2020

Living Upstream: Kennebec River Influence on Nutrient Regimes and Phytoplankton Communities in Harpswell Sound

Siena Brook Ballance
Bowdoin College

Living Upstream: Kennebec River Influence on Nutrient Regimes and Phytoplankton Communities in Harpswell Sound

An Honors Paper for the Department of Earth and Oceanographic Science

By Siena Brook Ballance

Bowdoin College, 2020

Acknowledgements

This work would not have been possible without the help from my advisor, Collin Roesler and my two Honors committee members, Michelle Fame and Michèle LaVigne. Thank you for all of your help and support this year and throughout my time in the Earth and Oceanographic Science Department at Bowdoin College.

Table of Contents

Abstract

Phytoplankton underpin marine trophic systems and biogeochemical cycles. Estuarine and coastal phytoplankton account for 40-50% of global ocean primary productivity and carbon flux making it critical to identify sources of variability. This project focuses on the Kennebec River and Harpswell Sound, a downstream, but hydrologically connected coastal estuary, as a case study of temperate river influence on estuarine nutrient regimes and phytoplankton communities. Phytoplankton pigments and nutrients were analyzed from water samples collected monthly at 8 main-stem rivers stations (2011-2013) and weekly in Harpswell Sound (2008-2017) during ice-free months. Spatial bedrock and land use impacts on river nutrients were investigated at sub-watershed scales using GIS. Spatial analysis reveals a 10-fold increase in measured phytoplankton biomass across the Kennebec River's saltwater boundary, which demonstrates ocean-driven phytoplankton variability in the lower river. The biomass pattern is accompanied by a transition in phytoplankton community structure with respect to which groups co-occur (diatoms, chlorophytes, and cryptophytes) and which are unique (dinoflagellates in Harpswell). Upstream, the timing of each community depends on land-use proximity and seasonal discharge. In Harpswell Sound, the nutrient regime and phytoplankton community structure vary systematically: first diatoms strip silicate, then dinoflagellates utilize nitrate, followed by chlorophytes and cryptophytes that utilize available phosphate. These findings reveal, for the first time, patterns in phytoplankton communities and nutrient dynamics across the fresh to salt water interface. Ultimately the Kennebec River phytoplankton communities and nutrient regimes are distinct, and the river is only a source of silicate to Harpswell Sound.

Introduction

As the base of the aquatic food chain, phytoplankton support all marine trophic systems and biogeochemical cycles in ocean and river ecosystems. The conduit between these two ecosystems, estuarine and coastal waters, accounts for 40-50% of global ocean primary productivity and subsequent carbon flux (Paerl, 1997). Coastal estuarine phytoplankton communities are influenced by the variability in nutrients delivered by salty ocean water masses and freshwater rivers. Increased development of urban and agricultural land along rivers adjoining costal estuaries has caused accelerated eutrophication and subsequent degradation of fisheries due to the alteration of phytoplankton communities (Paerl, 2006). Recent shifts in these phytoplankton communities and coastal fisheries demonstrate the importance in understanding

estuarine phytoplankton seasonal and interannual variations in response to quickly changing river systems.

Harpswell Sound and the Kennebec River

Harpswell Sound (HS) is a reverse and tidally impacted estuary in the Gulf of Maine (GOM) where the phytoplankton community varies in composition throughout the year, allowing for a diverse marine trophic environment and robust fishing industry. HS is a productive estuary with a median Total Chlorophyll a concentration of 3.30mg/m$_3$ with 4 dominant phytoplankton groups: chlorophytes, cryptophytes and dinoflagellates, haptophytes, and diatoms (Kramer and Siegel, 2019). Its freshwater source enters from the mouth of the estuary, the plume arising from the upstream Kennebec River (KR) (Figure 1). This reverse estuarine structure is relatively unique. The KR is deflected into the mouth of HS by Coriolis (Figure 1), creating strong stratification at the mouth of the estuary and a more mixed water column at the estuary head (Wolovick, 2009).

In addition to its unique structure, HS is a sentinel site for harmful algal blooms. Historically and presently, these blooms have been caused by *Alexandrium* spp., a dinoflagellate species that commonly causes toxic shellfish poisoning (Keafer, 2005; Barton et al., 2013; Hankinson, 2010; Townsend, Pettigrew and Thomas, 2005). In recent years, however, *Pseudo-nitzschia* spp. and *Dinophysis* spp., two diatom species that can cause amnesic shellfish poisoning and diarrhetic shellfish poisoning, have emerged in HS (HABON-NE, 2019-2023). This recent shift in phytoplankton community, along with the continued seasonal presence of toxic dinoflagellates demonstrate the importance in monitoring and gaining a better understanding of the river and ocean influence on HS phytoplankton and nutrients.

Harpswell Sound's freshwater source, the KR, flows from the headwaters of Moosehead Lake to Merrymeeting Bay where it then converges with five other rivers before flowing out to the coast and into the mouth of HS (Figure 1). Along this upstream to downstream path, the KR transports and transforms crucial elements that provide a primary nutrient source to the KR and potentially to HS phytoplankton communities (Xenopoulos et al., 2017).

River nutrients are derived from bedrock and land use, then utilized by river phytoplankton before entering the coastal waters (bold arrows, Figure 2). Ongoing chemical weathering of bedrock and surficial glacial sediment, which is greatly influenced by land use type, provides a primary source of dissolved silicate to rivers (Turner et al., 2003) (light arrows, Figure 2). Runoff from agricultural and urban land use represents a dominant nitrate and phosphate river nutrient source (Pratt and Chang, 2011; Sanders et al., 1997) (light arrows, Figure 2). Additionally, baseflow groundwater, containing nutrients from both bedrock and land use sources, is another important nutrient input into rivers and a direct input into coastal waters (Jung, 2020). Although the connection between river nutrients and coastal waters is not fully understood, it is clear that river phytoplankton nutrient utilization influences the concentration and composition of nutrients entering coastal waters (Devercelli and O'Farrell, 2013) (light arrows, Figure 2).

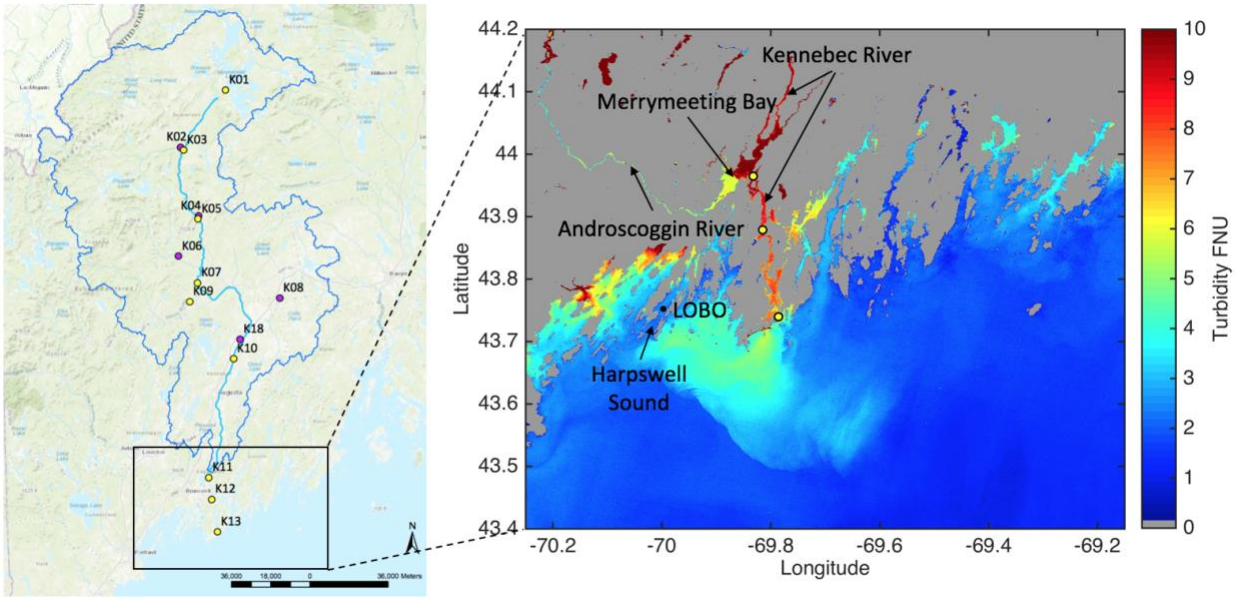


Figure 1. Map of the Kennebec River watershed (outlined in dark blue). Stations along the main stem (K01, K03, K05, K07, K10, K11, K12, K13) of the Kennebec River (light blue line) are labeled with a yellow dot and tributaries (K02, K04, K06, K08, K18) are labeled with a purple dot. Turbidity, the inset map (right), shows the convergence of the Kennebec and Androscoggin Rivers in Merrymeeting Bay and the path of the Kennebec River plume into Harpswell Sound (adapted from Figure 3 in Snyder et al., 2017). Yellow dots correspond to downstream stations K11, K12, and K13 on the map of the Kennebec watershed. The location of the LOBO buoy is also marked.

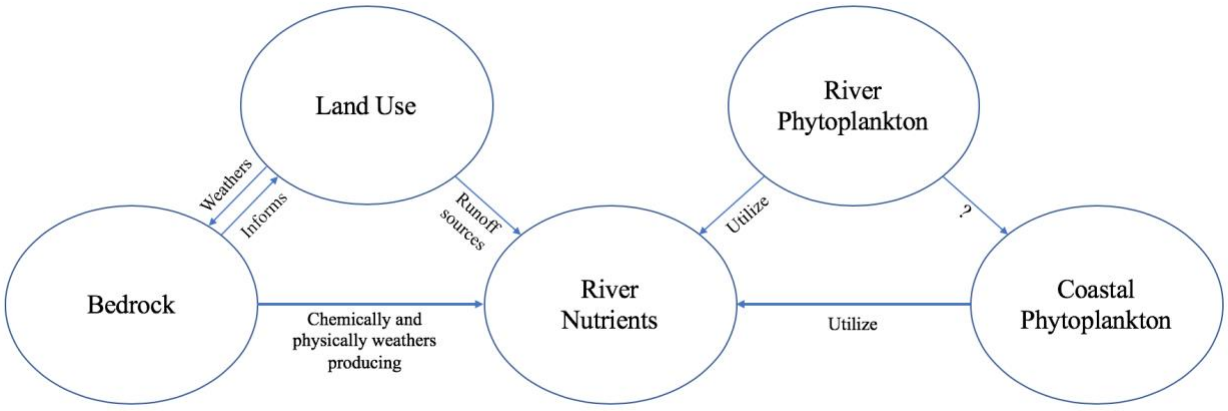


Figure 2. Schematic diagram of the influence of bedrock, land use, river nutrients, and river phytoplankton on coastal phytoplankton communities and the interplay between each. It is not yet understood how river phytoplankton influence or compare to coastal phytoplankton so the connection between is labeled with a question mark.

In previous studies of the GOM, rivers have been found to play a minor role in nutrient fluxes when compared to GOM water masses. In a study of nitrogen sources and cycling in the

GOM, Townsend, 1998 found that rivers only contribute $0.8x10^9$ gat N/year to the advective flux of nitrogen into and out of the GOM while the Scotian Shelf Waters and the Northeast Channel Slope Waters, the two primary water masses in the GOM, fluxed $31.5x10^9$ gat N/year and $147.9x10^9$ gat N/year, respectively. Based on these results, there have been extensive analyses conducted on ocean water mass dynamics in the GOM and water mass contribution to coastal water nutrients which then promote seasonal phytoplankton blooms.

Deep slope water, delivered to the GOM through the Northeast Channel, provides a primary source of ocean-derived nutrients to coastal waters (Townsend et al., 2010). These deep-water nutrients are mixed into the surface waters tidally and seasonally through winter convective mixing (Townsend et al., 2010; Ji et al., 2008). The nutrient enriched water column after the winter combined with the formation of a thermocline in the early spring allow for a large spring phytoplankton bloom (Ji et al., 2008). Although the nutrients are depleted in the summer, fall cooling and mixing promotes a smaller fall phytoplankton bloom (Song, 2010).

In addition to the importance of ocean water mass nutrient sources, an increase in anthropogenic nutrient runoff, transported to estuaries through river systems, has stimulated phytoplankton blooms throughout the year, altering this seasonal cycle (Fennel and Testa, 2019). Nutrient loading from rivers and subsequent eutrophication of coastal waters has been found to increase harmful algal blooms (HAB) (Anderson et al., 2008). In the GOM, paralytic shellfish poisoning (PSP) poses the greatest threat to local Maine fisheries (Mcgillicuddy et al., 2003). PSP events are almost always preceded by high discharge from the KR during the spring (Hunt et al., 2005). With the increasing threat of anthropogenic nutrient loading and eutrophication, it has become increasingly important to understand the impact rivers have on adjoining estuaries.

In the GOM shelf sea that is dominated by distributed river systems, we seek to reveal the drivers of one of the major river systems on the nutrient and phytoplankton regimes in the downstream coastal waters. As climate change warms the ocean and alters hydrologic patterns, understanding river to ocean connections is vital. We present an extensive analysis of bedrock, land use, nutrients, and phytoplankton to gain a deep understanding of the spatial and temporal variability in phytoplankton communities and nutrients within the KR. We then consider how these riverine communities and nutrients affect phytoplankton and nutrient variability in HS. In assessing the relationship between the KR and HS as a case study, results will help better understand how global rivers influence their coastal waters.

Questions and Hypotheses

In an effort to understand the spatial and temporal co-variance between the river and estuarine ecosystem nutrient regimes and phytoplankton communities, we present our questions and hypotheses in a spatial format. We begin with a broad analysis of the KR watershed bedrock and land use. We then consider how the bedrock and land use inform the nutrients observed in the river water column and how these nutrients influence patterns in river phytoplankton communities spatially and temporally (Figure 2). Phytoplankton community structure is assessed based on the functionality of phytoplankton present within an ecosystem as revealed by pigment-based taxonomic distinctions. Phytoplankton functionality is defined by their different biogeochemical roles; thus, tracking the community structure will help us track ecosystem function (IOCCG, 2006). Following the KR analysis, we transition downstream into HS and analyze the nutrient regime and phytoplankton community in the estuary using data collection years that overlap with the KR dataset and using the long-term HS dataset to track changes over a longer time period. Finally, we compare the river and estuarine phytoplankton communities to

gain an understanding of the connectivity between the KR and HS. Within this spatial and temporal format, we seek to answer the following questions:

1) What is the impact of bedrock and land use on nutrient regimes and phytoplankton community structure from upstream to downstream in the KR? How do the nutrients and phytoplankton communities change seasonally and interannually?
    a. Do silicate, nitrate, phosphate, and ammonium display the same spatial and temporal trends from upstream to downstream in the KR?
    b. Is the phytoplankton community structure the same throughout the KR and do the PFTs vary the same seasonally from upstream to downstream in the KR?
2) How does the phytoplankton community structure in HS vary seasonally in response to the nutrients? How has the HS phytoplankton community structure and nutrient regime shifted overtime from 2008-2017?
3) How does the phytoplankton community structure and nutrient regime in the KR compare to the nutrients and phytoplankton observed in HS? Further, are the nutrients and phytoplankton in HS driven by ocean variability, river variability, or both?

We present two contrasting hypotheses in response to these questions:

H1) There is complete connectivity between the coastal waters and the KR in which downstream trends in nutrients and phytoplankton communities are maintained across the freshwater-saltwater interface.

H2) Alternatively, there is no connectivity between the KR and HS, so the coastal waters can be considered a separate, ocean influenced ecosystem from the river ecosystem.

By addressing these questions and testing these hypotheses, this work will advance our understanding of the spatial and temporal connections between nutrients and phytoplankton in

the KR and HS. This will not only help us better understand if the HS phytoplankton communities are driven by river variability but will also help to evaluate and predict the impacts of regional bedrock weathering rates, land use changes, river phytoplankton and overall river nutrient input on coastal phytoplankton communities, in particular as the hydrologic patterns are predicted to change significantly in response to regional warming (Hayhoe et al., 2007).

Background

Phytoplankton Classes

Phytoplankton communities consist of different phytoplankton functional types (PFTs) that are grouped based on their biogeochemical role and their physiological characteristics (Kramer and Siegel, 2019; Lu et al., 2018). Each PFT carries out specific chemical processes (biogeochemical guilds) including, but not limited to, calcification, silicification, and nitrogen fixation (IOCCG, 2006). PFTs are used as proxies of ecosystem function and may reflect climate change and ocean carbon sequestration efficiency which makes them important to monitor (IOCCG, 2006).

Gulf of Maine, Harpswell Sound, and Kennebec River Phytoplankton

Diatoms, dinoflagellates, chlorophytes, and cryptophytes are four of the dominant PFTs in the Gulf of Maine (GOM) and in Harpswell Sound (HS). Litchman et al., 2015 provides a thorough assessment of each PFT. Diatoms, or phytoplankton silicifiers (20-200 μm) (Quéré et al., 2005), are part of the red algae superfamily and are a dominant PFT. As a member of the red algae superfamily, diatoms have low N:P ratios (10:1). However, diatoms require silicate in addition to iron and phosphate for growth, so they tend to dominate in nutrient-rich waters and are therefore responsible for most of the primary production during the spring bloom, when nutrients and light are abundant (Quéré et al., 2005; Kramer, Roesler and Sosik, 2018). Globally,

diatoms represent about 20-25% of primary productivity and due to their relatively large cell size, have a high sinking rate which exports carbon and silica to the deep ocean. Dinoflagellates, also belonging to the red algae superfamily, are part of the mixed PFT (2-200 μm) and are often related to more nutrient depleted environments (Barton et al., 2013; Hankinson, 2010). Since they have one of the largest cell sizes amongst the PFTs, they have slower growth rates, making them not as competitive for nutrients. Cryptophytes, a less dominant and less studied PFT, are also a member of the red algae superfamily and therefore also have a lower N:P ratio. Chlorophytes are a dominant member of the green algae superfamily. Phytoplankton within this family have higher C:N and N:P ratios than red algae, but do not perform any distinct biogeochemical functions like the other PFTs. Based on the Litchman et al., 2015 study and other related studies, there is a strong understanding of the PFTs within the GOM and HS.

It is important to remember, in the context of HS, that the change in diatom and dinoflagellate presence overtime is particularly important to monitor since each can cause, and has historically caused, harmful algal blooms in HS and in the GOM (Townsend, Pettigrew and Thomas, 2005; HABON-NE, 2019-2023).

Although there has not yet been a study assessing PFTs in the Kennebec River (KR), rivers have been found to be dominated by PFTs with faster growth rates. These include the diatom, cryptophyte, and chlorophyte PFTs described above (Reynolds, 1994; Bortolini and Bueno, 2013).

