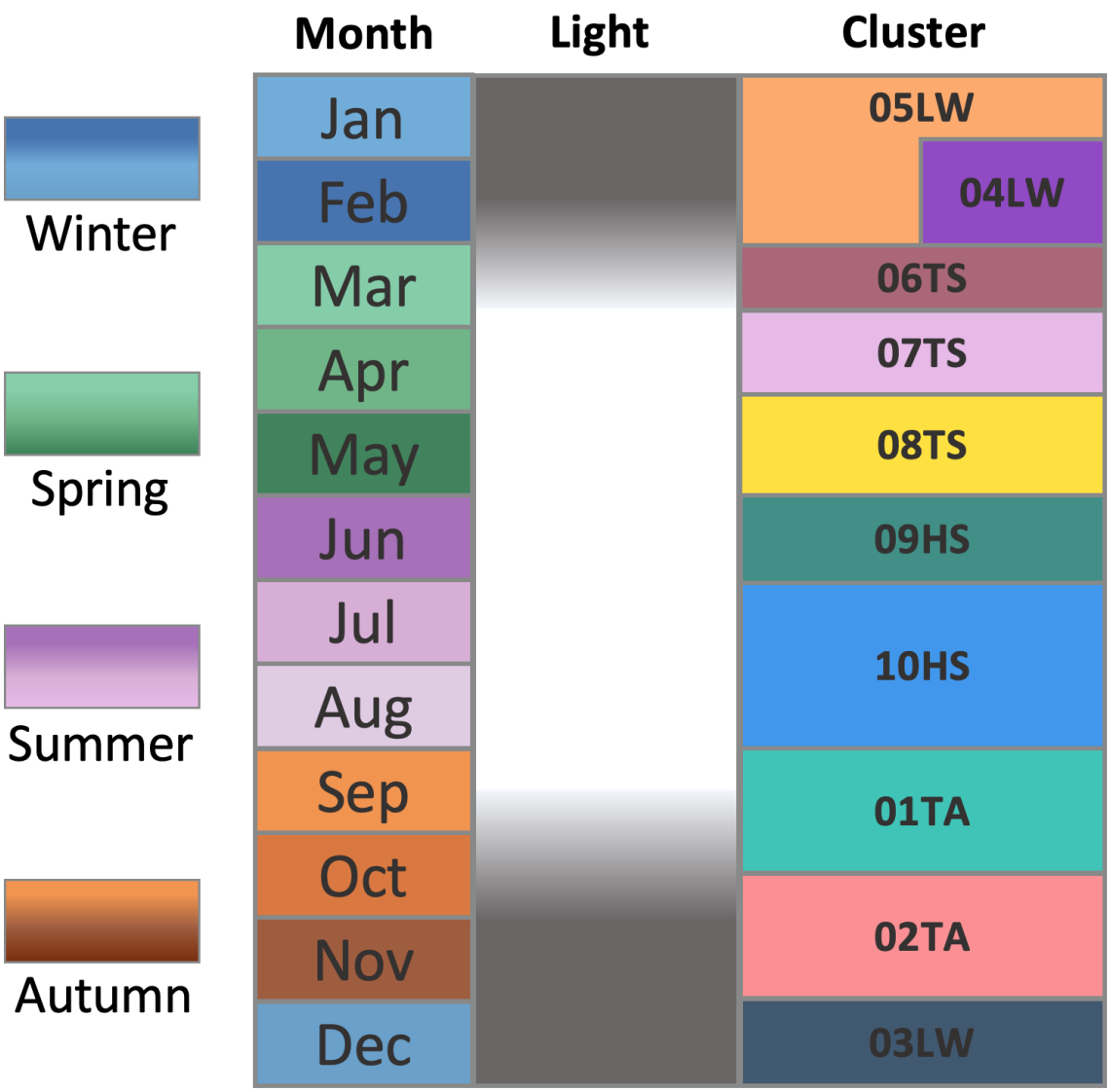

Figure 1. Cluster definition: The names of the cluster are based on the light availability, categorised in transition areas between dark and light(T), high light (H) and low light (L) based on the PAR parameter and the season spring (S), summer(S), autumn (A) and winter (W). Each season consist of three months.

3 RESULTS

We examined a dataset of 1,019 eukaryotic ASVs and eight environmental parameters compiled over four years at mooring site F4 in the West Spitsbergen Current (WSC) in Fram Strait. The aim was to characterize species communities, analysing causal relationships between ASVs and to identify keystone and resilient species with respect to the impact of various environmental conditions.

To accomplish this, we established a novel computational pipeline, coupling co-occurrence analysis with convergence cross mapping and energy landscape analysis. This allowed us to identify causal interactions among species in a co-occurring community and to identify stable community states across different environmental conditions.

3.1 Co-Occurrence Network reveals seasonal dynamics

The co-occurrence network (CON) comprised five connected components, with a major component accounting for 99% (905) of all nodes, which are connected by 8,429 edges. In the following, our analysis is restricted to this major connected component. The resulting undirected graph notably displays a clear seasonal cyclic pattern (Figure 2 A).

The network was partitioned using the Louvain community detection algorithm (Blondel et al., 2008), revealing ten discrete community clusters (Figure 2 B) labeled by the season in which the majority of clusters members had their maximum abundance (Table 1). To further group the clusters, we submerge each three month period to one season. Two clusters were assigned to the transition autumn period (01TA and 02TA), three clusters were associated with the low light winter period (03LW, 04LW and 05LW) and three clusters with the transition spring period (06TS, 07TS and 08TS). Finally, clusters 09 and 10 were allocated to the high light summer season (09HS, 10HS).

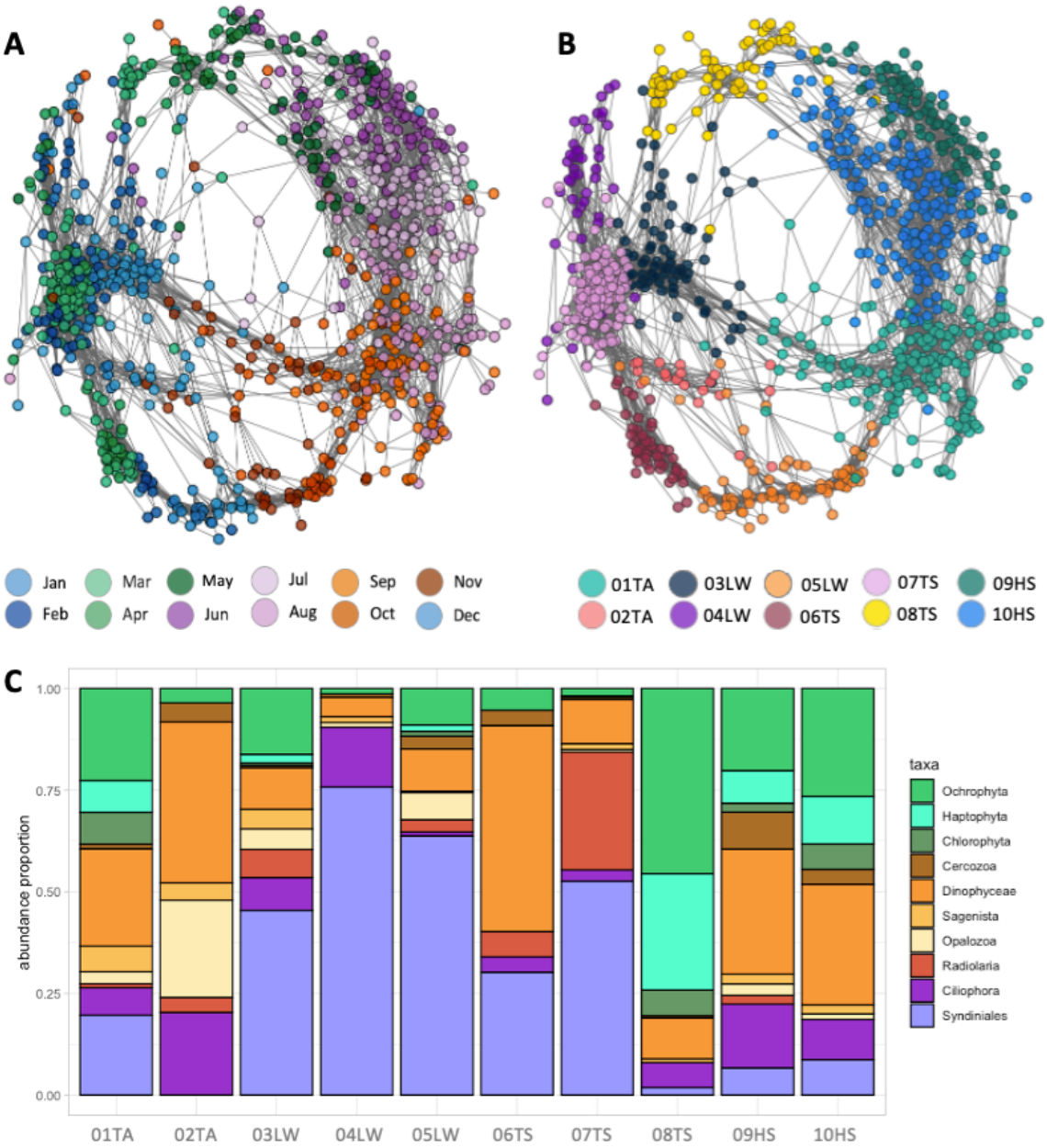

Figure 2. Co-Occurrence Networks and community composition of microeukaryotes at mooring F4 from 2016-08-01 to 2020-07-31.
Each network node represents an ASV, and each edge represents a similar temporal pattern of two ASVs. The edge weights correspond to the Pearson correlation coefficients determined from the comparison of the individual ASV temporal profiles. ASVs are connected if the coefficient is $r > 0.7, p < 0.05$. **A:** Node color reflect the month in which the ASV exhibit maximal abundance, calculated from the maximum abundance mode for each year ranging from January to December. **B:** In this representation, nodes are coloured based on the community membership that was determined by the Louvain community detection algorithm. **C:** The relative abundance of the top 10 taxonomic classes by cluster ('HS' denotes high light summer clusters, 'LW' represents low light winter clusters, 'TS' corresponds to transition spring clusters, and 'TA' indicates transition autumn). The percentage of relative abundance is expressed, with colour shades illustrating the assignment to auto- (green), mixo- (orange) or heterotroph (purple) .

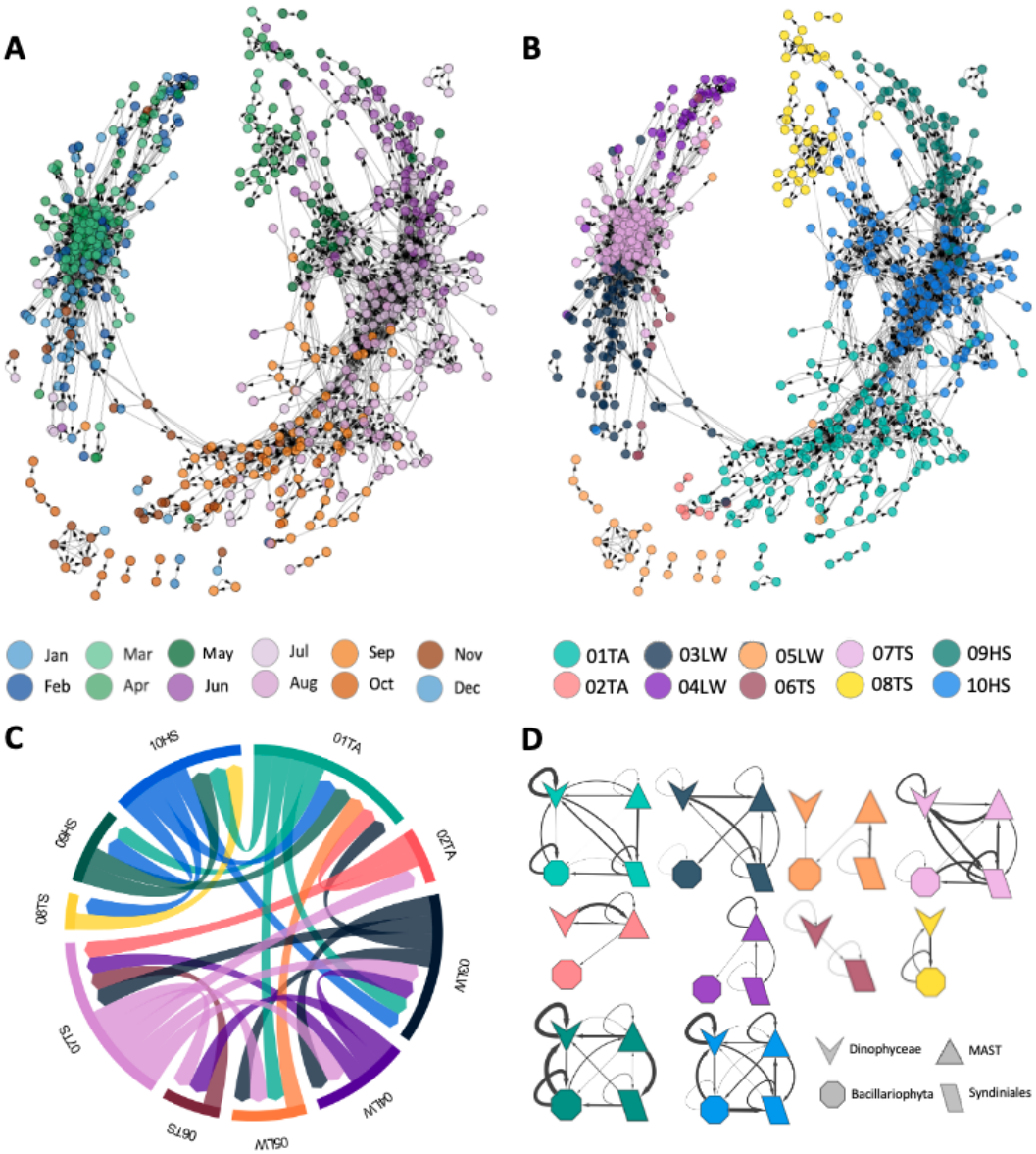

Figure 3. Convergence Cross Mapping Networks of microeukaryotes at mooring F4 from 2016-08-01 to 2020-07-31.
Each node in the CCM network represents an ASV, and each edge represents the causal influences. The edge weight corresponds to the normalized mutual information determined from the comparison of the individual ASV and their predicted representation in the shadow manifold. ASVs are connected if the smoothed p-value of the weight is $p < 0.05$. **A:** Node color reflects the month in which the ASV exhibits maximal abundance, calculated from the maximum abundance mode for each year ranging from January to December. **B:** In this representation, nodes are coloured based on the community membership that was determined by the Louvain community detection algorithm. 'HS' labels denote high light summer clusters, 'LW' represents low light winter clusters, 'TS' corresponds to transition spring clusters, and 'TA' indicates transition autumn clusters. **C:** The normalised mutual information aggregated across the edges between the clusters, visually represented by thickness of the arrows corresponding to their respective values. Corresponding colour visually represent the clusters. **D:** Interaction analysis between taxonomic clusters. For each of the ten clusters, the interactions between ASV groups are examined at class level, considering 'Syndiniales', 'Dinophyceae', 'Bacillariophyta' and 'MAST'. The cluster assignments are marked by different colors. The thickness of the arrows denotes the strength of the interaction, while the shapes represent the various taxa groups at the class level.

Table 1. Co-occurrence network clusters. Ten labeled clusters with their assigned season based on month in which the ASV exhibit maximal abundance, calculated from the maximum abundance mode (majority vote) for each year ranging from January to December, the number of ASVs per cluster, and the number of significant edges in network graph.

Cluster	MaxMonth	Season	Number of ASVs	Number of edges
01TA	9	Autumn	168	1024
02TA	11	Autumn	20	78
03LW	12	Winter	75	509
04LW	2	Winter	45	166
05LW	12	Winter	73	466
06TS	3	Spring	50	382
07TS	3	Spring	154	3062
08TS	4	Spring	62	260
09HS	6	Summer	83	492
10HS	7	Summer	175	1113

3.2 Community composition

We explored the taxonomic community per cluster to explore the seasonal associations of each specific taxonomic group. Alpha biodiversity, measured by Shannon entropy, decreases from summer through autumn and winter, gradually decreasing towards spring (Supplementary Material - Results Figure S1). The beta biodiversity, measured by Bray-Curtis distance (Supplementary Material - Results Figure S2), between the winter and spring clusters (03LW, 04LW, 05LW and 06TS) is notably lower than the most other scores, except that between 01TA and 10HS. Cluster 02TA exhibits on average a higher beta diversity compared to all other clusters, which can be explained by the fact that 02TA is the smallest cluster in terms of the number of ASVs (Supplementary Material - Results Figure S2).

In our study, we found distinct taxonomic compositions within various clusters. Photosynthetic organisms like Ochrophyta and Haptophyta dominate the light phases (Oldenburg et al., 2024). In late spring (cluster 08TS) phototrophs make up more than 75% of ASVs, while during summer (clusters 08TS and 09HS) and early autumn (10HS) they still comprise over 25% of all ASVs. Mixotrophs are highly abundant in most clusters, while they clearly dominate during the late autumn transition (cluster 02TA). Through the complete dark period (clusters 03LW, 04LW and 05LW) as well as in early spring (06TS and 07TS), heterotrophs, particularly Syndiniales, are dominant wiht a clear peak of abundance (more than 90% of ASVs) in mid winter (cluster 04LW). During early spring (cluster 06TS) when sunlight appears again, mixotrophs increase in their abundance, highlighting the nuanced trophic dynamics during the annual cycle. This comprehensive analysis at taxa level provides insights into the composition of these clusters, shedding light on the prevalence and distribution of specific classes within distinct seasonal communities (Figure 2 C).