Controls on Phytoplankton Functional Types

Nutrients

Seasonal variations in nutrients, along with temperature and irradiance, influence which estuarine phytoplankton communities thrive at different times of year based on their specific

nutrient requirements (Barraquand et al., 2018). The Redfield Ratio (C_{106}:N_{15}:P_1) is often used to generally describe the elemental composition of plankton biomass (Tyrell, 2019). As cell size varies between different PFTs, the number of atoms for each element changes, but the proportion remains relatively constant (Tyrell, 2019). Silicate, as a limiting nutrient for diatoms, is often incorporated into this ratio (N_{15}:P_1:Si_{15}). Since diatoms require silicate for growth, but green and blue algal groups do not, when Si:P and Si:N ratios are high, diatoms outcompete and dominate the phytoplankton community structure (D Tilman et al., 1982). In the Lagus et al., 2004 study, diatoms were also found to dominate at high N:P ratios while green algal groups tended to favor intermediate N:P ratios. In contrast, dinoflagellates were often found to dominate in nutrient-limited environments because many of the species are mixotrophic. When dinoflagellates are smaller than diatoms, they have a relatively higher surface area to volume ratio which allows them to survive in nutrient-depleted conditions (Fogg, 1986).

River Flow Regime

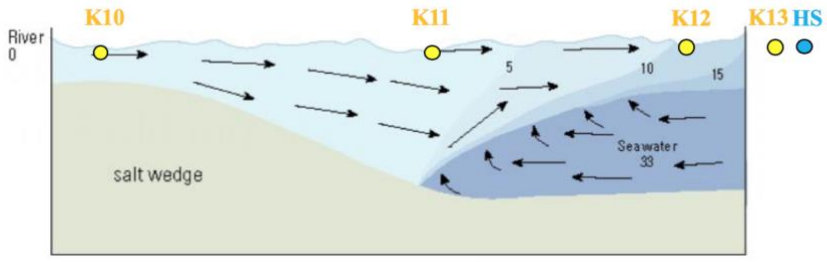
Although access to required nutrients plays a primary role in taxonomic composition of river phytoplankton, physical and hydrological factors have been found to be a key driver in river phytoplankton community structure (Descy et al., 2016; Chen et al., 2015). Physical/light limiting factors include turbidity, suspended sediment, and CDOM (Color Dissolved Organic Matter) and hydrological factors include discharge, flow velocity, water depth, and water residence time (Chen et al., 2015). With higher turbidity from suspended material and lack of vertical stratification, river phytoplankton are often more light limited than lake or ocean phytoplankton (Chetelat and Pick, 2006). These environmental factors set river phytoplankton apart from lake, estuarine, and ocean ecosystems. Due to constant river flow, turbulence, which resuspends sediment, promotes a phytoplankton community structure that is often dominated by

diatoms and green algal groups, taxonomic groups with faster growth rates (Reynolds, 1994; Bortolini and Bueno, 2013). In temperate rivers, diatoms have been found to dominate phytoplankton biomass due to their ability to adapt to changing flow regimes and low water residence times (Descy et al., 2016). Light levels have also been found to impact river phytoplankton size distribution (Chetelat and Pick, 2006). Lower light, deep, or turbid waters have been found to favor diatoms while a shallower water column and less light limitation allows for green algae growth (Chetelat and Pick, 2006; Descy et al., 2016).

River versus Ocean Influence on the Gulf of Maine and Harpswell Sound

Kennebec River Influence

The 240 km-long KR is among one of the largest rivers in the state of Maine with an average annual discharge of 452 m^3/year, draining an area of 14,790 km^2 (Kelley et al., 2011). The river flows out of Moosehead Lake, the largest lake in Maine, down through central Maine and out to the coast where it joins with 5 other rivers in Merrymeeting Bay before flowing into the coastal waters of the GOM. When the KR reaches the coast, a salt wedge is formed with a strong vertical and horizontal salinity gradient, limiting mixing between the fresh and salt water (Figure 3). Limited mixing allows the KR to enter HS as a freshwater plume.



Station	Salinity (psu)
K10	0
K11	0.02-1.6
K12	0.03-13.4
K13	0.19-31.6
HS	23.4-30.9

Figure 3. Schematic salinity salt wedge diagram (left) with the farthest downstream freshwater station (K10), tidally impacted stations (K11, K12, and K13), and HS marked. The diagram is adapted from https://www.amap.no/documents/doc/basic-circulation-and-salinity-distribution-in-salt-wedge-partially-mixed-well-mixed-and-fjord-type-estuaries-as-defined-by-wollast-and-duinker-1982-numbers-and-shading-show-salinity-values/340. Table (right) lists each station with each respective range in measured salinity.

Along this path from Moosehead Lake to HS, the KR accumulates nutrients from chemical and physical weathering of exposed bedrock and glaciomarine sediments and from anthropogenic land use runoff.

Bedrock – Silicate

Metasedimentary and igneous rocks from the Lower Paleozoic compose the bedrock beneath the KR watershed (Kelley et al., 2011). However, the retreat of the Late Wisconsinan Laurentide Ice Sheet between 16 kya and 12 kya left glacial deposits that scatter across this bedrock framework (Kelley et al., 2011). The upper KR watershed consists primarily of glacial till dated to the Late Pleistocene while the middle and lower watershed contains glaciomarine silt, clay, and sand that indicate the complex sea level history of coastal Maine (Kelley et al., 2011; Thompson and Borns, 1985). During glacial retreat, the depressed land surface was inundated with ocean water, allowing for the formation of these glaciomarine deltas (Kelley et al., 2011). The remaining high surface area, fine grained glaciomarine sediment along with bedrock exposure represent two important sources for weathering and subsequently, sources of nutrient input to the river water column.

The rate of silicate weathering is controlled by climate and specific silicate mineral compositions. The type and intensity of weathering within a certain region generally reflects the regional climate (Bierman and Montgomery, 2014). With an average annual temperature of ~8°C in Maine, similar rates of physical and chemical weathering occur (Bierman and Montgomery, 2014). Since bicarbonate and silica are not inputs associated with anthropogenic land use change, they are good indicators for chemical weathering (Bluth and Kump, 1994). Goldrich's Weathering Series establishes the stability of minerals at the Earth's surface. Minerals formed at high temperatures tend to be least stable at the Earth's surface when compared to minerals

forming at low temperatures (Bierman and Montgomery, 2014). This implies that granitic rocks are the most stable and most difficult to weather while ultra-mafic rocks are the least stable and first to weather. More generally, felsic rocks are harder to weather than mafic rocks (Lasaga et al., 1994).

Although silicate weathering is characteristic of many river systems, the KR displays above average silicate concentrations. In a study of global rivers, rivers at a similar latitude to the KR have average annual concentrations of dissolved silicate of about 1.2 uM (74.7 ug/L) while the KR has an average dissolved silicate concentration of about 28.7 uM (Turner et al., 2003). In this same study, silicate concentrations were found to be most dependent on discharge where higher discharge promotes lower silicate concentrations through dilution. Although the KR has seasonally high discharge in the spring, fall, and during dam releases in the summer, the presence of 6 dams along the KR (Figure 4) promote relatively higher silicate weathering (Sullivan et al., 2000). The impoundment formed behind a dam allows silicate to accumulate (Sullivan et al., 2000). Silicate accumulation promotes diatom blooms which leads to further biogenic silicate production and deposition (Sullivan et al., 2000). Ultimately, this promotes a silicate-rich water column along the transect of the KR. Highest silicate concentrations are observed in the winter when discharge is at a minimum, but silicate is never a limiting nutrient within the KR (Hunt et al., 2005).

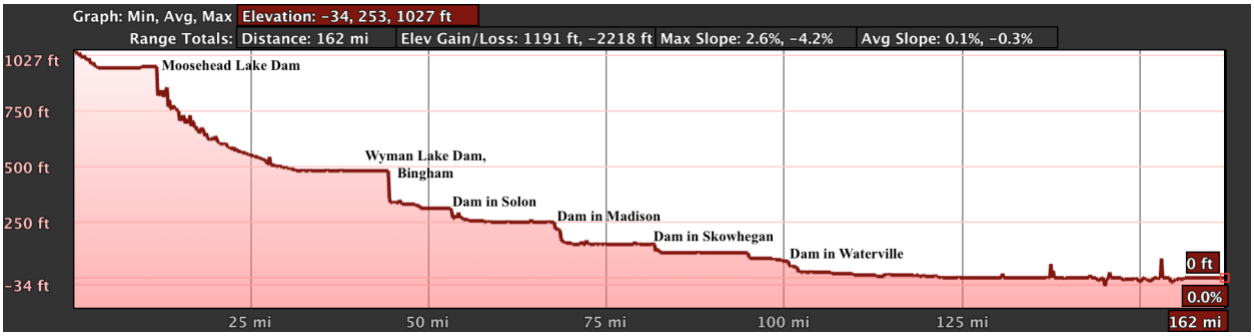


Figure 4. Google Earth generated elevation profile of the Kennebec River from the headwaters of Moosehead Lake to the coast. The 6 dams along the transect are labeled.

Land Use – Nitrate and Phosphate

In addition to bedrock and glacial sediment weathering, as open ecosystems, rivers depend heavily on the surrounding landscape and therefore, are drastically affected by changes in land use (Ileva et al., 2009). Particularly, the development of urban and agricultural land has been found to increase nutrient loads and runoff into rivers (Pratt and Chang, 2011; Sanders et al., 1997). Removal of forested land for urbanization reduces nutrient uptake and increases impervious surfaces, causing an increased nutrient flux to rivers (Mouri et al., 2011). Similarly, fertilizer use on agricultural lands increases river nitrate and phosphate loads (Pratt and Chang, 2011; Ileva et al., 2009).

Previous work by Hunt et al., 2005 provides important observations about nitrate and phosphate patterns in the KR. The KR experiences highest nitrate concentrations in the winter and spring periods, correlated with snowmelt runoff. Lowest concentrations in nitrate are observed in summer months since New England experiences the highest biological activity within the river during this time. Phosphate is limiting in all seasons and therefore, when urban land introduces phosphate into the water column, there is often a peak in biological activity. Spatially, the addition of point-source nutrient inputs from urban and agricultural land causes nitrate and phosphate concentrations to increase moving downstream. Based upon these results, we conclude that nutrient flux from urban and agricultural land along the KR shows both seasonal and spatial trends in relation to anthropogenic nutrient sources. These sources of nutrients include, but are not limited to, pulp/paper mills and 10 wastewater treatment plants that have been found to increase phosphate and subsequently, phytoplankton blooms.

Ocean Influence

In addition to river nutrient input, an alternation of characteristically different deep water masses provides a primary source of nutrients to phytoplankton in the GOM. The Labrador Slope Water (LSW), a cold fresh water mass, and Warm Slope Water (WSW), a relatively warmer saltier water mass, deliver nitrate and silicate in differing concentrations to the GOM (Townsend et al., 2010). WSW contains higher nitrate concentrations (>24 μM) than the LSW (15-17 μM) while both contain about 11-15 μM of silicate (Townsend et al., 2010). Relative amounts of each water mass may depend on the North Atlantic Oscillation (NAO) (Townsend et al., 2010; Smith et al., 2001). During high NAO years, an increase in LSW causes cooler, fresher surface waters that reflect an overall lower nutrient composition (Smith et al., 2001).

Biogeochemical nutrient cycling within GOM estuarine ecosystems also plays a role in fluxing nutrients into the water column. The structure of estuarine environments including insolation induced stratification and stratification from the amount of fresh and salt water influences how ocean nutrient cycling occurs within the estuary. In freshwater, phosphate generally limits primary production while nitrate is limiting in saltwater systems (Hartzell and Jordan, 2012). Since the amount of freshwater and saltwater shifts seasonally in estuarine ecosystems, the type of nutrient limitation also varies (Hartzell and Jordan, 2012).

Summary

While we understand the seasonal cycle and ocean influence on coastal phytoplankton communities in the GOM and HS which is crucial for predicting changes in upper trophic levels and fisheries, we do not yet understand phytoplankton communities and nutrient regimes in the KR and their impact on HS. This study will use a spatial and temporal analysis of phytoplankton

pigments and nutrients in the KR and HS as a case study to identify the relationship between river and coastal phytoplankton communities and nutrient regimes.

Methods

Water sampling for nutrients and phytoplankton pigments was conducted in two separate studies, one in the Kennebec River (KR) and one in Harspwell Sound (HS). Along the KR, samples were collected monthly from 2011-2013 during ice-free months March/April to November/December. Weekly sampling was conducted in HS from 2008-2017 during ice-free months March/April to October/November. To address the comparison between the KR and HS, the subset of HS data from 2011-2013 will be studied; however, the full-length time series for HS will also be assessed to understand changes within the estuary over the entirety of the HS dataset.

As part of this study, a GIS (Geographic Information System) analysis of the KR watershed and sub-watershed bedrock and land use was conducted to provide a background for interpreting the KR and HS trends in nutrient regime and phytoplankton community.

Sample Collection and Processing

Kennebec River

Eight stations (K01, K03, K05, K07, K10, K11, K12, K13) along the main stem of the KR (Figure 1) were used for bedrock, land use, nutrient, and pigment analysis. Stations are located near USGS (United States Geological Survey) gauging stations. Monthly surface water samples were collected from these main stem stations from May 17, 2011 to January 1, 2014, as part of a carbon study project conducted for NASA (Historical and Projected Changes in Carbon Export to the Gulf of Maine from Land Use and Climate Change) by Collin Roesler, Philip Camill, and

John Lichter. Water samples were syringe filtered through a 0.7 μm glass fiber filter and collected in dark, acid-washed bottles which were stored in a cooler for processing in the Bowdoin College lab. Filters were sent to NASA for filtrate analysis, generating nutrient data, and HPLC (High Pressure Liquid Chromatography) analysis, generating pigment concentration data for each station.

Harpswell Sound

All data used for HS was collected at the Land/Ocean Biogeochemical Observatory (LOBO) located at 43° 45.70N 69° 59.30W (Figure 1). Weekly water samples have been collected by Bowdoin students and faculty from 2008 to the present during ice-free months of February through November (monthly data changes with each year) from the LOBO buoy. Water samples are filtered at Bowdoin College and filters are sent to NASA for filtrate analysis, generating nutrient data, and HPLC analysis, generating pigment concentration data. For the purposes of this study, pigment and nutrient data from surface depths (0 to 2.5 m) from 2008-2017 was considered with an emphasis on water samples collected from a 2.5 m depth from 2011-2013 (the overlapping years with the KR data).

Phytoplankton Pigment as a Proxy for Phytoplankton Taxonomy

Phytoplankton pigments, obtained using HPLC, are used as a proxy for phytoplankton taxonomy (Table 1). HPLC is one of the most widespread methods available to measure phytoplankton taxonomic variability (Kramer and Siegel, 2019). Taxanomic groups of phytoplankton share similar sets of pigments, allowing for differentiation between taxonomy based on pigment.

Table 1. Primary phytoplankton pigments and their abbreviations, assumed taxonomic values, groupings, and taxonomic distributions. The pigments focused on in this analysis are highlighted in a color. Table cells are color coded based on phytoplankton groupings determined from the Harpswell Sound hierarchical cluster analysis results (Figure 16). Table 1 is adapted from Table 1 in Catlett and Siegel 2018.

Phytoplankton Grouping	Abbreviation	Pigment	Assumed Taxonomic Value	Taxonomic Distribution
Green Algae, Pelagophytes, Haptophytes, and Cyanobacteria	ButFuco	Butanoyloxyfucoxanthin	Pelagophytes	Diatoms, dictyophytes, pelagophytes, raphidophytes, haptophytes, dinoflagellates
	HexFuco	Hexanoyloxyfucoxanthin	Haptophytes	Haptophytes, dinoflagellates
	Zea	Zeaxanthin	Cyanobacteria, green algae	Cyanobacteria, rhodophytes, diatoms, chrysophytes, dictyophytes, eustigmatophytes, pelagophytes, raphidophytes, all green algae except euglenophytes
	TChlb	Total Chlorophyll b	All green algae	Cyanobacteria, green algae
Dinoflagellates and Cryptophytes	Perid	Peridinin	Dinoflagellates	Dinoflagellates
	Allo	Alloxanthin	Cryptophytes	Cryptophytes, dinoflagellates, chlorophytes
Red Algae and Diatoms	Caro	Carotene	All taxa	All algae except cryptophytes and rhodophytes
	TChlc	Total Chlorophyll c	All red algae	All red algae except rhodophytes and eustigmatophytes
	Diad	Diadinoxanthin	Red algae	Diatoms, bolidophytes, dictyophytes, pelagophytes, xanthophytes, haptophytes, dinoflagellates, euglenophytes
	Diato	Diatoxanthin	Red algae	Diatoms, bolidophytes, dictyophytes, pelagophytes, xanthophytes, haptophytes, dinoflagellates, euglenophytes
	Fuco	Fucoxanthin	Diatoms	Diatoms, bolidophytes, chrysophytes, dictylophytes, pelagophytes, raphidophytes, haptophytes, dinoflagellates

Data Analyses

KR watershed/sub-watershed bedrock and land use data were obtained from USGS. Bedrock and land use analyses within the KR watershed and sub-watersheds feeding into each station were processed using ArcMap GIS version 10.7. Matlab R2019a was used for pigment, nutrient, Hierarchical Cluster, and Empirical Orthogonal Function (EOF) analyses for both the KR and HS. Matlab R2019a was also used for a lag correlation analysis for the KR.

Prior to all pigment Hierarchical Cluster and pigment EOF analyses, pigment data was mean centered and normalized with respect to Total Chlorophyll a (Kramer and Siegel, 2019). For the nutrient EOF analyses, nutrient data was mean centered.

Lag Correlation Analysis

Before beginning the KR analyses, an observational lag correlation analysis was conducted to estimate the time between a peak discharge event at an upstream station to the time of peak runoff observed at a downstream station (Seyam and Othman, 2014). KR main stem sites with USGS gauging stations (K03, K05, K07, and K10) were used to calculate lag time. Following the same methodology as Seyam and Othman, 2014, discharge was plotted from January 1, 2011, to January 1, 2014, to gain a visual understanding of the lag time between stations (Figure 5). Discharge accumulated moving downstream, but there was a minimal lag time between stations. The amount of time between two peak discharge events that occurred during the 3-year time series was approximately 5 hours, indicating that water was traveling at a rate of about 18 km/hr. This implies that water traveling from Moosehead Lake reaches the coast on a scale of days to weeks. Since the KR data was collected monthly, pigment and nutrient analyses will be interpreted as snapshots of phytoplankton communities and nutrient regimes that are then flushed into the coastal waters.