3.3 Convergent Cross Mapping identifies Community Interactions

Mutual influence between ASVs was estimated via Convergent Cross Mapping (CCM). This method allows predicting causal relationships within and between seasonal clusters based on the underlying ASV dynamics. We project the weights derived from Convergent Cross Mapping (CCM) onto the co-occurrence network, resulting in a directed graph consisting of 16,858 directed edges and 905 nodes. Here, a directed edge indicates that knowledge of the dynamics of the source node allows for prediction of the dynamics of the target node.

A comparative analysis of edge weights within the CCM network was conducted. The connectivity derived from the co-occurrence network was compared with theoretical edge weights and randomly permuted connections. To perform this comparison, a two-sided Kolmogorov-Smirnov test was used. The theoretical edge weights were derived from all possible connections between pairs of nodes in the co-occurrence network, excluding the existing true links (Section 3.3). The findings clearly show that there is a stronger mutualistic influence (higher NMI values) between co-occurring species compared to random

or to unconnected nodes (Supplementary Material - Results Figure S8 and S9 and Table S1 and S2). The significance level was set at a nonparametric p-value of less than 0.05 calculated similar to Ma et al. (2016). Trimming edges with non-significant NMI values (Supplementary Material - Results Figure S10), produced a network graph consisting of 3,654 edges and 637 nodes, divided into 14 disconnected components, with the largest module encompassing 603 nodes (Fig. 3 A,B).

Of the total 13,204 edges eliminated during trimming, 18.79% represent connections between species reaching their peak abundance in March. This selective removal has a profound impact on the network structure, (Figure 3 A). In contrast to the co-occurrence network, the causal interaction network "breaks" during spring season, as demonstrated in Figure 3 A (months March and April). The winter cluster 05LW and the spring cluster 06TS collapse (see Figure 3 B), meaning that nodes disintegrate and large parts of the clusters are no longer connected to the rest of the network. This suggests that the corresponding connections in the co-occurrence network are not a result of causal interactions, but rather result from other factors, possibly caused by the prevailing environmental conditions. Analysis of the betweenness centrality reveals that species of Picozoa, Leegaardiella, Acantharea, Dinophyceae, MAST-1, and Syndiniales serve as essential hub nodes throughout the seasonal cycle in the network, highlighting their crucial function in maintaining network stability.

3.4 Community Interaction

Analyzing cluster interactions revealed distinct patterns. We measure distance of clusters by "network distance", a metric designed to evaluate the separation between clusters. This measure is computed by assessing the distance between the centroids of clusters within the dimension-reduced UMAP embedding space (Supplementary Material - Results). Distances between clusters thus determined range between 1 and 7. Proximate clusters (a network distance of two to four), exhibited notably higher connectivity compared to clusters situated further apart (distance five to eight). Figure 3 C provides a visual and quantitative representation of these interactions. Subsequent analysis revealed a prevalence of connections at a network distance of two (44% of total connections), followed by distances of three (29% of total connections), and four (26% of total connections) (Supplementary Material - Results Figure S4,S5,S6).

The average effects of two linked clusters revealed that summer has a greater influence on winter than the inverse association. This suggests that predicting community composition in winter, based on the preceding summer community, is more reliable than vice versa (Figure 3 C).

For each seasonal cluster, ee investigate in detail the mutual influence of four taxonomic groups selected from the top ten classifications (Figure 2 C): Bacillariophyta, Syndiniales, Dinophyceae, and MAST (MArine STramenopile; all MAST-X variants were classified under MAST). Bacillariophyta primarily comprises photo-autotrophic species (Mann et al., 2017), while Syndiniales include parasitic species, most of them characterized by their heterotrophic lifestyle (Suter et al., 2022). Dinophyceae are known for their diverse array of species and ecological roles, from symbionts to planktonic autotrophs (Lin et al., 2022). MASTs are heterotrophic protists and contribute substantially to protist abundances in the ocean. They play a crucial role in marine ecosystems and are one of the dominant eukaryotes in the Arctic Ocean (Thaler and Lovejoy, 2014; Lin et al., 2022).

These four taxonomic groups are primarily distinguished by their unique lifestyles and ecological roles as primary producers, consumers, parasites, or endosymbiotic interactors. These distinctions form the basis for our analysis of their contributions to the ecosystem. Hence, by summarizing the members of each group into single nodes, we analyzed their cross-interactions using the information obtained from the CCM network.

The strength and direction of interactions between these taxonomic groups varied over the annual cycle (see Figure 3 D). Notably, during the spring-summer and summer-fall transition periods, clusters 02TA and 06TS displayed fewer and weaker connections compared to other clusters. This suggests dynamic changes in community structure during these periods, with ecological interactions between individual species either yet to be established or no longer present. At the start of the polar night (cluster 03LW), we detected the most substantial influence from the dinoflagellates (shown by the thickest arrow in Figure 3 D) to the pico-eukaryotic heterotrophic groups Syndinales and MAST, suggesting a crucial ecological role of dinoflagellates for the establishment of the winter community. Overall, the strength of the links between the taxonomic groups decreases as the polar night progresses and has its minimum at the peak of the polar night in December (Cluster 05LW). The lack of strong connections between taxonomic groups during

the deepest polar night indicates that there are only a few ecologically significant interactions within the eukaryotic microbial community.

During polar day, as the sunlight becomes more intense, the connections between taxonomic groups become stronger and reach their peak during the zenith of the polar day (Cluster 10HS). Notably, Bacillariophyta (i.e. diatoms) showed the most robust connections during the polar day, owing to their role as predominat phototrophic biomass producer and form the foundation of the marine food web. However, towards the end of the growth period, the impact of MAST on Dinoflagellates becomes more pronounced (as shown by the thicker arrow), indicating an essential involvement of this pico-eukaryotic heterotroph in the ecosystem during the later growth phase. This is also a signature of the transition from primary production to recycling.

3.5 Community and Environment Interactions

For a more detailed understanding of which environmental conditions align with seasonal community clusters, we conducted a correlation analysis (Figure 4). Cluster 10HS displayed a significant positive correlation with Photosynthetically Active Radiation (PAR) (0.67) and temperature (0.49), but a significant negative correlation with Mixed Layer Depth (MLD) (-0.61). This cluster thrives in environments with high light and temperature levels but less deeper mixed layers. Cluster 03LW exhibits the opposite behavior, showing a moderately positive correlation with MLD (0.23) and polar water fraction (PW_frac) (0.24) while displaying an inverse relationship with PAR and temperature (-0.35 and -0.36, respectively).

3.6 Energy Landscape Analysis determines stability of microbial communities

We performed energy landscape analysis to study how stable the observed communities are under the prevailing environmental conditions. We focus on four clusters representing the four seasons (01TA for autumn, 03LW for winter, 08TS for spring and 10HS for summer). For each of these clusters, we determined the energy landscape, which is a highly complex function that depends on the abundances of all ASVs and the environmental parameters (see Figure 5). To approximately visualise this landscape, we plot an interpolated smooth surface as a function of the two most significant NMDS dimensions. In addition, for each time point in the time series data, we evaluate the energy landscape function and represent each energy value by a point in the three-dimensional diagram, where the z-axis represents the energy value.

For the landscape reconstructed for cluster 01TA (Figure 5 A) the communities observed in autumn display lower energy values than communities observed in other seasons. This demonstrates that the autumn communities exhibit a high stability. For the winter cluster 03LW (Figure 5 B), the picture is less clear. Whereas the interpolated energy landscape has a more pronounced minimum, the energy values of the observed communities are not clearly separated. As a tendency, the summer communities have a high energy value, demonstrating that summer communities are unstable in winter conditions. However, spring and autumn communities exhibit comparable energy values as winter communities, which indicates that stable community structures in winter conditions are not clearly defined. This trend is even more pronounced for the spring cluster 08TS (Figure 5 C). Here, the interpolated energy landscape shows a broad and shallow minimum, and the energy values of the all observed communities, regardless of the season in which they are found, are very similar. This suggests that under spring conditions community structures are not very stable and that community compositions show a high plasticity. As a consequence, many different communities may exist under spring conditions. This explains the observation of the CCM analysis, which demonstrated that knowledge of the composition of winter communities does not allow to predict the composition of spring communities. Finally, the energy landscape in summer (Figure 5 D) shows a pronounced minimum, in which the observed summer communities are also found. This indicates that summer conditions support well-defined community structures which exhibit a high degree of stability.

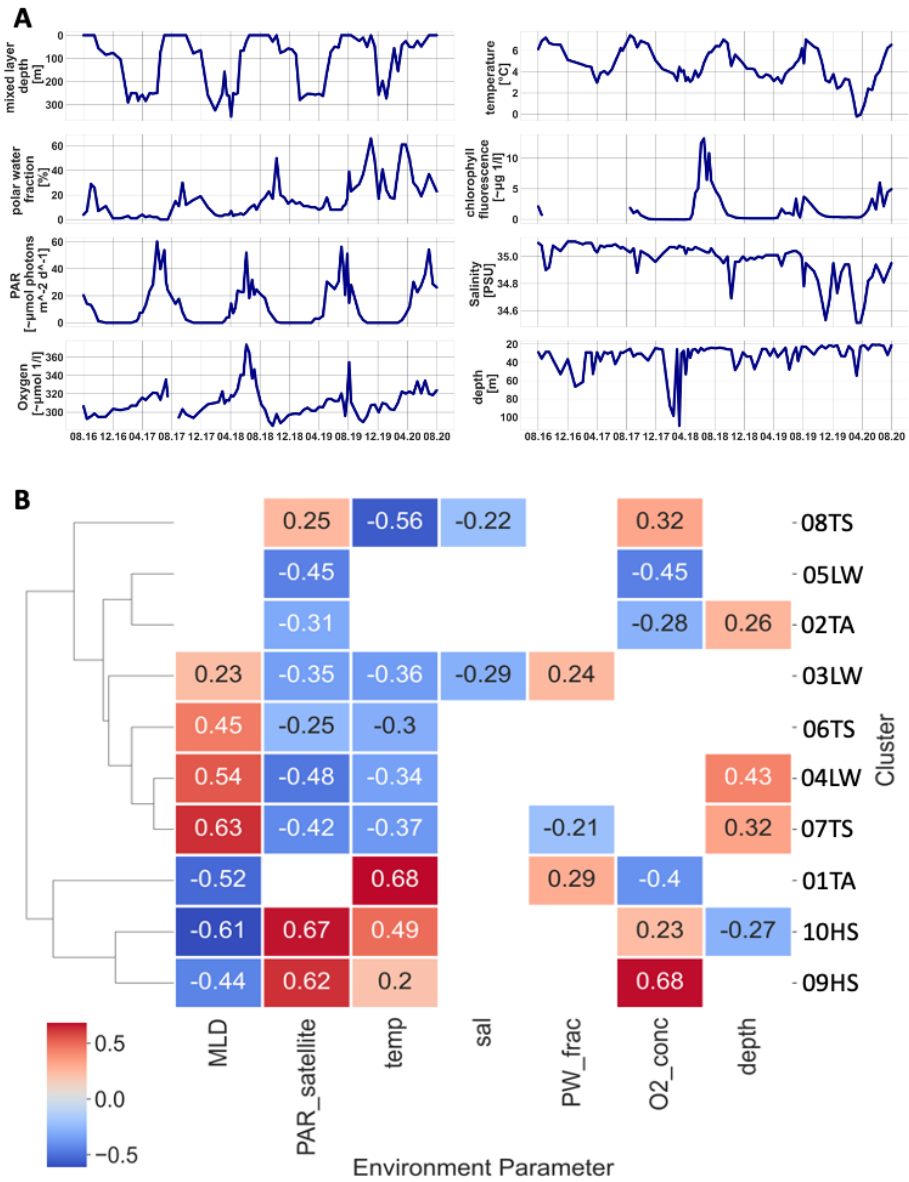

Figure 4. Environmental data and their correlation Louvain clusters. A): Environmental data for F4 from 2016-08-01 to 2020-07-31. The x-axis represents the time period, while the y-axis indicates the following parameters: Mixed Layer Depth (MLD) [m], Temperature [°C], Chlorophyll Fluorescence [μ g/L], Polar Water Fraction [%], Photosynthetically Active Radiation (PAR) [μ mol photons/m^2/d], Salinity [Practical Salinity Units (PSU)], Oxygen Concentration [μ mol/L], Depth of measurement [m]. **B): Correlations between environmental parameters and seasonal Louvain clusters.** The displayed chart shows the environmental parameters from section a) in relation to seasonal clusters. These clusters are characterized by the cumulative relative abundance of ASVs. Cluster labels include '01TA', '02TA', '03LW', '04LW', '05LW', '06TS', '07TS', '08TS', '09HS', and '10HS', where 'TA' denotes transition autumn clusters, 'LW' represents low light winter clusters, 'TS' pertains to transition spring clusters, and 'HS' signifies high light summer clusters. The colour gradient used in the heatmap illustrates the strength of the correlation visually, with blue shades indicating negative correlations and red shades indicating positive correlations. It is worth noting that a significance mask has been applied to show only correlations that are statistically significant.

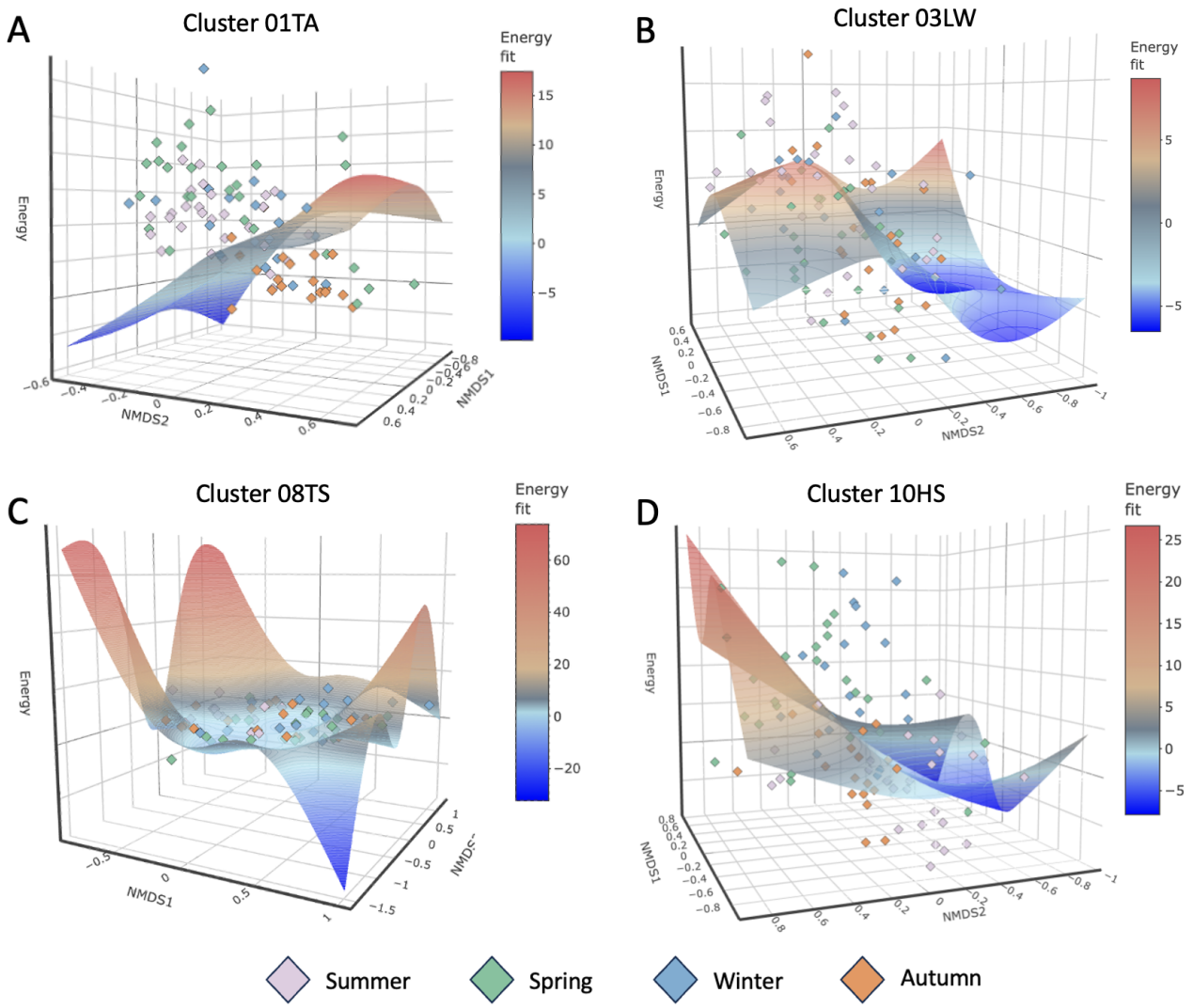

Figure 5. Energy landscapes depicting community structure dynamics. The plots display the reconstructed energy landscape on the NMDS surface for a cluster of each season. Community states, which are defined by ASV compositions and occupy lower-energy regions, indicate higher stability within the energy landscapes. **A):** The transition autumn cluster 01TA. **B):** The low light winter cluster 03LW. **C):** The transition spring cluster 08TS. **D):** The high light summer cluster 10HS. Environmental landscapes over the NMDS surface are reconstructed for each of the four example clusters. The z-axis displays the energy, while the x- and y-axes display the first and second NMDS dimensions. The landscape contours were estimated using a smoothing spline approach with optimized penalty parameters.