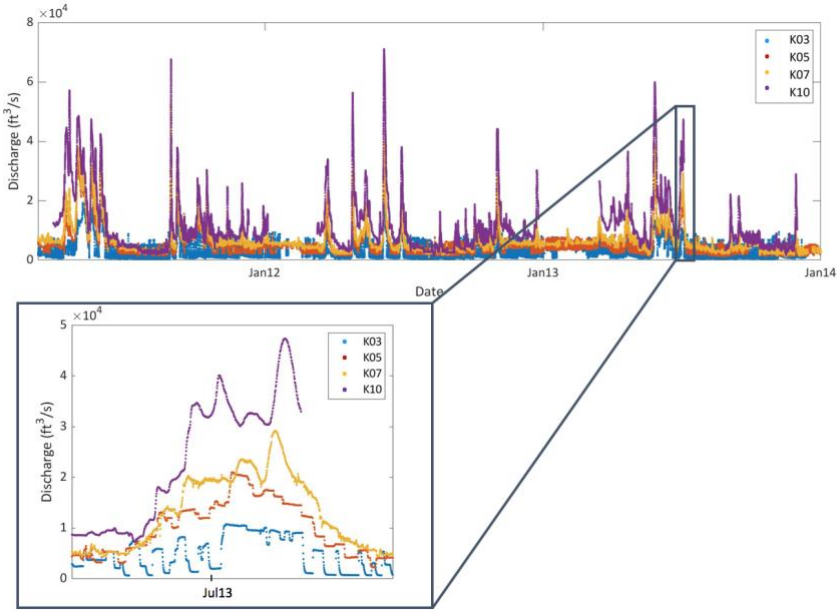


Figure 5. Hourly Kennebec River discharge from 2011-2013 measured at mainstem stations K03, K05, K07, and K10. Inset map displays a peak discharge event around July 2013.

Hierarchical Cluster Analysis

A hierarchical cluster analysis was performed on mean centered and normalized (with respect to TChla) phytoplankton pigments for KR mainstem stations K01-K10 (upstream freshwater stations), KR stations K11-K13 (downstream tidally impacted stations), and HS to identify the relatedness between taxa within each ecosystem. The hierarchical cluster analysis clusters the co-occurrences between pigments using a correlation distance 1-R (R=Pearson's correlation coefficient between pigments) and Ward's linkage method (Catlett and Siegel, 2018; Latasa and Bidigare, 1998).

Empirical Orthogonal Function (EOF) Analysis

To identify the space and time covariation of the phytoplankton communities identified in the cluster analysis, an EOF analysis was performed on mean centered and normalized (with respect to TChla) phytoplankton pigments for both the KR and HS from 2011-2013 in addition to surface (0-2.5 m) phytoplankton pigments from 2008-2017 for HS. The EOF function converts the pigments into a series of modes that represent the variability within each pigment dataset. The percent variance assigned to each mode decreases with increasing modes, meaning mode 1 describes the largest percent variance. Each mode is represented by loading values assigned to each pigment that indicate how much each pigment contributes to each mode. Each mode can therefore be interpreted as a distinct phytoplankton community. Additionally, the EOF amplitude function was used to plot a time series for the first 4 modes which shows the relative contribution of each pigment temporally and spatially. An EOF was also conducted on the HS mean-centered silicate, nitrate, and phosphate surface (0-2.5 m) nutrients from 2008-2017. In this case, each EOF mode can be interpreted as a distinct nutrient regime.

Results

Kennebec Watershed Bedrock and Land Use

The KR watershed is primarily composed of silicate-rich bedrock (Figure 6A). Sandstone, mudstone, quartzite, and granite comprise 56.5% of the total bedrock in the watershed (Table 1, Appendix). Sandstone, mudstone, quartzite, slate, and marble banding occurs in the southwest to northeast direction while granite, felsic gneiss, gabbro, and mafic metavolcanic rock types appear as isolated intrusions in the northwest and southern portions of the watershed.

Reflective of the entire KR watershed, all sub-watersheds display high percentages of silicate rich bedrock (Figure 7A, Table 1, Appendix). Overall, upstream stations K01, K03, K05, and K07 dominantly consist of sandstone, mudstone, and quartzite while downstream stations K10, K11, K12 and K13 have higher slate and felsic gneiss percent bedrock types. K01 is both the largest sub-watershed and contains the largest amount of bedrock types. Moving downstream, stations K03, K05, and K07 contain fewer bedrock types, mostly sandstone, mudstone, quartzite, mélange, and schist. In particular, K07 is almost entirely composed of sandstone. At stations K10, K11, K12, and K13, felsic gneiss and slate begin to appear. Moreover, there is a defined shift at K11 from mafic metavolcanic to slate, forming the Chops.

The bedrock informs the land use types observed within the watershed. The watershed is comprised of mostly forested land cover (65%) with pockets of harvested forest in the northwest and cultivated and developed land accumulating moving downstream toward K11 (Figure 6B, Table 1, Appendix). In addition to forested territory, swamps and open fresh water make up 13% of the watershed (Table 1, Appendix). Harvested forest, which makes up 4.6% of the KR watershed, is concentration in the NW region, in close proximity to the upstream stations K01-

K05 (Table 1, Appendix). Cultivated and developed land are concentrated near downstream stations K06-K11. Cultivated land is established dominantly on sandstone bedrock while developed areas are concentrated on granitic bedrock.

Similar to land use in the entire watershed, all sub-watersheds have the highest percent forested land coverage (Figure 7B, Table 1, Appendix). K01 contains the largest percent open fresh water. K03 and K05 have the highest percent forested land and also contain swamps. Cultivated and developed land coverage increase moving downstream, introduced significantly at K05 and maintained downstream. K10 contains the highest percentage of both cultivated and developed land. The K11, K12, K13 sub-watershed has the highest brackish water percentage with significant developed land and salt marsh/estuary presence.

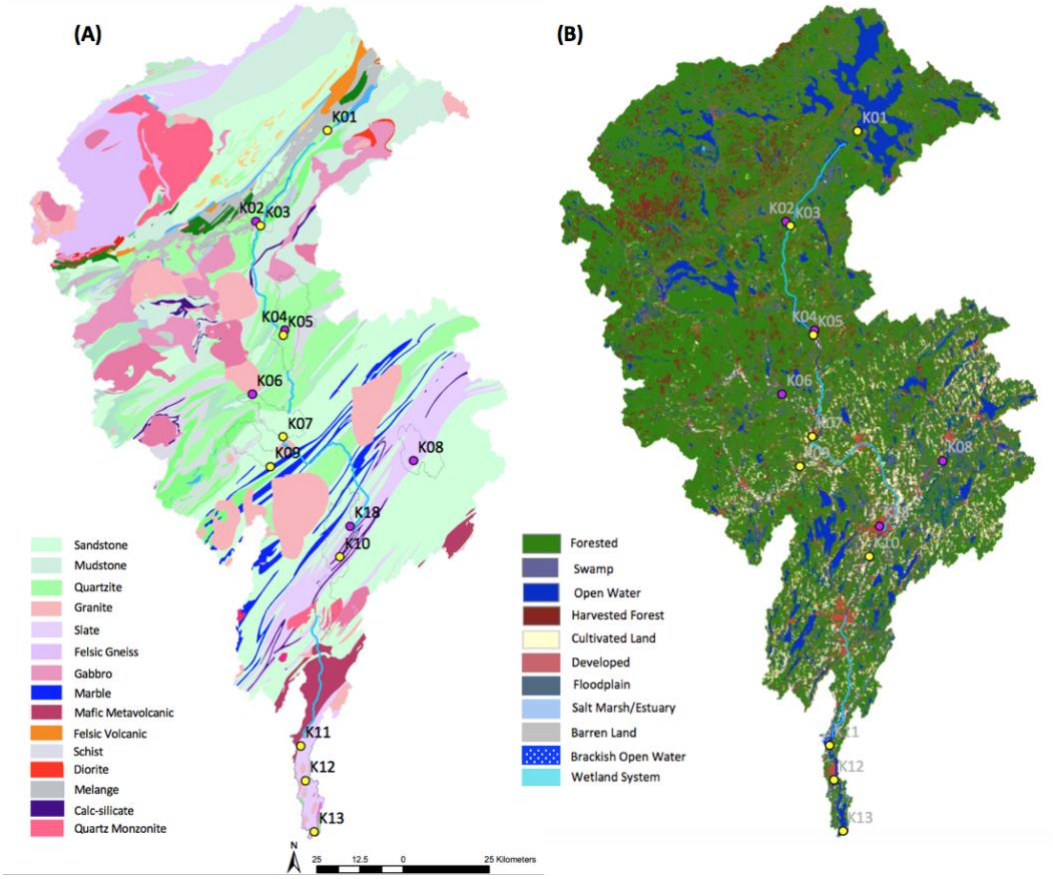


Figure 6. Map of the Kennebec River watershed (A) bedrock and (B) land use. The main stem of the Kennebec River is drawn in blue, main stem stations are labeled with a yellow dot, and tributary stations are labeled with a purple dot.

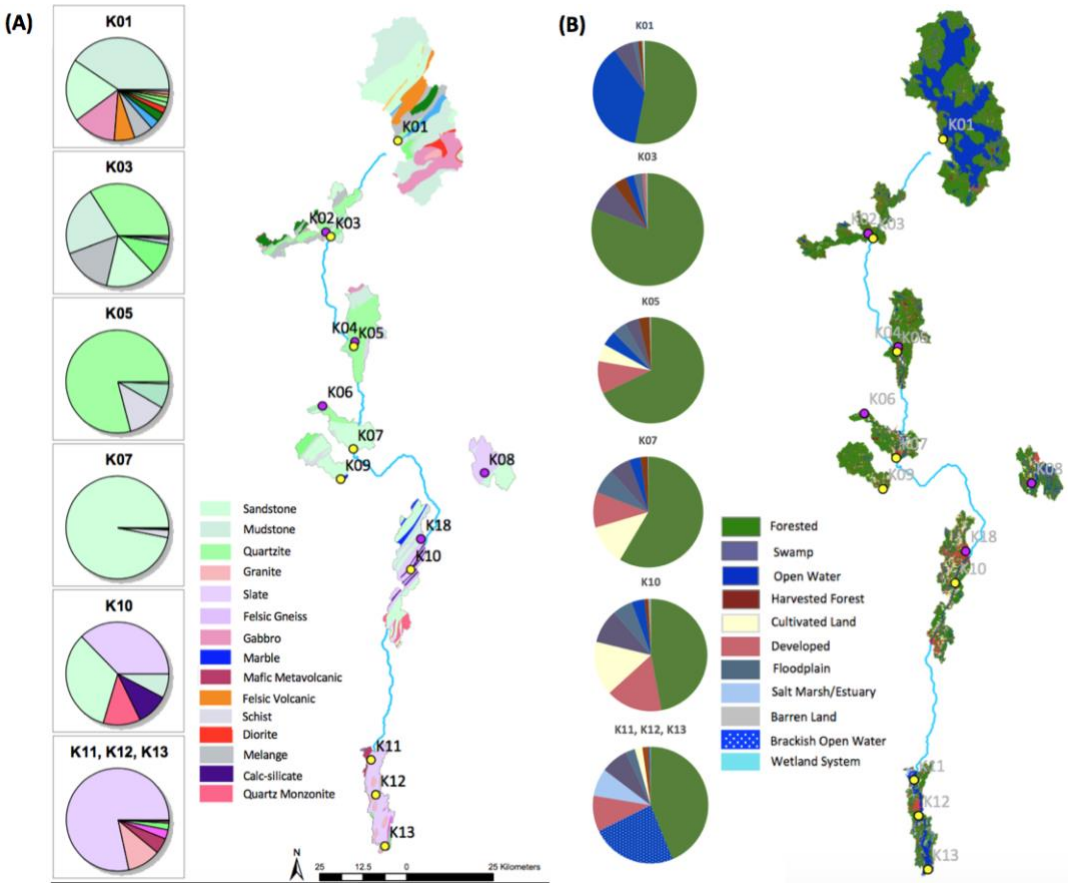


Figure 7. Map of Kennebec River sub-watershed (A) bedrock and (B) land use for each station. Pie charts display the percent (A) bedrock type and (B) land use type within each main stem sub-watershed. The main stem of the Kennebec River is drawn in blue and main stem stations are labeled with a yellow dot and tributary stations are labeled with a purple dot.

Nutrient Concentration and Composition

Spatial Trends

Each nutrient displays unique accumulation and/or depletion trends spatially from the headwaters of the KR to HS. Silicate concentrations are the greatest of all 4 measured nutrients in the KR and in HS. Silicate ranges from 0 μM to 40 μM in the KR and 0 μM to 15 μM in HS while phosphate, nitrate, and ammonium range from 0 μM to 0.1 μM, 0 μM to 2.5 μM, and 0 μM to 2 μM, respectively (Figure 8).

Silicate concentrations accumulate moving downstream from K01 to K10 and then decrease from K10 to 14 (HS) (Figure 8A). Highest silicate concentrations are observed at K07

and K10 with average concentrations of 35.1±11.6 μM and 41.3±14.5 μM, respectively (Table 2). After K10, silicate concentrations decrease from 41.3±14.5 μM to 24.4±12.6 μM at K13. Lowest silicate concentrations are found in HS with an average concentration of 8.1±4.2 μM and a maximum concentration of 17.5 μM.

Phosphate has the lowest concentration out of the 4 measured nutrients (Figure 8B). Upstream stations K01 through K10 range from 0.03±0.04 μM to 0.08±0.04 μM with highest concentrations in the upstream stations observed at K07 and K10 (Table 2). Phosphate is relatively higher at K01, then decreases slightly at K03 and increases again at K05. Beginning at K11, phosphate accumulates moving downstream and into station 14 (HS) where phosphate increases up to an average concentration of 2±2.18 μM. The highest maximum concentration of phosphate, 7.1 μM, is also observed at station 14.

Nitrate is depleted moving downstream from K01 to K07 with relatively high average concentrations at K01 (1.6±0.3 μM) and lower average concentrations downstream at K07 (1.2±0.4 μM) (Table 2, Figure 8C). K10 is a new source of nitrate with the highest observed average concentration of 1.6±0.6 μM and the highest maximum concentration, 3.5 μM. Nitrate is depleted again from K12 to station 14 with the lowest average (0.4±0.2 μM) and maximum (0.8 μM) nitrate concentrations at station 14.

Overall, ammonium accumulates from upstream to downstream excluding slight decreases in ammonium at K07 and K11 (Figure 8D, Table 2). The highest average concentration of ammonium is observed at K13, 3.5±1.7 μM. Ammonium decreases from K13 to station 14 where the average concentration is 1.7±2.1 μM. Overall lowest average concentrations of ammonium (0.8±0.4 μM) are found in the headwaters of the KR (K01).

Table 2. Mean ± standard deviation (SD) and maximum silicate, phosphate, nitrate, and ammonium concentrations for all mainstem KR stations and HS (station 14) from 2011-2013.

Station	Silicate (µM)		Phosphate (µM)		Nitrate (µM)		Ammonium (µM)	
	Mean±SD	Max	Mean±SD	Max	Mean±SD	Max	Mean±SD	Max
K01	23.4±8.7	35.3	0.03±0.04	0.2	1.6±0.3	1.9	0.8±0.4	2.2
K03	25.7±9.4	49.3	0.02±0.03	0.1	1.4±0.4	2.6	0.8±1	4.8
K05	31.5±15.8	81.4	0.04±0.06	0.3	1.2±0.5	3	1.6±2.3	12
K07	35.1±11.6	50.7	0.09±0.06	0.2	1.2±0.4	2.1	1.2±0.5	2.7
K10	41.3±14.5	68	0.08±0.04	0.2	1.6±0.6	3.5	1.8±1.3	7
K11	30.5±17.6	58.8	0.05±0.03	0.1	1.2±0.6	2	1.1±0.7	3
K12	30.9±19.3	66.3	0.07±0.05	0.3	1.3±0.8	2.7	2±1.2	5.6
K13	24.4±12.6	44.8	0.14±0.05	0.3	0.9±0.5	1.9	3.5±1.7	7.5
14 (HS)	8.1±4.2	17.5	2±2.18	7.1	0.4±0.2	0.8	1.7±2.1	5.4

Temporal Trends

In addition to the spatial distribution of nutrients from upstream to downstream, nutrients also vary seasonally and interannually. High silicate concentrations dominate during the summer months, with the exception of downstream stations K11-K13 in 2011 and 2012 (Figure 8A). 2013 has the highest silicate concentrations of downstream stations K10-K13. Silicate is most depleted during the spring and fall months for all three years. In HS, silicate is relatively depleted in all sampled months when compared to the silicate concentrations contained within the river.

In contrast to the silicate concentrations in HS, phosphate maintains the highest concentrations in HS from 2011 through 2013 (Figure 8B). K13 also maintains relatively high phosphate concentrations in all sampled months from 2011 through 2013. Phosphate concentrations are relatively lower in upstream stations K01, K03, and K05 during the late summer months, but relatively high in the late spring/early summer during all three years. Concentrations are higher in the summer at K07 for 2011 and 2012, but low in the summer and high in the spring and fall in 2013.

For each of the 3 sampling years, the highest nitrate concentrations were observed at K01 and lowest nitrate concentrations were observed in HS; however, significant seasonal variations are observed on top of this larger trend (Figure 8C). High nitrate concentrations are observed in the spring of 2012 and 2013 at stations K01-K12 and in the fall at K01 from 2011-2013. In addition to low nitrate concentrations in HS, markedly low concentrations of nitrate also appear at K13, excluding relatively higher nitrate concentrations in the fall of 2011 and 2013 and in the spring of 2013.

Ammonium differs more interannually from upstream to downstream than the other nutrients (Figure 8D). In all sampled months from 2011 through 2013, K12 and K13 have the highest ammonium concentrations. Ammonium concentrations appear to be the greatest overall in 2013, aside from station 14 which has the lowest ammonium values. Station 14 has highest concentrations throughout the year in 2011, but significantly lower concentrations in 2012 and 2013. From 2011 through 2013, K10 and K11 have anomalously high concentrations of ammonium in the spring and fall and lower concentrations in the summer. K05 and K07 display isolated high concentrations events in all three years with the addition of K01 and K03 during 2012 and 2013.