3.7 Characterisation of microeukaryote keystone species in the Fram Strait

According to our definition (Section 2.5) a keystone species is highly connected in the co-occurrence network, has a high influence on other species, and appears in a stable community. Applying this definition (see Section 2.5 for details) resulted in 19 keystone species across the complete annual cycle within the measured environmental profile (Table 2). Nine of these keystone species are associated with summer clusters (seven and two are found in clusters 09HS and 10HS, respectively), five with winter (four and one in the winter clusters 03LW and 04LW, respectively), four are associated with autumn (cluster 01TA) and one with spring (cluster 07TS). Remarkably, the nine keystone species from the summer clusters belong to the taxonomic groups Ochrophyta (5), Dinophyceae (2), and Ciliophora (2). These groups include Fragilariopsis, Pseudo-nitzschia, and Thalassiosira, which are known to contribute significantly to the biomass during this season (Von Quillfeldt, 2000) and to serve as prey for higher trophic levels, including microzooplankton (Cleary et al., 2016; Yang et al., 2015). Notably, Fragilariopsis and Thalassiosira exhibit the highest abundance within this cluster. The keystone species in the winter clusters comprise Syndiniales (3), Radiolaria (1), and Dinophyceae (1), autumn cluster keystone species are Ochrophyta (2), Chlorophyta (1) and Syndiniales (1) and the spring keystone species belongs to Radiolaria. The fact that we found only a single keystone species for the spring cluster indicates that communities in this period are most variable, which was also suggested by the energy landscape analysis (Section 3.6). These 19 keystone species reflect all major ecological roles including primary production, consumption, and parasitic interactions. Notably, the emergence of Chlorophyta as a keystone species during the beginning of autumn suggests a shift in primary production from Ochrophyta to Chlorophyta, species that may prefer colder temperatures (Tragin and Vaulot, 2018) and are better adapted to nutrient limitation (Maat et al., 2014).

Table 2. ASV identified as potential keystone species for clusters 10HS, 06TS, 03LW and 01TA. The taxonomic classifications, clusters, raw abundance, proportion of total raw abundance, cluster abundance, and proportion of cluster raw abundance are presented in summary form over the 4-year observation period. The column, significance, indicates if this ASV (Nodes) has at least on significant CCM connection measured in normalized mutual information.

Nodes	Phylum	Genus	Species	Cluster	relAbund	Closeness
euk_asv12	Ochrophyta	Fragilariopsis	Fragilariopsis uc	10HS	0.009	0.356
euk_asv35	Ochrophyta	Fragilariopsis	Fragilariopsis_sublineata	10HS	0.007	0.345
euk_asv24	Ochrophyta	Pseudo-nitzschia	Pseudo-nitzschia_sp.	10HS	0.006	0.312
euk_asv28	Ochrophyta	Thalassiosira	Thalassiosira uc	10HS	0.006	0.317
euk_asv54	Dinoflagellata	Gyrodinium	Gyrodinium_fusiforme	10HS	0.006	0.308
euk_asv52	Dinoflagellata	Peridiniales uc	Peridiniales uc	10HS	0.005	0.345
euk_asv125	Ciliophora	Dadayiella	Dadayiella_ganymedes	10HS	0.003	0.356
euk_asv115	Ochrophyta	Mediophyceae uc	Mediophyceae uc	09HS	0.004	0.284
euk_asv186	Ciliophora	Strombidiidae_H uc	Strombidiidae_H uc_sp.	09HS	0.002	0.313
euk_asv5	Radiolaria	RAD-C uc	RAD-C uc_sp.	07TS	0.014	0.307
euk_asv21	Dinoflagellata	Dino-II-10-and-11 uc	Dino-II-10-and-11 uc_sp.	04LW	0.009	0.291
euk_asv87	Radiolaria	RAD-B-Group-IV uc	RAD-B-Group-IV uc_sp.	03LW	0.002	0.338
euk_asv198	Dinoflagellata	Dino-II-6 uc	Dino-II-6 uc_sp.	03LW	0.001	0.314
euk_asv213	Dinoflagellata	Gymnodinium	Gymnodinium_sp.	03LW	0.001	0.309
euk_asv780	Dinoflagellata	Dino-I-5 uc	Dino-I-5 uc_sp.	03LW	0.000	0.299
euk_asv163	Dinoflagellata	Dino-I-1 uc	Dino-I-1 uc_sp.	01TA	0.003	0.368
euk_asv118	Chlorophyta	Chloroparvula	Chloroparvula_pacifica	01TA	0.002	0.369
euk_asv330	Ochrophyta	Pelagophyceae uc	Pelagophyceae uc_sp.	01TA	0.002	0.326
euk_asv401	Ochrophyta	Aureococcus	Aureococcus_anophagefferens	01TA	0.001	0.315

3.8 Predicting community stability in altered environments

One critical ecological question is how marine microbial communities change if they experience different climatic conditions. We employ energy landscape analysis to assess the stability of observed communities under environmental conditions different to those they usually experience. This is done by evaluating the energy landscape function that resulted from the ELA analysis (see Section 2.4) for given communities and environmental parameters. First, we test the reliability of this approach by evaluating the communities of the summer (10HS) and winter (03LW) clusters under typical Atlantic summer or winter days. Here, we

defined "typical" Atlantic days by taking the averages of the environmental parameters during a three-month period (see Supplementary Material - Methods Table S0). The energy values remain almost unchanged if we evaluate the summer cluster communities in a typical summer day and the winter cluster communities in a typical winter day (see Supplementary Material - Results Figure S13). This is to be expected, because the ASVs in the summer clusters assume their most stable community configuration during the summer months, and likewise for the winter cluster. When evaluating the cluster communities under the opposite conditions, i.e. the summer cluster communities with environmental parameters representing a winter day and vice versa, the energy values are drastically changed (see Supplementary Material - Results Figure S13). Those communities of the summer cluster, which were still present in the winter, exhibit an increased stability (lower energy) than those present in summer. The Atlantic projection panel (Fig. Supplementary Material - Results Figure S13 (C) and (E)) display almost opposite oscillatory time courses. Placing communities of the winter cluster in summer conditions has a more differentiated effect. The communities found in winter 2018 and 2020 are highly unstable in this environment, while communities found in summers appear to exhibit an increased stability. This observation allows speculating that some ASVs of the winter cluster might even benefit from warmer summer conditions.

We next investigate a scenario that might result from future temperature increase and a continued melting of the ice caps. Ice-free summers in the Arctic will allow Atlantic water, and with this Atlantic microbial communities, to enter polar regions. We therefore analyze how the stability of the communities found in the summer (10HS) and winter (03LW) clusters changes if they experience typical Arctic conditions. For this, we defined "typical" Arctic summer and winter days by selecting extreme environmental parameter values from the central Arctic (see Supplementary Material - Methods Table S0). Subsequently, we evaluate the energy landscape functions as above with the respective environmental parameters. Interestingly, simulating communities of the summer cluster under Arctic summer or winter conditions results in similar energy values as placing these communities in Atlantic winter days. Communities found in summer are highly unstable whereas communities present in winter show, as a tendency, increased stability (Supplementary Material - Results Figure S13 E,G,I - left panel, bottom three), indicating that summer cluster communities will face challenges to adapt to Arctic conditions. When comparing the ASVs, which are found in at least one community configuration predicted as stable, we observe that ASVs stable in Arctic environments are almost completely different from those stable in their natural (Altlantic) environment (Supplementary Material - Results Figure S14, Table S9 and S10). Interestingly, one ASV predicted to be stable under Arctic conditions is the phototroph *Chaetoceros brevis*, which agrees with the fact that Chaetoceros species are preferentially found under ice (Booth and Horner, 1997). To estimate the importance of the ASVs identified to be stable under natural and Arctic conditions, respectively, we determine their average closeness centrality (Supplementary Material - Results Figure S15). Clearly, ASVs stable in their natural conditions (and in Atlantic conditions) display a significantly (p-value = $6x10^{-5}$) higher centrality than average, while those predicted to be stable under Arctic conditions show a reduced centrality. Likewise, the average NMI score for outgoing edges, indicating the mean influence of an ASV on other ecosystem members, is higher than average for ASVs stable under natural (and Atlantic) conditions, but lower than average for ASVs predicted to be stable under Arctic conditions. The reduced connectivity and weaker influence on other ecosystem members suggests that species of the summer cluster, when exposed to Arctic conditions, will play a less important role in determining ecosystem dynamics and stability.

Simulating the effect of Arctic conditions on winter cluster communities reveals that neither communities present in summer nor winter display a pronounced stability. The only exception appear to be communities during spring 2020. Remarkably, this period was characterized by unusually low temperatures (Figure 4 A). In general the ASVs identified as members of stable communities in the original conditions, Altlantic winter, or Arctic winter show a great similarity, with the majority (17 of 36) found in stable communities under all three conditions. These observations suggests that communities of the winter cluster might easily adapt to colder Arctic conditions, which is also supported by the observation that closeness centrality and averaged NMI scores (for outgoing edges) are almost unchanged for these three groups of organisms.