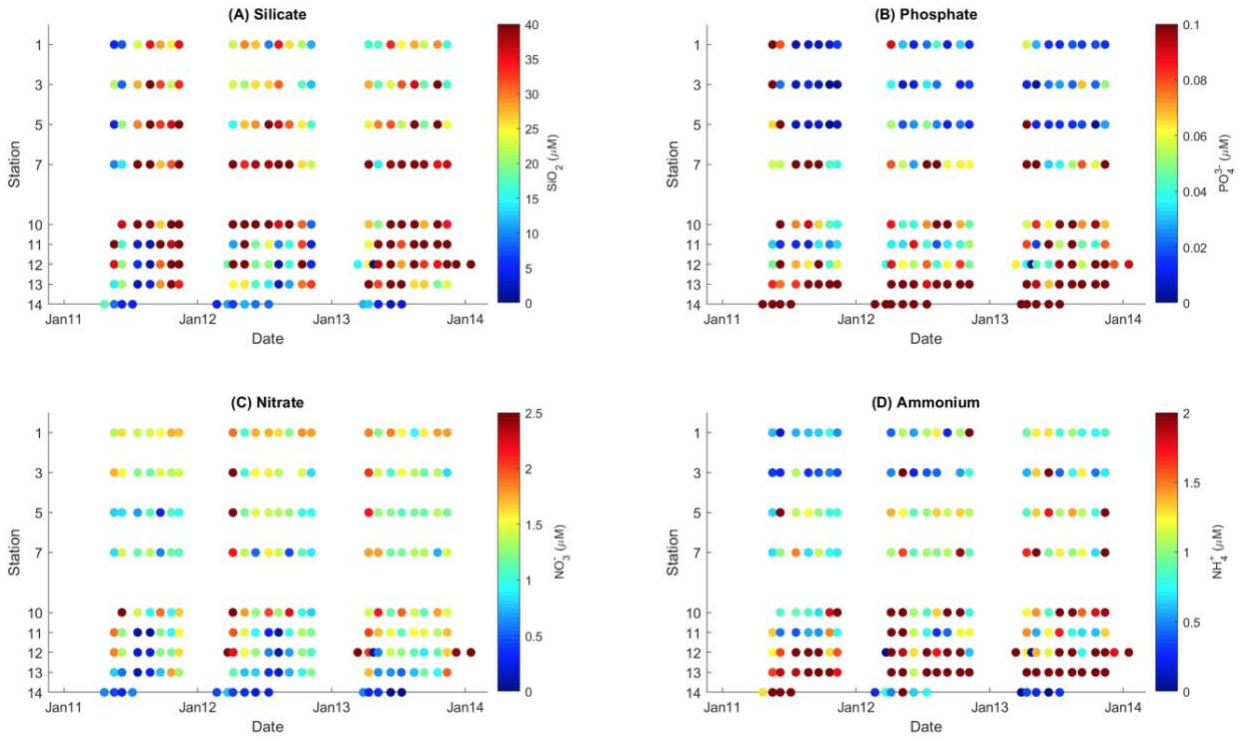


Figure 8. Time series from 2011-2013 of nutrients (A) silicate, (B) phosphate, (C) nitrate, and (D) ammonium from the Kennebec River (main stem stations 1-13) and Harpswell Sound (station 14). Each subplot is color coded by the nutrient concentration.

The nutrient ratios N:P and N:Si reveal similar trends in accumulation and utilization from upstream to downstream (Figure 9, Appendix, Table 2). From K01 to K07, the N:P ratio decreases from 115.6±94.4 to 17.7±9.0 which indicates a gain in phosphate and/or a decrease in nitrate moving downstream. The N:P ratio increases from K10 to K11 and then decreases again at station 14 where the N:P ratio is lowest, 0.3±0.2. The range in N:Si, 0.04±0.02 to 0.09±0.06, is significantly smaller than the range in N:P. The small range in N:Si demonstrates that all stations have a high silicate concentration when compared to nitrate. All changes in the N:Si ratio are small when compared to the changes observed in N:P.

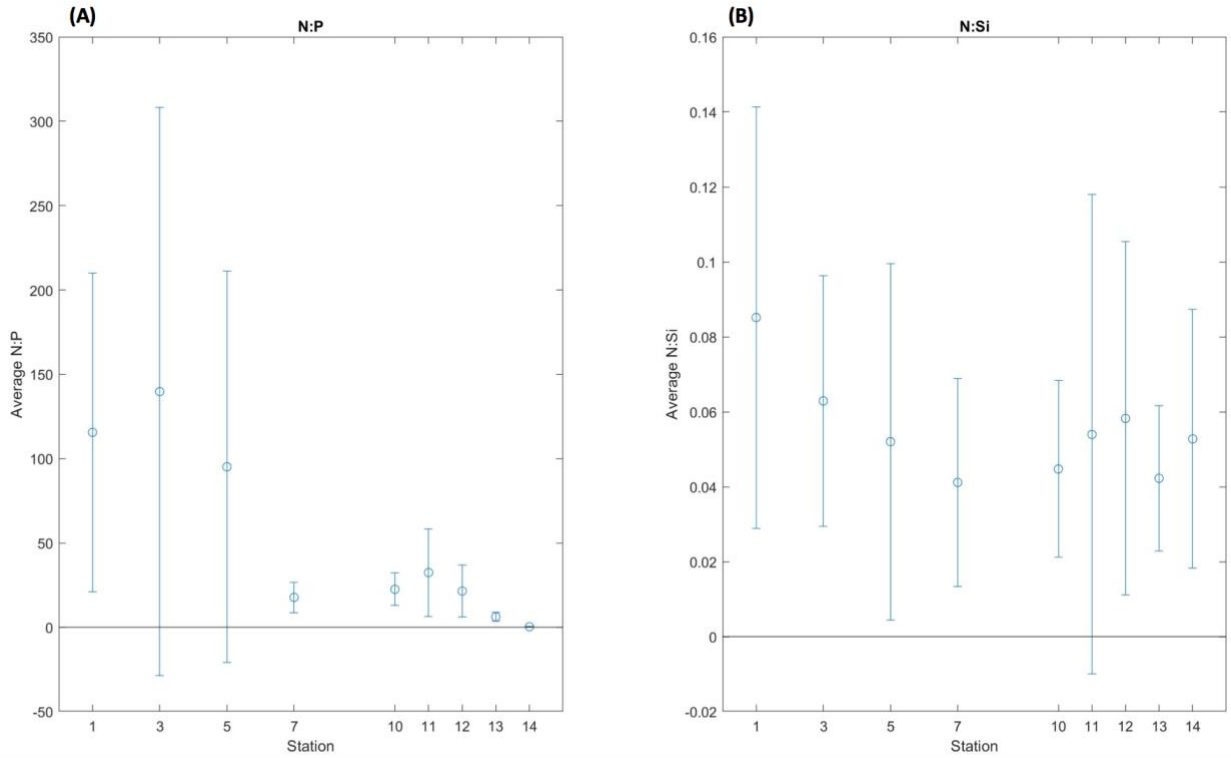


Figure 9. (A) Nitrate:phosphate (N:P) and (B) nitrate:silicate (N:Si) average molar concentrations plotted for each Kennebec River main stem station (1-13) and Harpswell Sound (station 14) for the 2011-2013 time period. Error bars represent the calculated standard deviation from first computing the nutrient ratio, then computing the standard deviation of taking the average of those nutrient ratios. A line of reference is plotted at 0.

Phytoplankton Pigment Concentration and Composition

Spatial and temporal trends in phytoplankton pigments correlate with nutrient patterns. Relatively low nutrient concentrations in the upstream stations when compared to downstream stations is reflected by the relatively low phytoplankton pigment concentrations in upstream versus downstream stations.

Spatial Trends

TChla (representing phytoplankton biomass) varies by an order of magnitude from an average concentration of 0.801±0.39 μg/L, measured at K05, to 8.544±7.745 μg/L, measured at K11, with the overall highest average and maximum concentrations recorded at downstream stations K11-14 (Figure 10A, Table 3). K11 has the highest observed maximum concentration in

TChla, 25.6 μg/L, with the second highest maximum and average concentrations recorded in HS. The overall lower TChla concentrations observed in upstream stations when compared to downstream stations are consistent with the overall lower nutrient concentrations in upstream stations.

TChlb (representing chlorophytes) reflects a similar upstream to downstream pattern as TChla with a shift in concentration from relatively low to high concentrations occurring between K10 and K11 (Figure 10B). Overall average concentration in TChlb ranges from 0.031±0.018 μg/L to 0.239±0.128 μg/L, with upstream stations K01-K10 ranging from 0.031±0.018 μg/L to 0.04±0.051 μg/L and downstream stations K11-14 ranging from 0.108±0.068 μg/L to 0.239±0.128 μg/L. The overall highest average and maximum concentrations of TChlb are observed at station 14, with TChlb increasing from K12 to 14.

Fuco (representing diatoms and red algal PFTs) also closely resembles the pattern of TChla with highest average concentrations observed at stations K11-14 (Figure 10C). Downstream stations range in average Fuco concentration from 0.611±0.391 μg/L to 2.597±2.806 μg/L while upstream stations range from 0.167±0.101 μg/L to 0.308±0.17 μg/L. Maximum (10.083 μg/L) and highest average concentrations occur at K11, the same station as the observed largest concentrations of TChla. Overall, each data point for TChla matches the relative magnitude of each point in the Fuco time series.

Perid (representing dinoflagellates) ranges from 0.005±0.006 μg/L, measured at K01, to 0.171±0.174 μg/L measured at station 14 (Figure 10D). Thus, smallest and largest concentrations are observed at the headwaters of the KR and in the coastal waters at HS, respectively. There is minimal Perid in stations K01-K12 with the largest average concentration, amongst these

stations, observed at K10. The highest maximum concentration of Perid is found at K13 with a value of 1.83 µg/L.

Table 3. Mean ± standard deviation and maximum Total Chlorophyll a, Total Chlorophyll b, Fucoxanthin, and Peridinin concentrations for all mainstem KR stations and HS (station 14) from 2011-2013.

Station	Total Chlorophyll a (µg/L)		Total Chlorophyll b (µg/L)		Fucoxanthin (µg/L)		Peridinin (µg/L)	
	Mean±SD	Max	Mean±SD	Max	Mean±SD	Max	Mean±SD	Max
K01	1.125±0.24	1.661	0.031±0.018	0.098	0.271±0.096	0.4815	0.005±0.006	0.016
K03	1.217±0.493	2.492	0.05±0.026	0.121	0.309±0.17	0.727	0.009±0.012	0.054
K05	0.801±0.39	1.806	0.029±0.016	0.069	0.167±0.101	0.458	0.009±0.016	0.074
K07	0.872±0.325	1.502	0.053±0.058	0.293	0.193±0.076	0.359	0.006±0.01	0.042
K10	1.337±0.508	3.01	0.084±0.051	0.236	0.257±0.089	0.428	0.019±0.012	0.048
K11	8.544±7.745	25.6	0.173±0.115	0.451	2.597±2.806	10.083	0.01±0.013	0.035
K12	2.458±1.502	7.663	0.108±0.068	0.288	0.62±0.412	2.164	0.014±0.015	0.056
K13	2.085±1.271	4.672	0.129±0.086	0.339	0.611±0.391	1.358	0.146±0.373	1.827
14 (HS)	3.091±1.262	4.949	0.239±0.128	4.949	0.846±0.472	1.8315	0.171±0.174	0.609

Temporal Trends

Temporally, the pigments appear to be overall less variable interannually than nutrients but display seasonal patterns that are co-related to the seasonal trends in nutrients. The temporal patterns in TChla and Fuco are the same, just with a smaller magnitude for Fuco concentrations (Figure 10A, C). Highest concentrations for both pigments are observed in almost all sampled months from 2011-2013 at K11. Anomalously high concentrations in each pigment are also observed at K03 in the fall and spring months for all three years. Overall, relatively high concentrations are observed in the fall and spring months in upstream stations K01-K10 while high concentrations are found in the summer in downstream stations K11-14. The high concentrations in Fuco and TChla seasonally corresponds with low concentrations of silicate in the fall and spring in upstream stations and summer in downstream stations (Figure 8A).

TChlb is highest during the summer months at all stations (Figure 10B). Downstream stations K11-14 maintain a strong presence of TChlb for the duration of the summer months in all 3 years while upstream stations K01-K07 have single high measurements in each year in the summer. K10 has anomalously high concentrations of TChlb during the mid-summer in 2011 and 2012, but lower concentrations in 2013. Overall, lower concentrations of TChlb are found in the spring and fall months in all stations except station 14.

Perid displays the lowest concentrations out of all 4 pigments in all 3 years from K01-K12 (Figure 10D). K13 has the highest concentrations of Perid in the late summer months for all 3 years and station 14 has high concentrations in the spring and summer in all 3 years. In May 2012, an anomalously high concentration of Perid was observed at K03, K05, and K07, reaching a peak concentration at K05.

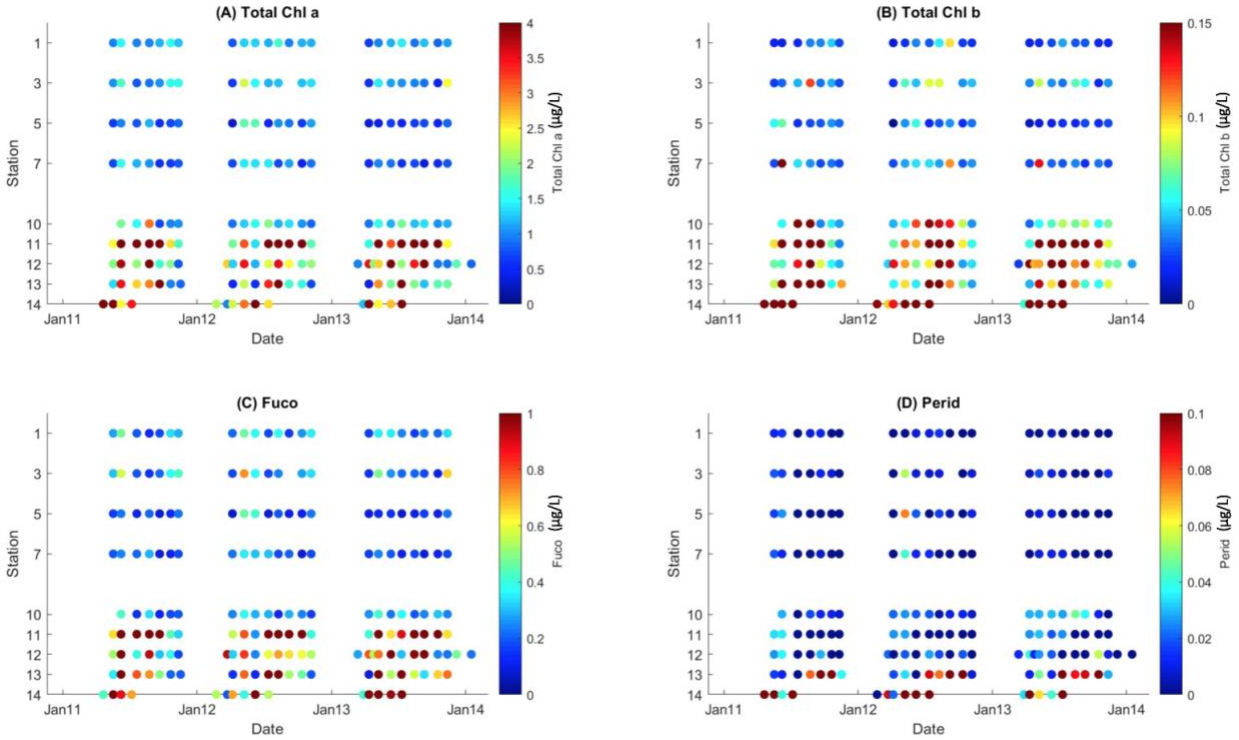


Figure 10. Time series from 2011-2013 of dominant phytoplankton pigments (A) Total Chlorophyll a, (B) Total Chlorophyll b, (C) Fucoxanthin, and (D) Peridinin from the Kennebec River (main stem stations 1-13) and Harspwell Sound (station 14). Each subplot is color coded by phytoplankton pigment concentration.

Kennebec River Phytoplankton Communities

Kennebec River Cluster Analysis

To begin to understand how phytoplankton communities are constructed and which pigments are co-located, a cluster analysis was conducted on all main stem stations from K01-K13 (Figure 11). The cluster analysis revealed two dominant phytoplankton pigment clusters: a green algae cluster and a red algae cluster. A cluster between Perid and HexFuco, and ButFuco and Allo form the green algae cluster and a cluster between Caro, TChlc, Fuco, Diad, and Diato and a slightly separate green algal cluster consisting of TChlb and Zea form the red algae cluster. Thus, TChlb and Zea are co-located with the red algae pigments while the other green algae pigments cluster separately.

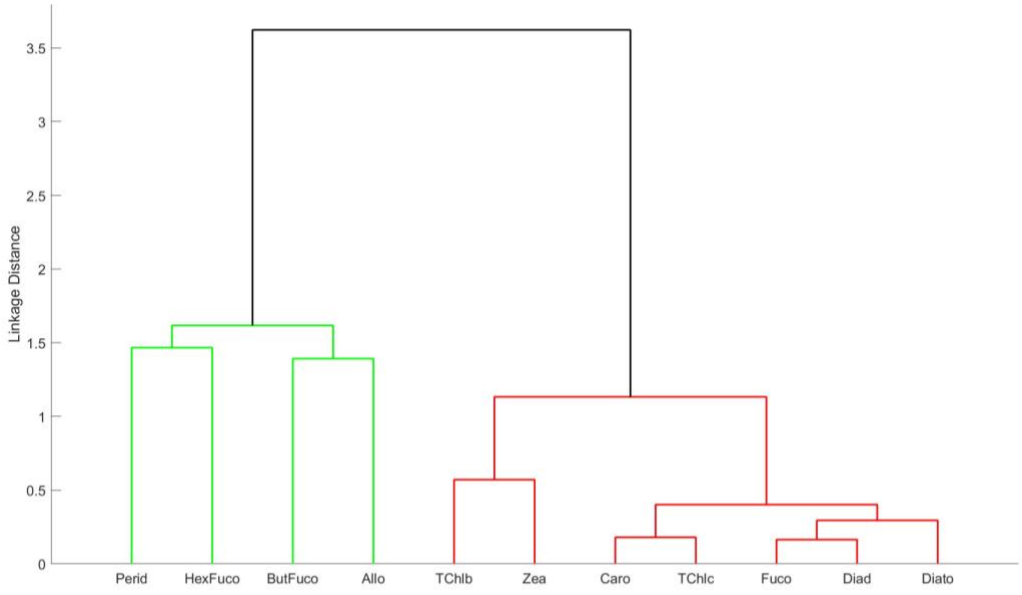


Figure 11. Hierarchical cluster analysis of KR mainstem stations K01-K13 pigments using a correlation distance of 1-R (R=Pearson's correlation coefficient between pigments) and Ward's linkage method. Each cluster is color coded based on the dominant taxonomic value.

Kennebec River Empirical Orthogonal Function Analysis

To understand the spatial and temporal distribution of the pigment clusters found in the KR, an EOF analysis was first conducted on all main stem stations from K01-K13. The EOF

modes were consistent with the groupings found in the cluster analysis. The first four EOF modes explained 99.5% of the variance in pigment data and represent four distinct phytoplankton communities in the Kennebec River: diatoms; dinoflagellates; chlorophytes, cryptophytes, and prymnesiophytes; and chlorophytes, euglena, and cryptophytes (Figure 12).

EOF Spatial Trends

There is a distinct spatial boundary for all 4 modes between stations K01-K10 and K11-K13 for modes 1 and 2 and K01-K07 and K10-K13 for modes 3 and 4. The diatom community, represented by EOF mode 1, is co-related to all other pigments and explains the greatest variance, 97.6% (Figure 12A). The phytoplankton within the river are therefore driven by diatom biomass and thus the other phytoplankton communities, while taxonomically distinct, are minor contributors to the overall variability. The EOF mode 1 amplitude ranges from -0.06 μg/L to 0.06 μg/L with highest amplitudes observed at K11. Downstream stations K11-K13 are out of phase with upstream stations K01-K10 where the EOF amplitude is almost always negative.

EOF mode 2 represents a dinoflagellate community co-related with chlorophytes and cryptophytes, explaining only 1.2% variance (Figure 12B). Diatoms are anti-correlated with these taxa. The EOF amplitude values range from -0.006 μg/L to 0.006μg/L. This community displays the overall highest EOF amplitude values at K13. Similar to mode 1, K01-K10 are out of phase with downstream stations K11-K13.

Mode 3 defines a chlorophyte, cryptophyte, and prymnesiophyte community, explaining just a small fraction of the variance (0.4%) (Figure 12C). Because there is a clear distinction in the community structure, this small fraction of the variance is biologically important. Diatoms and dinoflagellates are anti-correlated with this mode. EOF amplitudes range from -0.008 μg/L

to 0.008μg/L with the highest EOF amplitudes observed at K10-K13. In this mode, K01-K07 mostly display negative EOF amplitudes while K10-K13 have more positive EOF amplitudes.

Mode 4 only explains 0.3% of the overall variance and represents a chlorophyte, euglena, and cryptophyte community (Figure 12D). Diatoms are anti-correlated with this mode. Mode 4 EOF amplitudes have the smallest range compared to previous modes, ranging from -0.005 μg/L to 0.005 μg/L. Similar to mode 3, K01-K07 have overall lower, more negative EOF amplitudes than K10-K13 where the largest positive EOF amplitudes are observed.

EOF Temporal Trends

The diatom community (mode 1) displays very similar seasonal trends interannually (Figure 12A). Highest EOF amplitudes found in downstream stations K11-K13 dominate in the late spring and summer months. Lowest EOF amplitudes in the downstream stations occur in the early spring and fall particularly for K12 and K13 while relatively high amplitudes are observed in almost all sampled months for K11. Upstream stations K01-K10 display negative EOF amplitudes for almost all sampled months and all 3 years.

The dinoflagellate community (mode 2) varies interannually with highest EOF amplitudes in downstream stations K11-K13 in 2011 and lower amplitudes in 2012 and 2013 (Figure 12B). K13 has the highest EOF amplitude values for all 3 years which is consistent with the constant high Perid concentration (Figure 9). The overall highest amplitudes at K13 occur in the summer months with slightly lower amplitudes in the fall and spring. Upstream stations K01-K07 maintain negative amplitudes with less negative amplitudes in the summer months.

The chlorophyte, cryptophyte, and prymnesiophyte community (mode 3) demonstrates dominance in the summer months for all stations and across the 3 year time series (Figure 12C). K01-K07 have smaller magnitude positive amplitudes in the summer months, except for an

anomalously high amplitude observed in the early summer at K07 in 2011. K10-K13 have the highest EOF amplitudes in the summer for all 3 years with negative EOF amplitudes observed in the spring and fall.

The chlorophyte, euglena, and cryptophyte community (mode 4) also maintains high amplitude values in the summer months with lower amplitudes in the fall and spring at all stations (Figure 12D). Similar to the mode 3 community, highest amplitudes are found in K10-K13; however, some of the largest negative amplitudes are also observed during the summer in 2012 at K11 and K13. Negative amplitudes in the fall and the spring correspond with the anti-correlated diatom community.

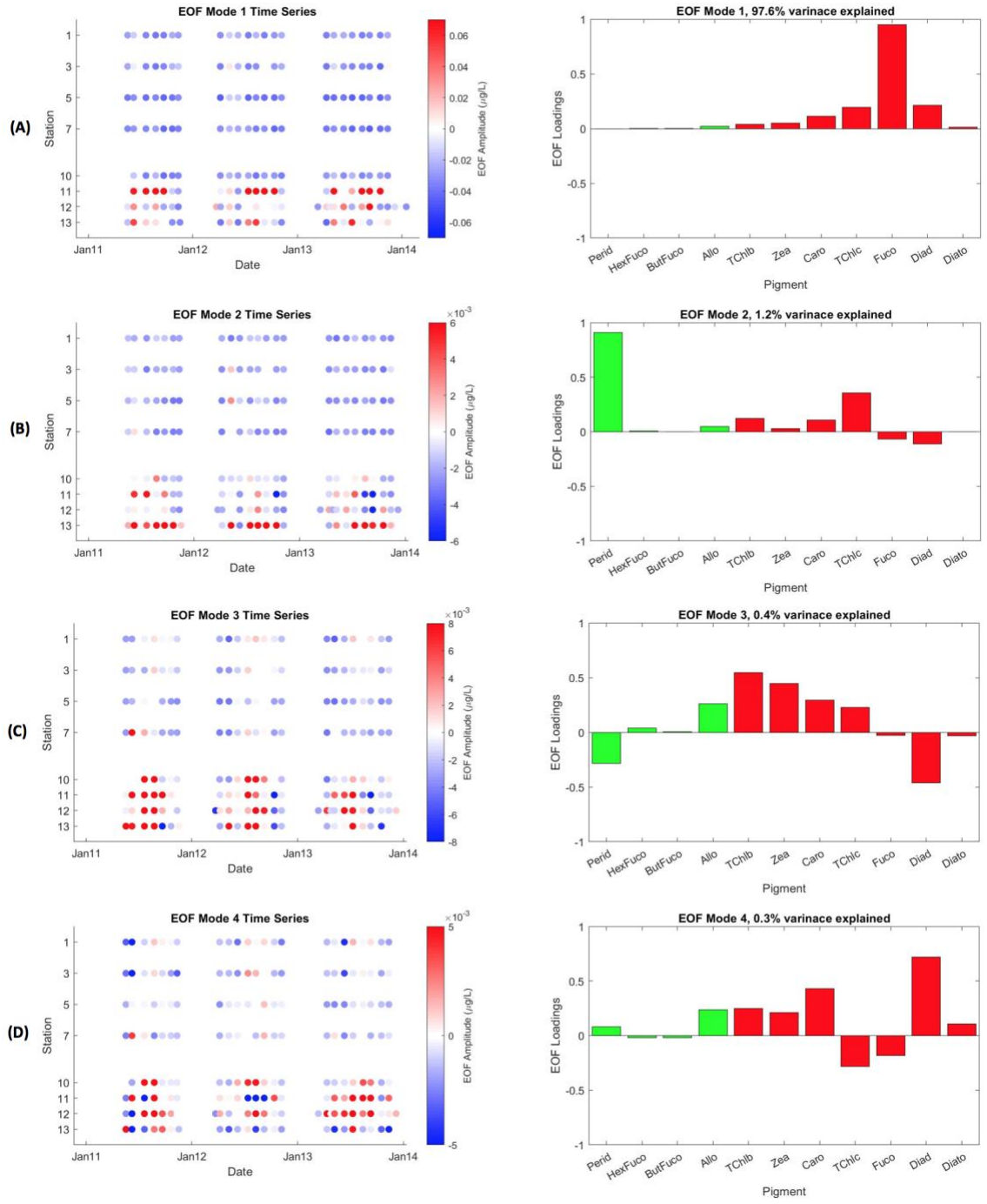


Figure 12. Pigment Empirical Orthogonal Function loadings of modes 1-4 (A-D) (right) and respective time series (left) for the Kennebec River main stem stations from 2011-2013. Bar plots are color coded to match the cluster analysis (Figure 11). The mode number and percent variance explained by each mode are displayed above each bar plot. Time series plots are color coded by EOF amplitude.

Overall, since upstream stations K01-K10 most often had significantly smaller EOF amplitude magnitudes and often were out of phase with downstream stations K11-K13, a

separate cluster and EOF analysis was conducted for downstream stations K11-K13 (Figure 1, 2, Appendix). Cluster and EOF results were almost identical to the cluster and EOF analysis conducted on the entire KR (Figures 11,12). This indicates that the downstream stations were driving the results for the entire river.

Upstream Kennebec River Phytoplankton Communities

To account for the clear differences between stations K01-K10 and K11-K13, a separate cluster and EOF analysis was conducted for upstream stations K01-K10.

Upstream Cluster Analysis

Upstream stations have two different clusters than those found in downstream stations (Figure 13). A cluster with TChlb, Zea, Caro, and Allo along with HexFuco and ButFuco construct a green algae cluster and Fuco, TChlc, and Diad along with Diato and Perid construct a red algae cluster. This reveals that in upstream stations the green algae cluster separately from the red algae, but Perid clusters along with the red algae.

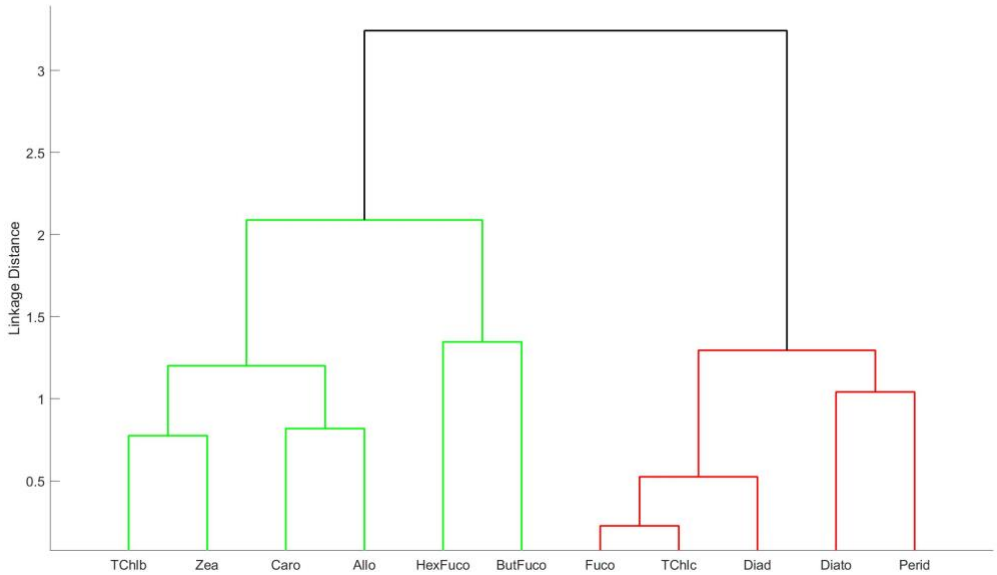


Figure 13. Hierarchical cluster analysis of upstream KR mainstem stations K01-K10 pigments using a correlation distance of 1-R (R=Pearson's correlation coefficient between pigments) and Ward's linkage method. Each cluster is color coded based on the dominant taxonomic value.

Upstream Empirical Orthogonal Function Analysis

Expanding on the results of the cluster analysis, the EOF analysis conducted on upstream stations reveals an upstream phytoplankton community structure that is distinct from the downstream community. The first four EOF modes explain 97.8% of the variance in pigment data in upstream KR stations (K01-K10) and represent four different distinct phytoplankton communities: diatoms, chlorophytes and cryptophytes, chlorophytes, and cryptophytes and euglena.

EOF Spatial Trends

The diatom community, represented by mode 1, explains 76.9% variance (Figure 14A). All pigments are co-variable in this mode. EOF amplitude ranges from -0.02 μg/L to 0.2 μg/L with the overall highest amplitudes observed in upstream stations K01 and K03 while lowest amplitudes are observed at K05.

Mode 2, representing the chlorophyte and cryptophyte community, explains 12.7% variance (Figure 14B). This community is anticorrelated with diatoms, but co-related to dinoflagellates. Although dinoflagellates co-vary with this community, the loading value is significantly smaller than the dinoflagellate loading value observed in EOF mode 2 for all KR main stem stations (Figure 12B). EOF amplitude ranges from -0.008 μg/L to 0.008μg/L with highest amplitudes observed at K10.

Mode 3 represents a chlorophyte community which explains 6.1% variance (Figure 14C). This community is anti-correlated with cryptophytes. EOF amplitudes range from -0.008 μg/L to 0.008 μg/L. All stations contain negative and positive EOF amplitudes with highest amplitudes at K07 and lowest at K05. The lowest amplitudes at K05 indicate the dominance of cryptophytes, the anti-correlated community.

Mode 4, explaining 2.1% variance, represents a cryptophyte and euglena community (Figure 14D). EOF amplitudes range from -0.005 μg/L to 0.005μg/L. Chlorophytes and dinoflagellates co-vary with mode 4 while diatoms and cyanobacteria are anti-correlated. Positive amplitude values are maintained at K05. K01 has consistently negative EOF amplitude values, indicating the dominance of cyanobacteria and/or diatoms.

EOF Temporal Trends

Diatoms (mode 1) exhibit a strong seasonal pattern with high positive amplitude values occurring in the spring and fall and negative amplitudes in the summer (Figure 14A). K01 and K03 exhibit the largest positive amplitudes in the spring and fall with negative amplitudes during the summer for all 3 years. K05 has the most negative amplitudes but exhibits large positive amplitudes in the spring of 2012. All stations below K05 have relatively lower magnitude positive amplitudes in the spring and summer and more negative amplitudes overall when compared to upstream stations K01 and K03.

The chlorophyte and cryptophyte community (mode 2) exhibits high positive amplitudes in the mid-summer months for all 3 years and in all upstream stations (Figure 14B). The overall highest EOF amplitudes are observed during the summer months at K10 and during the summer of 2011 at K07. Anomalously low concentrations of the chlorophyte and cryptophyte community occur when the diatom community (mode 1) exhibits positive EOF amplitudes in the fall and spring (Figure 14A, B).

The chlorophyte community (mode 3) positive amplitudes have discrete peaks that vary monthly and annually (Figure 14C). The highest positive amplitude values are observed in June 2011 and May 2013 at K07; however, the most consistent positive amplitude values are found at

K03 and K10. The anti-correlated cryptophyte community, represented by negative amplitudes is consistently present at K05, particularly in 2012 and at K01 in 2011.

The cryptophyte and euglena community is consistently present at K05 in all 3 years with constant positive amplitude values (Figure 14D). At K01, K03, and K07, distinct spring and fall positive amplitudes are observed in all 3 years with the overall largest amplitude occurring in the spring, 2011 at K07. Negative amplitudes, representing cyanobacteria, are consistently present during the summer and fall at K01 for all 3 years. Distinct negative amplitudes can also be found during the summer in all 3 years at K07 and K10. This could relate to cyanobacteria or the diatom community which is also anti-correlated with this mode and displays positive loadings in mode 1 during the summer at these stations.

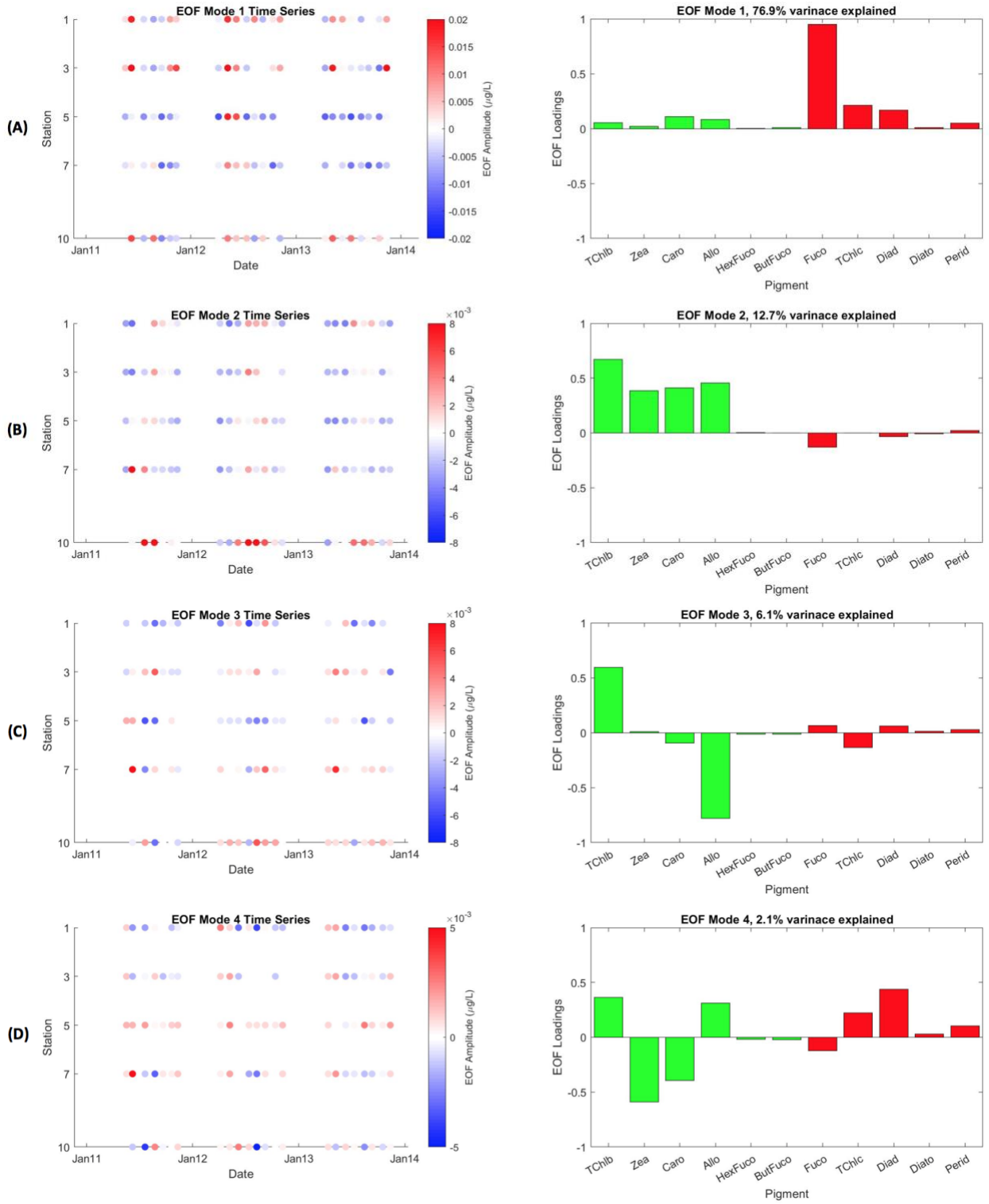


Figure 14. Pigment Empirical Orthogonal Function loadings of modes 1-4 (A-D) (right) and respective time series (left) for the Kennebec River main stem stations 1, 3, 5, 7, and 10 from 2011-2013. Bar plots are color coded to match the cluster analysis for upstream KR stations (Figure 13). The mode number and percent variance explained by each mode are displayed above each bar plot. Time series plots are color coded by EOF amplitude.

Harpswell Sound Nutrient Regime

In understanding the KR phytoplankton community structure and nutrient regime, we transition into the coastal waters of HS, beginning with an initial assessment of HS nutrients through an EOF analysis of the years overlapping with the KR data, 2011-2013.