4 DISCUSSION

In this study, we developed a new method for investigating ecological time series data based on 18S metabarcoding derived abundance information, by combining three data analysis methods: Co-Occurrence Networks (CON) (Stephens et al., 2009), Convergent Cross Mapping (CCM) (Sugihara et al., 2012), and

Energy Landscape Analysis (ELA) (Suzuki et al., 2021). By integrating three methodological approaches, we aimed to identify and characterize keystone species in the West Spitsbergen Current across different seasons, environmental parameters and in relation to other organisms. In addition, we examine how these keystone species are affected by changing environmental conditions, providing insights into potential responses to Arctic warming and Atlantification. We also investigate how different taxa groups affect other taxa groups, and how their effects vary with seasonal shifts and environmental factors.

We generated a co-occurrence network using a novel similarity measure based on Fourier decomposition. This results in a co-occurrence network in which species cluster together, which occurred at similar times of the year. This approach differs from previous methods that rely directly on the raw time series signals (Ma et al., 2016; Lima-Mendez et al., 2015; Ma et al., 2020). By utilizing Fourier decomposition, we generated comprehensive temporal profiles for each ASV, providing a more nuanced understanding of co-occurrence dynamics. Consequently, the constructed network accurately reflects the temporal dynamics of each measured species' abundance, all in all revealing a clear seasonal cycle (see Figure 2). The taxonomic composition of the communities defined by the partitioning of the co-occurrence network is a reflection of the prevailing food-web (Dunne et al., 2002): it starts in spring (08TS) with photosynthetic primary producers such as Bacillariophyta and continues through the summer (09HS,10HS), while mixotrophs gain the upper hand in autumn (01TA,02TA, 03LW), leading to almost exclusively heterotrophic groups in winter dominated by the heterotrophic Syndiniales (04LW,05LW,07TS). It is notable that the spring clusters were taxonomically different from the other clusters (Supplementary Material - Results Figure S2). This diversity can be explained by the rapid environmental changes during this period. Early spring clusters have high connections to winter clusters and are dominated by mixotrophic Dinophyceae, while late spring clusters (02TA) show more similarities to summer clusters (Stoecker et al., 2017). The "late" spring cluster (08TS) reflects a bloom of primary producers with Ochrophyta (especially Bacillariophyta species), the dominant family (Obertegger et al., 2022). It is remarkable that Dinophyceae are present in almost all seasons, with the exception of the low light period, which is predominantly occupied by heterotrophic Syndiniales. This indicates that these mixotrophic organisms play a crucial role particularly in the intermediate phases of spring and autumn (Jassey et al., 2015; Bruhn et al., 2021; Mitra et al., 2014).

We used Co-Occurrence Networks (CON), Convergent Cross Mapping (CCM), and Energy Landscape Analysis (ELA) to analyze the complex dynamics of Arctic microbial communities. According to traditional ecological theory, keystone species are often defined as those with the most biomass. It should be noted that there are different seasonal progressions due to the annual cycle (von Appen et al., 2021; Wietz et al., 2021). However, this subjective evaluation does not necessarily indicate a central role in the ecosystem. Here, we use network analysis methods to define keystone species as those with the most significant role in the interaction network. This approach allows us to identify keystone species not only in summer but also in winter and other seasons. Our research has shown that indeed some of the keystone species show a high abundance, for example Fragilariopsis or Chaetoceros. However, in addition to these common species, there are also other species that may not have been detected using more traditional methods.

However, temporal co-occurrence does not imply interactions. To understand the causal influence between species we apply Convergent Cross-Mapping to derive a network with directed edges, whose weight indicate the strength of influence between ASVs. In general, most of the directed inter-cluster links are found between neighboring clusters. Our analysis revealed that by far not all species pairs, which are present at the same time, actually influence each other (see Fig. 3). A striking example is between clusters 06TS and 05LW, which were closely connected in the CON, while there are hardly any connections in the CCM network (see Fig. 3). This co-occurrence without apparent causal connections could be explained by unique environmental conditions shaping both of these clusters, such as polar water influx. Even more pronounced is the separation of cluster 03LW, mainly heterotrophs, and 08TS, mainly phototrophs, which are tightly connected by co-occurrence but show not a single causal link in the CCM network . The organisms in these two clusters are primarily influenced by environmental parameters, particularly light. Additionally, these photosynthetic and heterotrophic organisms are sometimes preyed upon by the same predators (Zhao et al., 2022) such as Syndiniales. This explains the simultaneous occurrence and similar seasonality of these phytoplankton taxa but indicates that they do not have a direct influence on each other. The lack of mutual influence during the transition from polar night to day is clearly visible in the CCM network (see Fig. 3). We interpret this gap between winter and spring clusters as a 'winter reset' (Supplementary Material - Results Figure S11). This phase is characterized by the predominance of parasites and Dinophycea. With the emergence of light, a new period of growth of primary producer

communities begins, shaped by the prevailing environmental conditions. These conditions determine which species will prevail in the future. By reflecting causal interactions between species, the CCM network even stronger reflects the cyclic microbiome structure than the co-occurrence network. The cycle begins with photoautotrophs (cluster 08TS) in early spring and ends with the hetero- and mixotrophs (cluster 01TA) in late autumn. As light intensity decreases, mixotrophs become more prevalent than photoautotrophs, leading to a shift towards a heterotrophic lifestyle and a transition from carbon fixation to consumption. This transition into a low light period is characterized by parasitic species, suggesting an ”eat and be eaten” scenario. The causal links from fall to winter are significantly less than between other seasons (except winter to spring). All of these causal links are related to Syndiniales, which can be explained by the parasitic lifestyle of Syndiniales, foraging upon mixotrophic species that are active during the transition autumn phase.

We also contrast our approach with two previous studies using Cross Convergence Mapping. While our approach focuses on the specific interactions within and between clusters, Ushio et al. provides a more general framework for predicting community diversity based on interaction capacity, temperature and abundance. The emphasis on mechanistic explanations for observed ecological patterns distinguishes the two approaches. Our methodology provides a comprehensive understanding of keystone species in a specific context, while Ushio’s study provides broader insights into the factors influencing community diversity in different ecosystems. Both studies use similar techniques such as correlation and CCM (Ushio, 2022). Fujita’s study used controlled experiments with six isolated community replicates, subjected to diverse treatments over 110 days. Regarding Takens’ Theorem and Convergence Cross Mapping, Fujita et al. used Simplex projection to forecast population size, while our study utilized pairwise Convergence Cross Mapping on ASV time series signals within clusters to predict potentially keystone species (Fujita et al., 2023).

The reconstructed landscape for autumn cluster 01TA shows that autumn communities are highly stable. However, for the winter cluster 03LW, the energy values of observed communities lack clear separation, making the situation less straightforward. Communities of the winter clusters which are still present in summer tend to display high energy values, indicating instability in winter conditions. The spring cluster 08TS shows an even more notable trend, indicating that community structures lack stability and exhibit high plasticity under spring conditions (Supplementary Material - Results Figure S12). This variability results in the coexistence of numerous community compositions during spring. This observation aligns with the findings of the CCM analysis, which highlight the inability to predict spring community compositions based on winter data. Finally, the analysis of the energy landscape for summer communities shows a distinct minimum, indicating well-defined community structures and high stability under summer conditions (Figure 5).

Taxonomically, the summer keystone species were mainly composed of five types of Ochrophyta, two types of Dinophyceae, and two types of Ciliophora. The most abundant species within this cluster were Fragilariopsis and Thalassiosira. In contrast, the winter keystone species consisted of three Syndiniales, one Radiolaria, and one Dinophyceae. The autumn cluster comprised two Ochrophyta, one Chlorophyta, and one Syndiniales. No keystone species were found for the spring period, indicating that it may be the most variable and selective season, as supported by the energy landscape analysis. Keystone species represent the ecological roles played by network members in primary production, consumption, and parasitic interactions. During the beginning of autumn, Chlorophyta emerges as a keystone species, indicating a shift in primary production from Ochrophyta to Chlorophyta. This shift may be explained by the preference of Chlorophyta for colder temperatures and a better adaptation to nutrient limitation.

Energy landscape analysis was used to assess the stability of observed communities when exposed to environmental conditions different from their typical settings. Our results showed that communities maintained stable configurations during typical summer (for communities of the summer clusters) or winter (for winter clusters) days, indicating their adaptability to seasonal variations. However, significant changes in energy values occurred when communities were assessed under opposing conditions, suggesting diverse responses to environmental shifts. Subsequent simulations of typical Arctic conditions revealed interesting patterns within the microbial communities. The summer cluster communities showed a clearly decreased stability under Arctic conditions, in contrast to the winter cluster communities, which displayed a tendency towards increased stability. This suggests that winter communities may have a greater capacity to adapt to colder Arctic conditions compared to their summer counterparts (Supplementary Material - Results Figure S15). The Arctic projection notably harbored a higher number of unique species, indicating varied responses

to environmental changes. Notably, the absence of shared keystone species among different datasets in the summer cluster 10HS suggests a lower robustness of Atlantic ASVs to environmental shifts, with keystone species candidates exhibiting variability. Under Atlantic environmental conditions, the closeness centrality of the summer cluster 10HS increased more compared to the Arctic setup. In the winter cluster community 03LW, closeness centrality is very similar in both projections (Supplementary Material - Results Figure S15).