The nutrient EOF defines 3 distinct nutrient regimes (Figure 15). Mode 1, which explains 88.3% of the variance, is represented dominantly by phosphate with a loading value of 1 (Figure 15A). Silicate and nitrate slightly co-vary with this mode. Mode 2 explains 11.5% variance and represents a silicate regime with co-varied nitrate and anti-correlated phosphate (Figure 15B). In mode 3, which explains only 0.3% variance, nitrate dominates with correlated phosphate, only minimally contributing to the variance, and anti-correlated silicate (Figure 15C). Although some of the nutrients are co-variable within each mode, each nutrient regime displays a unique temporal trend.

EOF Temporal Trends

Mode 1, the phosphate nutrient regime, has the largest EOF amplitude range, -10 μM to 11 μM, and has minimal interannual variability (Figure 15A). The largest positive EOF amplitudes for each year are observed in late-April/early-May with the overall largest positive amplitude observed in the spring of 2011. More negative amplitudes are observed in the late summer. A mid-summer positive amplitude, however, is observed in all 3 years. There is similar interannual variability in the phosphate regime with all 3 years sharing a similar curve shape with only a slight lag time between peak positive and negative amplitudes for each year.

The silicate nutrient regime (mode 2) ranges in amplitude from -3 μM to 7 μM and demonstrates more interannual variability when compared to the phosphate regime (Figure 15B) The largest positive amplitude value is observed in the early spring (late-March) of 2012.

Positive amplitude peaks are also observed in the spring (late-April) and summer (mid-May and early-June) of 2011 and in the summer of 2013 (late-May). Negative amplitudes are also found during the summer months for all three years.

Mode 3, representing the nitrate nutrient regime, has the smallest amplitude range from -0.4 μM to 0.7 μM and has the greatest interannual variability (Figure 15C). The largest positive amplitude value is observed in the late summer (early-July) of 2011. Largest positive amplitude values for each year are observed during the summer months. Each year has a different pattern in positive and negative amplitudes, but in all years the transition from positive to negative appears to occur relatively quickly when compared to the other nutrient regimes.

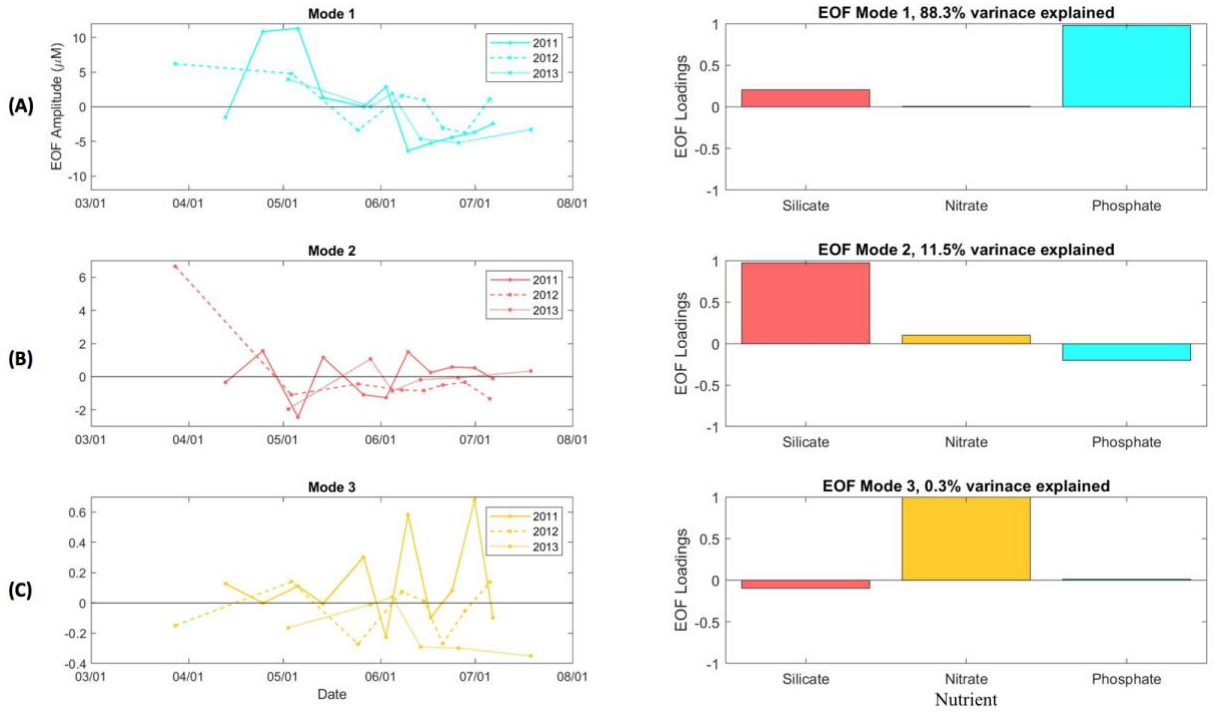


Figure 15. Nutrient Empirical Orthogonal Function loadings (right) and corresponding EOF amplitude time series (left) of modes 1-3 (A-C) for Harpswell Sound 2.5m depth measurements from 2011-2013. The mode number and percent variance explained by each mode are displayed above each bar plot. Time series are color coded by the largest loading value observed in each respective mode (mode 1: phosphate, mode 2: silicate, mode 3: nitrate). Data points for each year are connected with a line to discern trends. Line type indicates year: 2011 is a solid line, 2012 is a dashed line, and 2013 is a dotted line. A black line is plotted at 0 to help distinguish between positive and negative amplitudes.

Harpswell Sound Phytoplankton Communities

Harpswell Sound Cluster Analysis

Similar to the KR, a cluster analysis was first conducted to understand which pigments co-locate and form communities. The hierarchical cluster analysis revealed 3 pigment clusters that dominate the co-variability of the HS pigment data: green algae, pelagophytes, haptophytes, and cyanobacteria; dinoflagellates and cryptophytes; and red algae and diatoms. Fuco, Diato, Diad, TChlc, and Caro form a diatom cluster, Perid and Allo form a cryptophyte and dinoflagellate cluster, and ButFuco, HexFuco, Zea, and TChlb form a green algae cluster (Figure 16). Clusters also represent two size classes of phytoplankton: (1) nano-plankton (green algae) and pico-plankton (cyanobacteria) and (2) micro-plankton (diatoms).

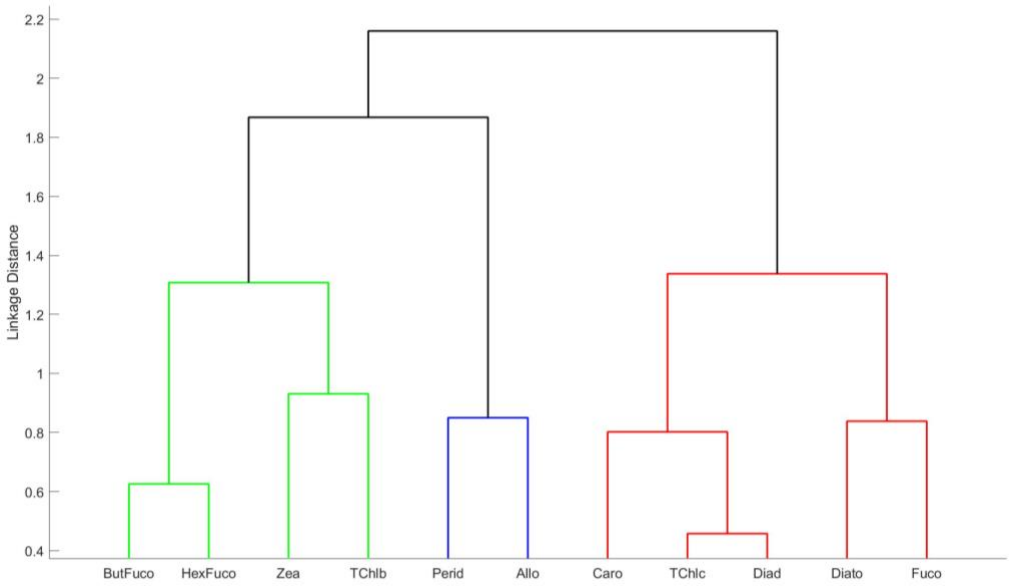


Figure 16. Hierarchical cluster analysis of Harpswell Sound pigments using a correlation distance of 1-R (R=Pearson's correlation coefficient between pigments) and Ward's linkage method. Each cluster is color coded based on the dominant taxonomic value.

Harpswell Sound Empirical Orthogonal Function Analysis 2011-2013

The same major taxonomic groups generated using the cluster analysis were found in the EOF analysis. The first four EOF modes explain 98.2% of the variance in pigment data and

represent four distinct phytoplankton communities in HS: diatoms; dinoflagellates, chlorophytes, and cryptophytes; chlorophytes and cryptophytes; and cryptophytes (Figure 17). Pigments representing a diatom community contribute most to EOF mode 1 which explains 82.8% of the overall variance (Figure 17A). All of the other pigments are co-related, but diatoms dominate the mode. Mode 2, which explains 11.0% variance, is represented by covariance between dinoflagellates, chlorophytes, and cryptophytes (Figure 17B). Prymnesiophytes minimally contribute to mode 2 as well. Diatoms display negative loading values and are therefore anti-correlated with these taxa. In mode 3, dinoflagellates become strongly anti-correlated while chlorophytes and cryptophytes co-vary with diatoms which only minimally contribute to the variance of this mode (2.7%) (Figure 17C). Mode 4 only explains 1.7% variance and is represented by a cryptophyte dominated community (Figure 17D). In this mode, chlorophytes and dinoflagellates are anti-correlated while diatoms co-vary.

EOF Temporal Trends

The diatom community (mode 1) has the overall largest EOF amplitude values compared to the other phytoplankton communities, ranging from -0.5 μg/L to 1 μg/L (Figure 17A). The overall largest EOF amplitude was observed in late-May of 2011. Positive amplitudes for 2012 and 2013 were also recorded during the summer months. Negative amplitudes were found during the spring and late summer of 2011 and 2012 and in the mid-summer of 2013.

The dinoflagellate, chlorophyte, and cryptophyte community (mode 2) have amplitudes ranging from -0.2 μg/L to 0.4 μg/L with largest positive amplitudes anti-correlating with mode 1 large positive amplitudes (Figure 17B). In 2011, the maximum peak positive amplitude occurs in early-May, preceding the mid-summer peak in the diatom community in 2011. Another positive peak amplitude event occurs in July for all 3 years while the diatom community has relatively

lower positive or negative amplitudes. The largest negative amplitude in the summer of 2012 occurs during the same time of a peak positive amplitude value for the diatom community in 2012. Since diatoms are anti-correlated with mode 2, the temporal anti-correlation between the two communities is expected.

Mode 3, representing the chlorophyte and cryptophyte community ranges from -0.1 μg/L to 0.1μg/L (Figure 17C). Largest positive amplitudes are observed in June and early-July for all 3 years. The largest positive peak for 2011 and 2012 occur in early-July when the diatom community maintains negative amplitudes. The peak positive amplitude for 2013 occurs in mid-June, preceding the late-June peak in the diatom community. In 2013, the mode 2 and mode 3 communities display clear anti-correlation.

The mode 4 cryptophyte community ranges from -0.1 μg/L to 0.1μg/L and appears to precede blooms in the chlorophyte and cryptophyte mode 3 community (Figure 17D). In 2012 and 2013, largest positive amplitude peaks occur in late-June and late-May, respectively, just preceding the peak positive amplitudes in the mode 3 community. Similar to all previous communities, the overall largest positive peak amplitude is observed in 2011. This peak occurs at a similar time as the mode 2 2011 peak positive amplitude event. Relatively lower positive or negative amplitudes occur when the mode 3 community has relatively higher positive values, representing the anti-correlation between these two communities.

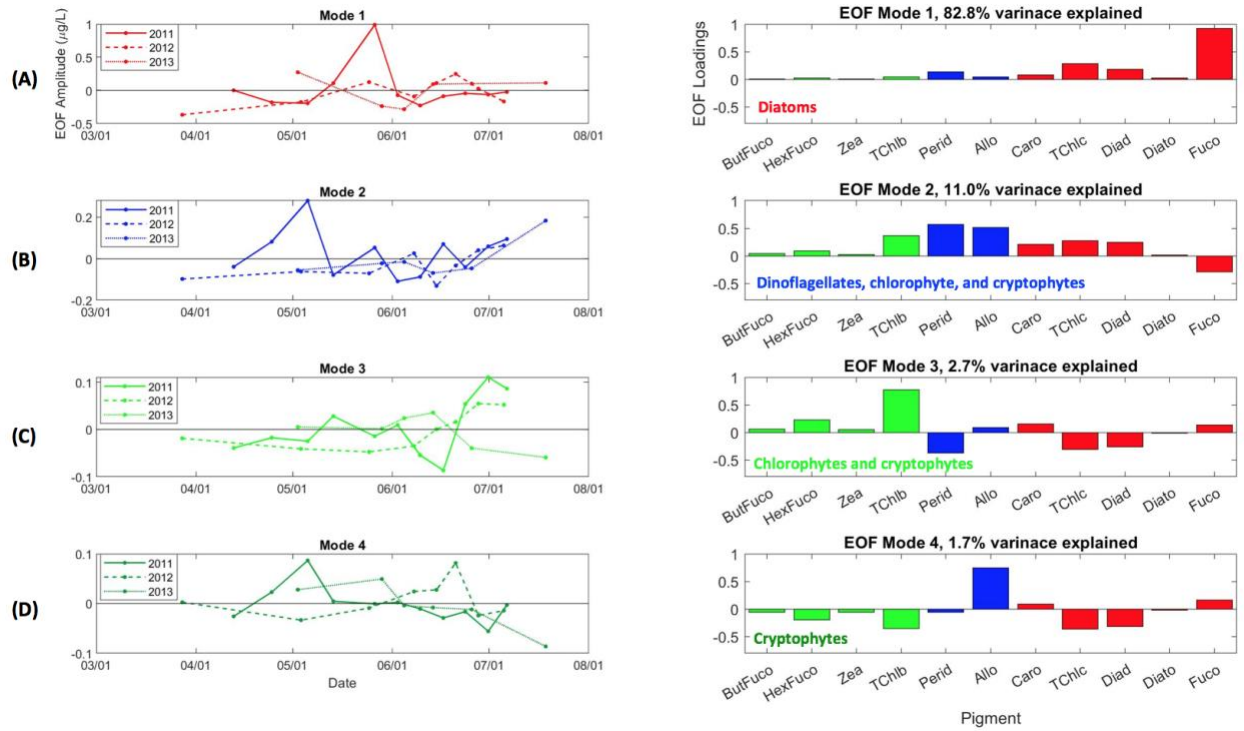


Figure 17. Pigment Empirical Orthogonal Function loadings (right) and corresponding EOF amplitude time series (left) of modes 1-4 (A-D) for Harpswell Sound 2.5m depth measurements from 2011-2013. Bar plots are color coded to match the Harpswell Sound cluster analysis (Figure 16). The mode number and percent variance explained by each mode are displayed above each bar plot. The assumed taxonomic values of each mode are described within each subplot. Text color matches the EOF amplitude time series color which is based on the dominant loading value determined by each bar plot. Data points for each year are connected with a line to discern trends. Line type indicates year: 2011 is a solid line, 2012 is a dashed line, and 2013 is a dotted line. A black line is plotted at 0 to help distinguish between positive and negative amplitudes.

Harpswell Sound Long Term Trends

Since the 2011-2013 data is only a subset of the nutrient and phytoplankton pigment data for HS, EOF analyses for the HS nutrient and pigment data from 2008-2017 were conducted to assess the long-term trends in the HS nutrient regime and phytoplankton community structure.

Harpswell Sound Nutrient Empirical Orthogonal Function Analysis 2008-2017

Over the long-term HS dataset from 2008-2017, silicate dominates the variance within the estuary (Figure 18). Mode 1 represents a silicate nutrient regime with co-variable nitrate and

phosphate, explaining 79.9% of the variance (Figure 18A). Mode 2 explains 18.8% variance and represents a nitrate regime with co-varied phosphate and anti-correlated silicate (Figure 18B). Mode 3, which explains only 1.3% variance, represents a phosphate regime with anti-correlated silicate and nitrate (Figure 18C).

From 2008 through 2017, silicate displays a significant shift in concentration after 2014, nitrate displays the greatest interannual variability, and phosphate remains at relatively lower amplitude magnitudes (Figure 18D). Mode 1, representing a silicate regime, has the largest range in EOF amplitude, -4.4 μg/L to 27.5 μg/L. The overall highest positive amplitude is observed in 2013, but relatively high amplitudes are observed from 2009 through 2013. In 2014, however, a shift in the silicate nutrient regime occurs with much lower EOF amplitudes observed from 2014 through 2017. Mode 2, representing a nitrate regime, ranges in EOF amplitude from -2.7 μg/L to 20.0 μg/L. Highest EOF amplitudes are observed in 2011 and 2016, but overall, the nitrate regime is quite variable interannually. Mode 3 (representing a phosphate regime) has the smallest range and maximum EOF amplitude value, ranging from -1.0 μg/L to 8.5 μg/L. Similar to mode 2, the highest positive amplitudes are observed in 2011 and 2016. Throughout the time series, however, phosphate has relatively low amplitudes when compared to silicate and nitrate.

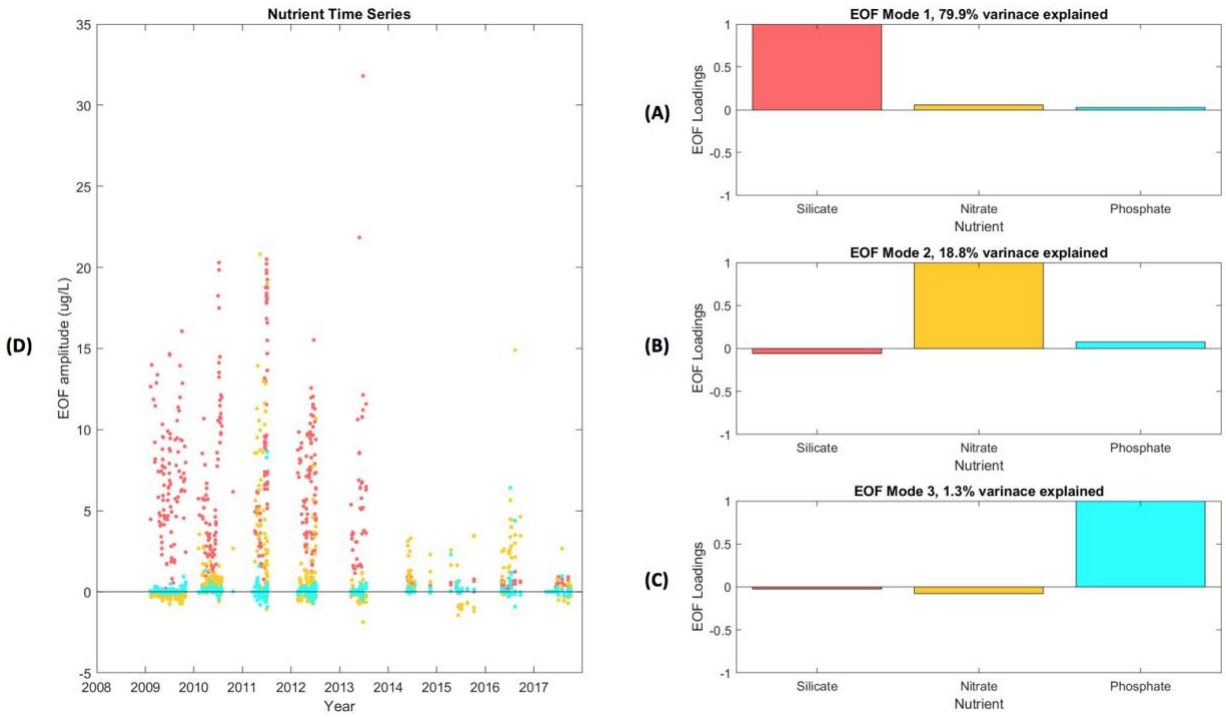


Figure 18. Mean-centered nutrient Empirical Orthogonal Function loadings of modes 1-3 (A-C) and the corresponding EOF amplitude time series (D) for Harpswell Sound from 2008-2017. EOF loading bars are color coded by nutrient. The mode number and percent variance are indicated above each bar plot. The EOF amplitude time series for each mode are plotted together with points color coded by the nutrient with the largest loading value (mode 1 is pink, mode 2 is yellow, and mode 3 is cyan). A black line is plotted at 0 to help distinguish between positive and negative amplitudes.

Harpswell Sound Pigment Empirical Orthogonal Function Analysis 2008-2017

The pigment EOF modes for the long-term analysis represent the same four distinct phytoplankton communities generated by the EOF conducted on the subset of data from 2011-2013: diatoms; dinoflagellates, green algae, and cryptophytes; green algae and cryptophytes; and cryptophytes (Figure 19). The first four EOF modes explain 94.2% of the variance in pigment data with mode 1, representing the diatom community, explaining 69.8% variance.

All phytoplankton communities identified by the EOF modes demonstrate an increase in the diatom community and decrease in the dinoflagellate, chlorophyte, and cryptophyte communities. Mode 1 (representing a diatom dominated community) has the highest positive amplitude, ranging from -0.3 µg/L to 0.9 µg/L (Figure 19A). The overall largest positive

amplitudes were observed in 2014 and 2017. Mode 2 (representing a dinoflagellate, chlorophyte, and cryptophyte community) ranges in amplitude from -0.2 μg/L to 0.6 μg/L with largest positive amplitudes observed in 2009 and 2011 and consistently lower positive amplitudes from 2014 through 2017 (Figure 19B). The chlorophyte and cryptophyte community, representing mode 3, displays a similar interannual trend with lowest positive amplitudes observed from 2014 through 2017 and largest positive amplitudes observed in 2009 and 2011 (Figure 19C). The mode 4 (representing a cryptophyte community) time series displays a similar amplitude trend as mode 1, but only ranges from -0.2 μg/L to 0.6 μg/L (Figure 19D).

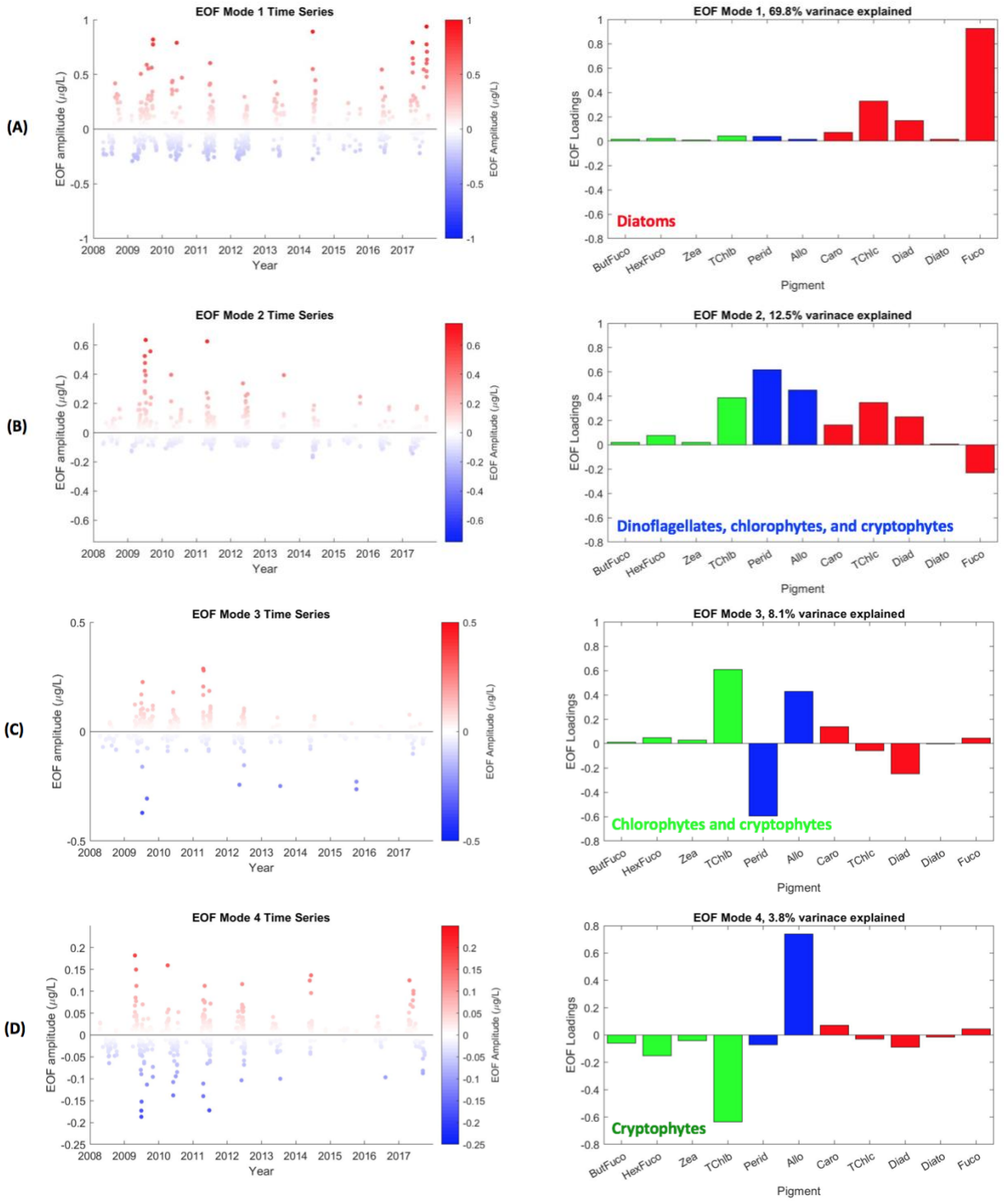


Figure 19. Pigment Empirical Orthogonal Function loadings (right) and corresponding EOF amplitude time series (left) of modes 1-4 (A-D) for Harpswell Sound 2.5m depth measurements from 2008-2017. Bar plots are color coded to match the Harpswell Sound cluster analysis (Figure 16). The mode number and percent variance explained by each mode is displayed above each bar plot. The assumed taxonomic value of each mode is described in the bottom left corner of each bar plot. The EOF amplitude time series is color coded by EOF amplitude. A black line is plotted at 0 to help distinguish between positive and negative amplitudes.

Discussion

Distinct River and Ocean Nutrient Regimes and Phytoplankton Communities

Differences in raw nutrient and pigment concentrations along with differing community structures between the upstream stations and downstream stations including HS indicate that the KR upstream freshwater stations should be considered as a separate ecosystem from the tidally impacted downstream stations and HS. Thus, we reject our first hypothesis and confirm hypothesis 2: there is no connectivity (excluding silicate) between the KR and HS so the coastal waters can be considered a separate, ocean influenced ecosystem from the river ecosystem. The tidal influence at and downstream of K11 provides a flux of phosphate, nitrate, and ammonium, while silicate is primarily river-derived. Silicate concentrations accumulate moving downstream with highest concentrations observed in upstream stations K05 and K07 along with downstream stations. The presence of high silicate concentrations in both upstream and downstream stations indicate that silicate is a river-derived nutrient source to the coastal waters. In contrast, phosphate exhibits significantly higher concentrations in tidally-impacted stations when compared to the depleted upstream stations, indicating that phosphate is an ocean-derived nutrient source to the coastal phytoplankton communities. Because nitrate is depleted from upstream to downstream in the KR, nitrate at the mouth of the KR is also an ocean-derived nutrient. Ammonium exhibits lowest concentrations in the headwaters and highest concentrations at K13 which is reflective of the overall increase in primary productivity in the downstream stations (Clark et al., 2008).

The flux of ocean-sourced nutrients sustains phytoplankton biomass that is 10 times greater in the tidally impacted stations and in HS when compared to upstream freshwater stations. This trend is not only observed in TChla (representing biomass), but also TChlb (green algae), Fuco (diatoms), and Perid (dinoflagellates) phytoplankton pigments. This indicates that

the biomass of phytoplankton taxa represented by these pigments increases significantly in downstream stations and in HS. In addition to the change in raw phytoplankton biomass, the composition of the phytoplankton communities in each ecosystem is distinct. In upstream stations, the four dominant communities consist of diatoms, chlorophytes, cryptophytes, and euglena while downstream stations and HS have the addition of a dominant dinoflagellate community. In addition to structural differences, the upstream diatoms behave differently in response to silicate concentrations when compared to the tidally impacted diatom communities.

What drives these fundamental differences in nutrient regime and phytoplankton community structure between the river and ocean ecosystems?

Upstream, Freshwater Nutrients and Phytoplankton Communities

Upstream phytoplankton communities display a seasonal turn-over in phytoplankton community related to bedrock, agriculture and urban land use types, environmental limitations, and the presence or absence of competitive taxa. Temporally, diatoms bloom in the spring, chlorophytes and cryptophytes bloom in the summer, followed by just cryptophytes and euglena in the late-summer, and diatoms again in the fall.

Seasonality in the upstream chlorophyte and cryptophyte phytoplankton communities and patterns in nutrient concentrations are governed by proximity to cultivated and developed land. The phytoplankton communities are spatially and temporally distributed in response to competition and access to necessary nutrients. Overall, the source of nitrate and phosphate transitions from upstream to downstream as an increase in cultivated and developed land starts establishing and influencing the nutrient regime beginning at K05. This is not only reflected by raw nitrate and phosphate concentrations, but also by the decrease in the N:P ratio moving downstream. These results corroborate the mechanisms described in Hunt et al., 2005 who found

that an increase in point-source inputs moving downstream would increase nitrate and phosphate concentrations, promoting biological activity. Increased biological activity due to point source nutrient input is also reflected by the isolated high concentrations of ammonium at K05 and K07 (Clark et al., 2008). Unlike the stations downstream of K05, upstream stations K01 and K03 rely heavily on the nitrate and phosphate supply that accumulates in Moosehead Lake, the headwaters of the KR, and swamp input from the K01 and K03 sub-watersheds (Atkinson et al., 2019).

The chlorophyte and cryptophyte community (upstream EOF mode 2) dominates in the summer months at K07 and K10, sub-watersheds that contain a relatively high percentage of cultivated and developed land. Temporally, blooming in the summer, and spatially, blooming at K07 and K10, allow for minimal competition with diatoms. Comparing the upstream mode 1 and mode 2 time series reveals that diatoms display a larger presence at K01 and K03 and precede the chlorophyte and cryptophyte blooms observed in the mid-summer months. In addition to the lack of competition in the summer, K07 and K10 provide ample amounts of nitrate and phosphate from cultivated land fertilizer runoff and from the accumulation of these nutrients as the KR flows downstream. Along with the highest concentration of cultivated land at K07, there is also a sewage treatment plant along the river that has been found to discharge many unregulated contaminants into the KR. The chlorophyte and cryptophyte community is then able to maximize on this new flux of nutrients. The chlorophyte community (upstream EOF mode 3) also dominates during the summer months at K07 for these similar reasons. The anti-correlated cryptophyte community however, dominates almost year-round at K05, just upstream of the chlorophyte community. Lack of competition from the chlorophyte community in addition to a presence of cultivated and developed land at K05 provides nutrient runoff and allows cryptophytes to dominate here.

The cryptophyte and euglena community (upstream EOF mode 4) also dominates during the summer months at all upstream stations, particularly at K05. The anti-correlated cryptophyte community in upstream EOF mode 3 and the upstream EOF mode 4 cryptophyte and euglena community are able to coexists at K05 with limited competition from the diatoms upstream and the chlorophytes downstream.

The anti-correlated cyanobacteria community appears throughout most of the year for all 3 years at K01 and during the summer at K07 and K10. Cyanobacteria often dominate in warmer water temperatures and nitrate/phosphate enriched environments (Havens et al., 2019). K01 is Moosehead Lake, the headwaters of the KR. The cyanobacteria community dominates at K01 in the summer months when the water is warmer and stagnant (Zhang et al., 2016). Moosehead Lake also sustains some of the highest concentrations of nitrate year-round amongst the upstream stations, providing the cyanobacteria with ample nutrient supply to maintain a dominant community almost year-round. Cyanobacteria also dominate at K07 and K10 from the surplus of nutrients delivered by cultivated land, and in the case of K07, unregulated nutrients from the sewage treatment plant. K07 has the lowest N:P ratio of the upstream freshwater stations which reflects the high phosphate inputs from the treatment plant. Pulses of these nutrients in the summer and fall allow cyanobacteria blooms to flourish.

With high biological activity of the chlorophyte and cryptophyte community during the summer months, the seasonal trend in nitrate is consistent with Hunt et al., 2005 findings of lower nitrate concentrations in the summer and higher nitrate in the spring and fall. The trend in phosphate also agrees with Hunt et al., 2005. Phosphate limits productivity upstream, but movement downstream and an increase in point-source phosphate inputs promotes chlorophyte and cryptophyte blooms.

Contrary to the varying upstream to downstream source of nitrate and phosphate to the KR that is utilized by the chlorophytes and cryptophytes, the source of silicate remains constant throughout the watershed; silicate is derived from the bedrock and surficial geology of the KR watershed and accumulates from upstream to downstream. All KR stations contain a plethora of silicate (Turner et al., 2003), demonstrated by the low N:Si ratios for all stations, due to silicate rich bedrock exposure, easily weatherable high surface area glacial sediment deposits, and the presence of dams that promote further silicate weathering and accumulation (Sullivan et al., 2000).

Although there is a sufficient amount of silicate in the river water column (up to 40 μM), the diatom community (mode 1 for the upstream EOF) does not dominate at all stations; rather, they appear in anomalously high concentrations at K01, K03, and K10 during the spring, fall, and occasionally during the mid-summer. In the spring and fall, KR discharge tends to be high with snowmelt and rain. Discharge can also be high at select times during the summer due to recreational dam releases. Lower silicate during the spring and fall and during select times in the summer is attributed to dilution and also due to utilization by diatoms. Based on Sullivan et al., 2000 findings, we expect dam water to be relatively enriched in silicate, with a longer water residence time, indicating that lower silicate concentrations in the mid-summer is most likely attributed to diatom utilization. In previous studies, diatoms have been found to dominate during periods of high discharge since they are able to withstand high turbulence waters while other taxonomic groups cannot (Reynolds, 1994; Bortolini and Bueno, 2013). It is surprising, however that much of the silicate during other seasons is unutilized by the diatoms and that silicate is, in fact, increasing moving downstream due to additive tributary sources.

Silicate is highest at K07 and K10 since these stations have accumulated unutilized silicate that originated upstream and are also positioned below dams, but diatom presence is still limited. Although the K07 sub-watershed almost entirely consists of sandstone, which is more difficult to weather based on Goldrich's Weathering Series (Bierman and Montgomery, 2014) and results from Lasaga et al., 1994, the Madison dam just upstream of K07 most likely accumulates high silicate concentrations with a longer water residence time behind the impoundment (Sullivan et al., 2000). This is reflected by the 40μM of silicate in the water column in almost all sampled months. Diatom presence here, however, is limited. Much of this unused silicate is probably transported to K10 where high silicate concentrations are also observed in all sampling months. K10 also has a high sub-watershed percent sandstone, but also is positioned below the Waterville dam, promoting silicate accumulation (Sullivan et al., 2000).

Perhaps there is a light limitation at K05-K10 that prevents the diatoms from maximizing on the silicate nutrient load (Chetelat and Pick, 2006). Although diatoms have been found to withstand lower light levels, the high turbidity of the KR in downstream stations may limit light too much for diatom survival. Further research on the specific species of diatoms dominating in the river is necessary to understand why their presence is limited with ample silicate.

Overall, the phytoplankton communities found in the upstream, freshwater stations corroborate previous studies on river phytoplankton. The KR upstream phytoplankton community was dominated by green and red algal groups which have faster growth rates and are therefore able to better withstand constant river flow, turbulence, and lower light (Reynolds, 1994; Bortolini and Bueno, 2013). Diatoms were found to dominate the upstream biomass, explaining 76.9% of the variance in upstream, freshwater stations which is consistent with the Descy et al., 2016 study of temperate river phytoplankton communities. Additionally, lower light

levels in the fall and spring due to higher discharge and more resuspended sediments allowed diatoms to dominate while less light limitation during the lower discharge summer months allowed for green algal group dominance (Chetelat and Pick, 2006; Descy et al., 2016). Dinoflagellates were not a dominant phytoplankton community which may indicate that the dinoflagellate in the river have a larger cell size, slower growth rate, and therefore, an inability to withstand the unstable river water column (Litchman et al., 2015).

Downstream, Tidal Nutrients and Phytoplankton Communities

Upon the entrance into the coastal waters, a shift in nutrient regime and phytoplankton community occurs, therefore rejecting H1, the hypothesis that there is complete connectivity between the KR and coastal waters. The lower KR gains nitrate and phosphate from the ocean and silicate from the river which allows for an increase in phytoplankton biomass. Additionally, the river widens and river flow rate decreases. Here, dinoflagellates are introduced into the community structure (mode 2 of the entire KR EOF).

The phytoplankton community within the downstream KR stations, K11, K12, and K13, is more spatially variable based on competition between taxa for specific nutrients. Diatoms, representing the mode 1 community for the EOF on downstream stations, dominate during the mid-summer months at K11 while dinoflagellates dominate just downstream at K13. With abundant silicate from the river and a plethora of ocean nutrients transported tidally to K11, diatoms dominate. As silicate becomes depleted by the diatoms, the dinoflagellates also maximize on the available ocean nitrate and phosphate but dominate spatially separately from the diatoms. Since dinoflagellates do not require silicate for growth, they sustain high concentrations where there are sufficiently available nutrients and minimal competition at K13. Chlorophytes and cryptophytes vary temporally with the diatoms and dinoflagellates. They

utilize any residual nitrate and phosphate during the summer months that was previously left over by the diatoms and/or the dinoflagellates (Margalef, 1958).