The results presented in this study not only have practical implications for ecosystem management by improving our understanding and ability to predict change in complex ecological systems but also provide systematic and mechanistic insights into the mechanisms responsible for shaping and maintaining spatiotemporal heterogeneity in ecosystem composition (Suzuki et al., 2020).

CONFLICT OF INTEREST STATEMENT

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

AUTHOR CONTRIBUTION

WJvA, CB, MW and KM are responsible for the sampling design. KM contributed 18S metabarcoding data. MW processed raw metabarcoding reads into ASVs. WJvA contributed oceanographic data. EO devised the project, the main conceptual ideas and study outline. EO and RMK designed the model and the computational framework and analysed the data. EO, RMK, OP and KM wrote the initial draft. RMK carried out the implementation. EO, RMK and OP interpreted the data and conceptualized the manuscript. OE and OP advised data evaluation and data interpretation and revised and finalized the manuscript. All authors contributed to the final manuscript, the scientific interpretation and the discussion of results.

ACKNOWLEDGEMENTS

Computational infrastructure and support were provided by the Centre for Information and Media Technology at Heinrich Heine University Düsseldorf. Thanks to Kenta Suzuki for supporting us in adapting rELA to our needs. Thanks to Taylor Priest providing the PAR data. We thank Theresa Hargesheimer, Jana Bäger, Jakob Barz and Daniel Scholz for the RAS support. Moreover we thank the captains and crews of RV Polarstern for excellent support at sea, and the chief scientists for leading the various expeditions conducted for this study. Ship time for RV Polarstern was provided under grants AWI_PS99_00, AWI_PS100_01, AWI_PS107_05, AWI_PS114_01, AWI_PS121_01 of RV Polarstern. The work was supported by the infrastructure project FRAM (Frontiers in Arctic Marine Monitoring) funded by the Helmholtz Association. Our special thanks go to S. Neuhaus for bioinformatic support, Kerstin Korte and Swantje Ziemann for excellent technical support in the laboratory as well as Martina Löbl for coordination of the FRAM-project.

FUNDING

This work was further supported by The Deutsche Forschungsgemeinschaft (DFG) under grant number EB 418/6-1 (From Dusk till Dawn) and under Germany's Excellence Strategy - EXC-2048/1 - project ID 390686111 (CEPLAS)(EO and OE).

SUPPLEMENTARY MATERIAL - METHODS

SUPPLEMENTARY MATERIAL - RESULTS

DATA AVAILABILITY STATEMENT

The python codes will be available at `https://github.com/rakro101/otter.git` after acceptance.

REFERENCES

Arrigo, K. R., Perovich, D. K., Pickart, R. S., Brown, Z. W., Van Dijken, G. L., Lowry, K. E., et al. (2012). Massive phytoplankton blooms under arctic sea ice. *Science* 336, 1408–1408

Arrigo, K. R., van Dijken, G. L., and Strong, A. L. (2015). Environmental controls of marine productivity hot spots around a ntarctica. *Journal of Geophysical Research: Oceans* 120, 5545–5565

Assmy, P., Fernández-Méndez, M., Duarte, P., Meyer, A., Randelhoff, A., Mundy, C. J., et al. (2017). Leads in arctic pack ice enable early phytoplankton blooms below snow-covered sea ice. *Scientific reports* 7, 40850

Barber, D. G., Hop, H., Mundy, C. J., Else, B., Dmitrenko, I. A., Tremblay, J.-E., et al. (2015). Selected physical, biological and biogeochemical implications of a rapidly changing arctic marginal ice zone. *Progress in Oceanography* 139, 122–150

Blondel, V. D., Guillaume, J.-L., Lambiotte, R., and Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment* 2008, P10008

Bluhm, B. A., Gradinger, R., and Hopcroft, R. R. (2011). Arctic ocean diversity: synthesis. *Marine Biodiversity* 41, 1–4

Bluhm, B. A., Swadling, K. M., and Gradinger, R. (2017). Sea ice as a habitat for macrograzers. *Sea ice* , 394–414

Booth, B. C. and Horner, R. A. (1997). Microalgae on the arctic ocean section, 1994: species abundance and biomass. *Deep Sea Research Part II: Topical Studies in Oceanography* 44, 1607–1622

Bruhn, C. S., Lundholm, N., Hansen, P. J., Wohlrab, S., and John, U. (2021). Transition from a mixotroph/heterotroph protist community during the dark winter to a photoautotrophic spring community in arctic surface waters

Cardinale, B. J., Duffy, J. E., Gonzalez, A., Hooper, D. U., Perrings, C., Venail, P., et al. (2012). Biodiversity loss and its impact on humanity. *Nature* 486, 59–67

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* 134, 181–189

Clement Kinney, J., Maslowski, W., Osinski, R., Jin, M., Frants, M., Jeffery, N., et al. (2020). Hidden production: On the importance of pelagic phytoplankton blooms beneath arctic sea ice. *Journal of Geophysical Research: Oceans* 125, e2020JC016211

Duffy, J. E. (2009). Why biodiversity is important to the functioning of real-world ecosystems. *Frontiers in Ecology and the Environment* 7, 437–444

Dunne, J. A., Williams, R. J., and Martinez, N. D. (2002). Network structure and biodiversity loss in food webs: robustness increases with connectance. *Ecology letters* 5, 558–567

Eppley, R. W. and Peterson, B. J. (1979). Particulate organic matter flux and planktonic new production in the deep ocean. *Nature* 282, 677–680

Fernández-Méndez, M., Katlein, C., Rabe, B., Nicolaus, M., Peeken, I., Bakker, K., et al. (2015). Photosynthetic production in the central arctic ocean during the record sea-ice minimum in 2012. *Biogeosciences* 12, 3525–3549

Fortier, M., Fortier, L., Michel, C., and Legendre, L. (2002). Climatic and biological forcing of the vertical flux of biogenic particles under seasonal arctic sea ice. *Marine Ecology Progress Series* 225, 1–16

Frey, K. (2017). Arctic ocean primary productivity. *In: Arctic Report Card 2017, NOAA, http://www. arctic. noaa. gov/Report-Card/Report-Card-2017/ArtMID/7798/ArticleID/701/Arctic-Ocean-Primary-Productivity*

Fujita, H., Ushio, M., Suzuki, K., Abe, M. S., Yamamichi, M., Iwayama, K., et al. (2023). Alternative stable states, nonlinear behavior, and predictability of microbiome dynamics. *Microbiome* 11, 63

Gosselin, M., Levasseur, M., Wheeler, P. A., Horner, R. A., and Booth, B. C. (1997). New measurements of phytoplankton and ice algal production in the arctic ocean. *Deep Sea Research Part II: Topical Studies in Oceanography* 44, 1623–1644

Guimera, R. and Nunes Amaral, L. A. (2005). Functional cartography of complex metabolic networks. *nature* 433, 895–900

Hardge, K., Peeken, I., Neuhaus, S., Lange, B. A., Stock, A., Stoeck, T., et al. (2017). The importance of sea ice for exchange of habitat-specific protist communities in the central arctic ocean. *Journal of Marine Systems* 165, 124–138

Hooper, D. U., Adair, E. C., Cardinale, B. J., Byrnes, J. E., Hungate, B. A., Matulich, K. L., et al. (2012). A global synthesis reveals biodiversity loss as a major driver of ecosystem change. *Nature* 486, 105–108

Ives, A. R. and Carpenter, S. R. (2007). Stability and diversity of ecosystems. *science* 317, 58–62

Jassey, V. E., Signarbieux, C., Hättenschwiler, S., Bragazza, L., Buttler, A., Delarue, F., et al. (2015). An unexpected role for mixotrophs in the response of peatland carbon cycling to climate warming. *Scientific reports* 5, 16931

Javier, P. J. E. A., Liponhay, M. P., Dajac, C. V. G., and Monterola, C. P. (2022). Causal network inference in a dam system and its implications on feature selection for machine learning forecasting. *Physica A: Statistical Mechanics and its Applications* 604, 127893

Joyce, K. E., Laurienti, P. J., Burdette, J. H., and Hayasaka, S. (2010). A new measure of centrality for brain networks. *PloS one* 5, e12200

Kraft, A., Nöthig, E.-M., Bauerfeind, E., Wildish, D. J., Pohle, G. W., Bathmann, U. V., et al. (2013). First evidence of reproductive success in a southern invader indicates possible community shifts among arctic zooplankton. *Marine Ecology Progress Series* 493, 291–296

Leu, E., Søreide, J., Hessen, D., Falk-Petersen, S., and Berge, J. (2011). Consequences of changing sea-ice cover for primary and secondary producers in the european arctic shelf seas: timing, quantity, and quality. *Progress in Oceanography* 90, 18–32

Lewis, K., Van Dijken, G., and Arrigo, K. R. (2020). Changes in phytoplankton concentration now drive increased arctic ocean primary production. *Science* 369, 198–202

Lima-Mendez, G., Faust, K., Henry, N., Decelle, J., Colin, S., Carcillo, F., et al. (2015). Determinants of community structure in the global plankton interactome. *Science* 348, 1262073

Lin, I., Liu, W. T., Wu, C.-C., Wong, G. T., Hu, C., Chen, Z., et al. (2003). New evidence for enhanced ocean primary production triggered by tropical cyclone. *Geophysical Research Letters* 30

Lin, Y.-C., Chin, C.-P., Yang, J. W., Chiang, K.-P., Hsieh, C.-h., Gong, G.-C., et al. (2022). How communities of marine stramenopiles varied with environmental and biological variables in the subtropical northwestern pacific ocean. *Microbial ecology* , 1–13

Loreau, M. and De Mazancourt, C. (2013). Biodiversity and ecosystem stability: a synthesis of underlying mechanisms. *Ecology letters* 16, 106–115

Ma, B., Wang, H., Dsouza, M., Lou, J., He, Y., Dai, Z., et al. (2016). Geographic patterns of co-occurrence network topological features for soil microbiota at continental scale in eastern china. *The ISME journal* 10, 1891–1901

Ma, B., Wang, Y., Ye, S., Liu, S., Stirling, E., Gilbert, J. A., et al. (2020). Earth microbial co-occurrence network reveals interconnection pattern across microbiomes. *Microbiome* 8, 1–12

Maat, D. S., Crawfurd, K. J., Timmermans, K. R., and Brussaard, C. P. (2014). Elevated co2 and phosphate limitation favor micromonas pusilla through stimulated growth and reduced viral impact. *Applied and Environmental Microbiology* 80, 3119–3127

Mann, D. G., Crawford, R., Round, F., et al. (2017). Bacillariophyta. *Handbook of the Protists* 7, 205–266

Meredith, M., Sommerkorn, M., Cassota, S., Derksen, C., Ekaykin, A., Hollowed, A., et al. (2019). Polar regions. *Special report on ocean and cryosphere in a changing chapter Intergovernmental Panel on Climate Change (IPCC). In Chapter 3 Polar Issues*

Mitra, A., Flynn, K. J., Burkholder, J. M., Berge, T., Calbet, A., Raven, J. A., et al. (2014). The role of mixotrophic protists in the biological carbon pump. *Biogeosciences* 11, 995–1005

Mori, A. S., Furukawa, T., and Sasaki, T. (2013). Response diversity determines the resilience of ecosystems to environmental change. *Biological reviews* 88, 349–364

Nöthig, E.-M., Ramondenc, S., Haas, A., Hehemann, L., Walter, A., Bracher, A., et al. (2020). Summertime chlorophyll a and particulate organic carbon standing stocks in surface waters of the fram strait and the arctic ocean (1991–2015). *Frontiers in Marine Science* 7, 350

Obertegger, U., Flaim, G., Corradini, S., Cerasino, L., and Zohary, T. (2022). Multi-annual comparisons of summer and under-ice phytoplankton communities of a mountain lake. *Hydrobiologia* 849, 4613–4635

Oldenburg, E., Popa, O., Wietz, M., von Appen, W.-J., Torres-Valdes, S., Bienhold, C., et al. (2024). Sea-ice melt determines seasonal phytoplankton dynamics and delimits the habitat of temperate Atlantic taxa as the Arctic Ocean atlantifies. *ISME Communications* 4, ycae027. doi:10.1093/ismeco/ycae027

Priest, T., von Appen, W.-J., Oldenburg, E., Popa, O., Torres-Valdés, S., Bienhold, C., et al. (2023). Atlantic water influx and sea-ice cover drive taxonomic and functional shifts in arctic marine bacterial communities. *The ISME Journal* 17, 1612–1625

Richter-Menge, J. A. and Farrell, S. L. (2013). Arctic sea ice conditions in spring 2009–2013 prior to melt. *Geophysical Research Letters* 40, 5888–5893

Sala, O. E., Stuart Chapin, F., Armesto, J. J., Berlow, E., Bloomfield, J., Dirzo, R., et al. (2000). Global biodiversity scenarios for the year 2100. *science* 287, 1770–1774

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome research* 13, 2498–2504

Soltwedel, T., Bauerfeind, E., Bergmann, M., Budaeva, N., Hoste, E., Jaeckisch, N., et al. (2005). Hausgarten: multidisciplinary investigations at a deep-sea, long-term observatory in the arctic ocean. *Oceanography*

Spreen, G., Kaleschke, L., and Heygster, G. (2008). Sea ice remote sensing using amsr-e 89-ghz channels. *Journal of Geophysical Research: Oceans* 113

Stephens, C. R., Heau, J. G., González, C., Ibarra-Cerdeña, C. N., Sánchez-Cordero, V., and González-Salazar, C. (2009). Using biotic interaction networks for prediction in biodiversity and emerging diseases. *PLoS One* 4, e5725

Stoecker, D. K., Hansen, P. J., Caron, D. A., and Mitra, A. (2017). Mixotrophy in the marine plankton. *Annual Review of Marine Science* 9, 311–335

Strass, V. H. and Nöthig, E. M. (1996). Seasonal shifts in ice edge phytoplankton blooms in the barents sea related to the water column stability. *Polar Biology* 16, 409–422

Sugihara, G., May, R., Ye, H., Hsieh, C.-h., Deyle, E., Fogarty, M., et al. (2012). Detecting causality in complex ecosystems. *science* 338, 496–500

Suter, E. A., Pachiadaki, M., Taylor, G. T., and Edgcomb, V. P. (2022). Eukaryotic parasites are integral to a productive microbial food web in oxygen-depleted waters. *Frontiers in Microbiology* 12, 764605

Suzuki, K., Nagaoka, S., Fukuda, S., and Masuya, H. (2020). Energy landscape analysis of ecological communities elucidates the phase space of community assembly dynamics. *Ecological Monographs*

Suzuki, K., Nakaoka, S., Fukuda, S., and Masuya, H. (2021). Energy landscape analysis elucidates the multistability of ecological communities across environmental gradients. *Ecological Monographs* 91, e01469

Thaler, M. and Lovejoy, C. (2014). Environmental selection of marine stramenopile clades in the arctic ocean and coastal waters. *Polar biology* 37, 347–357

Tragin, M. and Vaulot, D. (2018). Green microalgae in marine coastal waters: The ocean sampling day (osd) dataset. *Scientific Reports* 8, 14020

Ushio, M. (2022). Interaction capacity as a potential driver of community diversity. *Proceedings of the Royal Society B* 289, 20212690

van Leeuwe, M. A., Fenton, M., Davey, E., Rintala, J.-M., Jones, E. M., Meredith, M. P., et al. (2022). On the phenology and seeding potential of sea-ice microalgal species. *Elem Sci Anth* 10, 00029

Veilleux, B. (1979). An analysis of the predatory interaction between paramecium and didinium. *The Journal of Animal Ecology* , 787–803

von Appen, W.-J., Waite, A. M., Bergmann, M., Bienhold, C., Boebel, O., Bracher, A., et al. (2021). Sea-ice derived meltwater stratification slows the biological carbon pump: results from continuous observations. *Nature Communications* 12, 7309

von Appen, W.-J., Wekerle, C., Hehemann, L., Schourup-Kristensen, V., Konrad, C., and Iversen, M. H. (2018). Observations of a submesoscale cyclonic filament in the marginal ice zone. *Geophysical Research Letters* 45, 6141–6149

Von Quillfeldt, C. (2000). Common diatom species in arctic spring blooms: their distribution and abundance. *Botanica Marina*

Wietz, M., Bienhold, C., Metfies, K., Torres-Valdés, S., von Appen, W.-J., Salter, I., et al. (2021). The polar night shift: seasonal dynamics and drivers of arctic ocean microbiomes revealed by autonomous sampling. *ISME Communications* 1, 76

Yang, E. J., Ha, H. K., and Kang, S.-H. (2015). Microzooplankton community structure and grazing impact on major phytoplankton in the chukchi sea and the western canada basin, arctic ocean. *Deep Sea Research Part II: Topical Studies in Oceanography* 120, 91–102

Zhao, H., Zhang, Z., Nair, S., Zhao, J., Mou, S., Xu, K., et al. (2022). Vertically exported phytoplankton (¡ 20 μm) and their correlation network with bacterioplankton along a deep-sea seamount. *Frontiers in Marine Science* 9, 862494

Zhao, Q., Van den Brink, P. J., Xu, C., Wang, S., Clark, A. T., Karakoç, C., et al. (2023). Relationships of temperature and biodiversity with stability of natural aquatic food webs. *Nature Communications* 14, 3507