The increase in primary production in the downstream, tidal stations is reflected by the accumulation and increase in ammonium. The furthest downstream KR station, K13, has the highest ammonium concentration which is an indicator of phytoplankton nitrate uptake (Clark et al., 2008).

Harpswell Sound Nutrients and Phytoplankton Communities

Phytoplankton communities are uniform across the freshwater to saltwater interface between the lower KR tidally impacted stations and HS. KR stations K11-K13 and HS share the same 3 dominant phytoplankton communities: diatoms, dinoflagellates, and chlorophytes and cryptophytes, although the timing of each of these communities differs from the Chops to HS. While in the downstream stations, phytoplankton communities are more spatially variable, in HS, the phytoplankton communities vary seasonally.

In the analysis of HS from 2008-2017, the nutrient regime and phytoplankton community structure were found to vary systematically: diatoms bloomed first, stripping the waters of silicate, and dinoflagellates, chlorophytes, and cryptophytes bloomed subsequently, utilizing available nitrate and phosphate. In isolating a limited amount of surface data from 2011-2013, the nutrient regime and turn-over of the phytoplankton communities was contrary to this expected systematicity. While often diatoms are found to dominate in the spring and fall, the EOF time series displayed a diatom bloom in the mid-summer that was preceded by a dinoflagellate, chlorophyte, and cryptophyte bloom. In HS, however, there are often large diatom blooms that occur in February and therefore, this data captures and places emphasis on the secondary bloom (personal communication with Collin Roesler;

https://mclanelabs.com/imaging-flowcytobot/). Thus, data was not able to be collected early enough in the spring to observe the spring diatom bloom that would then be followed by the dinoflagellate, chlorophyte, and cryptophyte bloom, as expected in phytoplankton species succession (Margalef, 1958).

Additionally, while the nutrient EOF for 2008-2017 displayed the dominance of a silicate nutrient regime, followed by nitrate, then phosphate which correlated with the nutrient requirements for diatoms, followed by dinoflagellates, chlorophytes, and cryptophytes, the 2011-2013 time series presented phosphate as the nutrient regime explaining the highest percent variance, followed by silicate, then nitrate. This is most likely due to the dinoflagellate community preceding the diatom bloom within this set of data. A compilation of more surface water data points could help resolve this correlation between nutrients and phytoplankton community for the overlapping KR and HS years.

2011 was clearly driving the patterns shown in the nutrient and phytoplankton EOF analysis for the 2011-2013 time period. In the EOF nutrient analysis from 2008-2017, nitrate and phosphate EOF amplitudes displayed anomalously high positive amplitudes and silicate also displayed higher positive amplitudes. There were also larger EOF amplitude magnitudes observed in 2011 for the pigment EOF analysis from 2008-2017. Thus, when isolating the years 2011-2013, 2011 dominates the EOF results in nutrient regime and phytoplankton community. Higher phytoplankton biomass is also reflected by the anomalously high concentrations of ammonium observed in HS in 2011. Highest EOF amplitude magnitudes were also observed in upstream phytoplankton communities, indicating that 2011 was overall an anomalously high year for phytoplankton biomass and nutrients.

The nutrient EOF for 2011-2013, however, still displayed the non-correlation between nutrients, confirming that either the sources or utilization of the nutrients is sequential and not covariant. In other words, each nutrient EOF time series for 2011-2013 reflected the variability in nutrient source or utilization. Considering source, phosphate displayed minimal interannual variability which we might expect because not only is it an ocean-sourced nutrient to the estuary, but it also has the longest ocean residence time when compared to other biologically limiting nutrients (Martiny et al., 2019). In contrast, river-derived silicate is more variable interannually due to the variation in KR discharge seasonally and interannually. Nitrate displays the greatest interannual variability, as it is primarily sourced by deep ocean water replenishment through mixing across the pycnocline (Rebuck and Townsend, 2014). In considering utilization, since the dinoflagellate, chlorophyte, and cryptophyte bloom precedes the diatom bloom, we would expect nitrate and phosphate to be utilized first, quickly followed by silicate uptake by diatoms which is reflected in the nutrient EOF for 2011-2013.

Although the HS surface water data was relatively limited from 2011-2013, the same 4 dominant phytoplankton communities observed in the 10-year time series were still obtained within this 3-year time series and allowed us to confirm the consistency in phytoplankton community structure between downstream stations K11, K12, and K13 and HS. This consistency reveals that both downstream, tidally-impacted river stations and estuarine ecosystem phytoplankton communities are driven by ocean variability while upstream, freshwater phytoplankton communities are governed by land use, bedrock, and physical river properties.

Change Overtime in Harpswell Sound Phytoplankton Community and Nutrient Regime

Over the past decade, the phytoplankton community structure and nutrient regime has shifted in HS. With a dramatic increase in diatom concentrations and a slight decrease in

dinoflagellates, chlorophytes, and cryptophytes, in 2014, the estuary has observed a shift from a more diverse phytoplankton community to a diatom-dominated structure. With this increase in diatoms, *Pseudo-nitzchia* spp., an amnesic shellfish poisoning diatom species, is showing up in significant populations at the Bowdoin College Schiller Coastal Studies Center (HABON-NE, 2019-2023). The increased concentration in diatoms forced a decline in silicate concentration, with values remaining well-below the mean from 2014-2017. These findings have negative implications for marine upper trophic levels and the Maine fishing industry (Sellner et al., 2003).

Conclusion

Overall, this study sought to understand the influence of the Kennebec River (KR) on Harpswell Sound (HS) through a spatial and temporal approach. Results revealed the minimal connectivity between the KR upstream, freshwater ecosystem and the KR downstream, tidally impacted and HS estuarine ecosystems. Seasonality and spatial distribution of upstream phytoplankton communities are driven by river-variability based on bedrock, surrounding agriculture and urban land use types, environmental limitations, and the presence or absence of competitive taxa while seasonality in downstream and HS phytoplankton communities is driven by ocean variability and KR silicate input. These differences are reflected by the minimal influence of river nutrients on coastal waters (excluding silicate), the increase in phytoplankton biomass by a factor of 10 in the KR tidal downstream stations and in HS, the absence of a dominant dinoflagellate community in the freshwater upstream stations, and the difference in seasonality between each ecosystem.

These findings are not to say that the river does not have an influence on the estuarine ecosystems, as silicate from the KR does reach and impact the HS phytoplankton community; rather, this study reveals the necessity to strongly consider changes taking place in the ocean that

may have a greater impact on coastal Maine estuaries and other tidally impacted environments. It is therefore important to continue to monitor river systems as increased anthropogenic nutrient input and increased high discharge events with climate change could increase river impact on coastal waters; however, future studies must strongly consider variation in the ocean nutrient reservoir along with other nutrient sources to coastal waters including groundwater.

References

Anderson, D. M., Burkholder, J. M., Cochlan, W. P., Glibert, P. M., Gobler, C. J., Heil, C. A., Vargo, G. A. (2008). Harmful algal blooms and eutrophication: Examining linkages from selected coastal regions of the United States. *Harmful algae, 8*(1), 39-53.

Barraquand F., Picoche C., Maurer Daniele, Carassou L., Auby Isabelle (2018). Coastal phytoplankton community dynamics and coexistence driven by intragroup density-dependence, light and hydrodynamics. Oikos, 127(12), 1834-1852.

Barton, A. D., Finkel, Z. V., Ward, B. A., Johns, D. G., & Follows, M. J. (2013). On the roles of cell size and trophic strategy in North Atlantic diatom and dinoflagellate communities. *Limnology and Oceanography, 58*(1), 254-266.

Bierman, P. R., Montgomery, D. R., University of Vermont., & University of Washington. (2014). *Key concepts in geomorphology*.

Bluth, G. J. S., & Kump, L. R. (1994). Lithologic and climatologic controls of river chemistry. *Geochimica et Cosmochimica Acta, 58*(10), 2341-2359.

Bortolini, J. C., & Bueno, N. C. (2013). Seasonal variation of the phytoplankton community structure in the Sao Joao river, Iguacu National Park, Brazil/Variacao sazonal da estrutura da comunidade fitoplanctonica no Rio Sao Joao, Parque Nacional do Iguacu, Brasil. *Brazilian Journal of Biology, 73*, 1+.

Catlett, D., & Siegel, D. A. (2018). Phytoplankton Pigment Communities Can be Modeled Using Unique Relationships With Spectral Absorption Signatures in a Dynamic Coastal Environment. *Journal of Geophysical Research: Oceans, 123*(1), 246-264.

Chen, N., Liu, L., Li, Y., Qiao, D., Li., Y., Zhang, Y., Lv, Y. (2015). Morphology-based classification of functional groups for potamoplankton. *Journal of Limnology,74(3).*

Chételat, J., Pick, F. R., & Hamilton, P. B. (2006). Potamoplankton size structure and taxonomic composition: Influence of river size and nutrient concentrations. *Limnology and Oceanography, 51*(1part2), 681-689.

Clark D.R., Rees A.P., Joint I. (2008) Ammonium regeneration and nitrification rates in the oligotrophic Atlantic Ocean: Implications for new production estimates. *Limnology and Oceangography*, 53: 52–62.

Descy, J.-P., Darchambeau, F., Lambert, T., Stoyneva-Gaertner, M. P., Bouillon, S., & Borges, A. V. (2017). Phytoplankton dynamics in the Congo River. *Freshwater Biology, 62*(1), 87-101.

Devercelli, M., & O'Farrell, I. (2013). Factors affecting the structure and maintenance of phytoplankton functional groups in a nutrient rich lowland river. Limnologica, 43(2), 67-78.

D Tilman, S S Kilham, a., & Kilham, P. (1982). Phytoplankton Community Ecology: The Role of Limiting Nutrients. *Annual Review of Ecology and Systematics, 13*(1), 349-372.

Fennel, K., & Testa, J. M. (2019). Biogeochemical Controls on Coastal Hypoxia. *Annual Review of Marine Science, 11*(1), 105-130.

Fogg, G.E. (1986). "Estuarine and coastal pollution: Detection, research and control: Volume 18, numbers 4/5, 1986. Water Science and Technology. Proceedings of the IAWPRC/NERC Specialised Conference held at Plymouth 16–19 July 1985. Editors: D. S. Moulder & P. Williamson. 364 pp. Pergamon Press, Oxford. ISBN 0-08-033669-8". *Marine pollution bulletin*(0025-326X), 17 (11), p. 521.

HABON-NE. (2019-2023). An Adaptive Observing Network for Real-Time, In Situ HAB Monitoring and Data Sharing Across New England. Retrieved from https://coastalscience.noaa.gov/project/habon-ne-an-adaptive-observing-network-for-real-time-in-situ-hab-monitoring-and-data-sharing-across-new-england/.

Hankinson, S.J. (2010). Oceanographic and hydrologic effects on harmful algal blooms in Harpswell Sound, Gulf of Maine. *Bowdoin College*.

Hartzell, J. L., & Jordan, T. E. (2012). Shifts in the relative availability of phosphorus and nitrogen along estuarine salinity gradients. *Biogeochemistry, 107*(1), 489-500.

Havens, Karl E., Gaohua Ji, John R. Beaver, Fulton,Rolland S., I.,II, and Catherine E. Teacher. (2019). Dynamics of cyanobacteria blooms are linked to the hydrology of shallow florida lakes and provide insight into possible impacts of climate change. *Hydrobiologia* 829, no. 1: 43.

Hayhoe, K., Wake, C.P., Huntington, T.G. *et al.* (2007). Past and future changes in climate and hydrological indicators in the US Northeast. *Clim Dyn* **28,** 381–407.

Hunt, C. W., Loder, T., & VÖrÖsmarty, C. (2005). Spatial and Temporal Patterns of Inorganic Nutrient Concentrations in the Androscoggin and Kennebec Rivers, Maine. *Water, Air, and Soil Pollution, 163*(1), 303-323.

Ileva, N. Y., Shibata, H., Satoh, F., Sasa, K., & Ueda, H. (2009). Relationship between the riverine nitrate–nitrogen concentration and the land use in the Teshio River watershed, North Japan. *Sustainability Science, 4*(2), 189.

IOCCG (International Ocean Colour Coordinating Group) (2006). Phytoplankton Functional Types. Retrieved from https://ioccg.org/group/pft/.

Ji, R., Davis, C. S., Chen, C., Townsend, D. W., Mountain, D. G., & Beardsley, R. C. (2008). Modeling the influence of low-salinity water inflow on winter-spring phytoplankton dynamics in the Nova Scotian Shelf–Gulf of Maine region. *Journal of Plankton Research, 30*(12), 1399-1416.

Jung, H. B. (2020). Geochemical and hydrological study of coastal groundwater discharging to an urban estuary in northern New Jersey. Environmental Earth Sciences, 79(6), 158.

Keafer, B. A., Churchill, J. H., McGillicuddy, D. J., & Anderson, D. M. (2005). Bloom development and transport of toxic Alexandrium fundyense populations within a coastal plume in the Gulf of Maine. Deep Sea Research Part II: Topical Studies in Oceanography, 52(19), 2674-2697.

Kelley, A. R., Kelley, J. T., Belknap, D. F., & Gontz, A. M. (2011). Coastal and Terrestrial Impact of the Isostatically Forced Late Quaternary Drainage Divide Shift, Penobscot and Kennebec Rivers, Maine, USA. *Journal of Coastal Research, 27*(6), 1085-1093, 1089.

Kramer, S. J., Roesler, C. S., & Sosik, H. M. (2018). Bio-optical discrimination of diatoms from other phytoplankton in the surface ocean: Evaluation and refinement of a model for the Northwest Atlantic. *Remote Sensing of Environment, 217*, 126-143.

Kramer, S. J., & Siegel, D. A. How Can Phytoplankton Pigments Be Best Used to Characterize Surface Ocean Phytoplankton Groups for Ocean Color Remote Sensing Algorithms? *Journal of Geophysical Research: Oceans, n/a*(n/a).

Lagus, A., Suomela, J., Weithoff, G., Heikkilä, K., Helminen, H., & Sipura, J. (2004). Species-specific differences in phytoplankton responses to N and P enrichments and the N:P ratio in the Archipelago Sea, northern Baltic Sea. *Journal of Plankton Research, 26*(7), 779-798.

Latasa, M., & Bidigare, R. R. (1998). A comparison of phytoplankton populations of the Arabian Sea during the Spring Intermonsoon and Southwest Monsoon of 1995 as described by HPLC-analyzed pigments. *Deep Sea Research Part II: Topical Studies in Oceanography, 45*(10), 2133-2170.

Lasaga, A. C., Soler, J. M., Ganor, J., Burch, T. E., & Nagy, K. L. (1994). Chemical weathering rate laws and global geochemical cycles. *Geochimica et Cosmochimica Acta, 58*(10), 2361-2386.

Litchman, E., de Tezanos Pinto, P., Edwards, K., Klausmeier, C., Kremer, C., & Thomas, M. (2015). Global biogeochemical impacts of phytoplankton: A trait-based perspective. Journal of Ecology, 103, 1384–1396.

Lu, L., Jiang, T., Xu, Y., Zheng, Y., Chen, B., Cui, Z., & Qu, K. (2018). Succession of phytoplankton functional groups from spring to early summer in the central Bohai Sea using HPLC–CHEMTAX approaches. *Journal of Oceanography, 74*(4), 381-392.

Margalef, R. (1958). Temporal succession and spatial heterogeneity in phytoplankton. In: Buzzati-Traverso, A. A. (ed.) Perspective in marine biology. *University of California Press*, Berkeley and Los Angeles, p. 323-349.

Martiny, A. C., Lomas, M. W., Fu, W., Boyd, P. W., Chen, Y.-l. L., Cutter, G. A., . . . Moore, J. K. (2019). Biogeochemical controls of surface ocean phosphate. Science Advances, 5(8), eaax0341.

Mcgillicuddy, D. J., Jr, Signell, R. P., Stock, C. A., Keafer, B. A., Keller, M. D., Hetland, R. D., & Anderson, D. M. (2003). A mechanism for offshore initiation of harmful algal blooms in the coastal Gulf of Maine. *Journal of Plankton Research, 25*(9), 1131-1138.

Mouri, G., Takizawa, S., & Oki, T. (2011). Spatial and temporal variation in nutrient parameters in stream water in a rural-urban catchment, Shikoku, Japan: Effects of land cover and human impact. *Journal of Environmental Management, 92*(7), 1837-1848.

Paerl, H. W. (1997). Coastal eutrophication and harmful algal blooms: Importance of atmospheric deposition and groundwater as "new" nitrogen and other nutrient sources. *Limnology and Oceanography, 42*(5part2), 1154-1165.

Paerl, H. W. (2006). Assessing and managing nutrient-enhanced eutrophication in estuarine and coastal waters: Interactive effects of human and climatic perturbations. *Ecological Engineering, 26*(1), 40-54.

Pratt, B., & Chang, H. (2012). Effects of land cover, topography, and built structure on seasonal water quality at multiple spatial scales. *Journal of Hazardous Materials, 209-210*, 48-58.

Quéré, C. L., Harrison, S. P., Prentice, I. C., Buitenhuis, E. T., Aumont, O., Bopp, L., . . . Wolf-Gladrow, D. (2005). Ecosystem dynamics based on plankton functional types for global ocean biogeochemistry models. *Global Change Biology, 11*(11), 2016-2040.
Rebuck, N. D., & Townsend, D. W. (2014). A climatology and time series for dissolved nitrate in the Gulf of Maine region. Deep Sea Research Part II: Topical Studies in Oceanography, 103, 223-237.

Reynolds, C. S. (1994). The long, the short and the stalled: on the attributes of phytoplankton selected by physical mixing in lakes and rivers. *Hydrobiologia, 289*(1), 9-21.

Sanders, R., Klein, C., & Jickells, T. (1997). Biogeochemical Nutrient Cycling in the Upper Great Ouse Estuary, Norfolk, U.K. *Estuarine, Coastal and Shelf Science, 44*(5), 543-555.

Seyam, M., & Othman, F. (2014). The Influence of Accurate Lag Time Estimation on the Performance of Stream Flow Data-driven Based Models. *Water Resources Management, 28*(9), 2583-2597.

Smith, P. C., Houghton, R. W., Fairbanks, R. G., & Mountain, D. G. (2001). Interannual variability of boundary fluxes and water mass properties in the Gulf of Maine and on Georges Bank: 1993–1997. *Deep Sea Research Part II: Topical Studies in Oceanography, 48*(1), 37-70.

Snyder, J., Boss, E., Weatherbee, R., Thomas, A. C., Brady, D., & Newell, C. (2017). Oyster Aquaculture Site Selection Using Landsat 8-Derived Sea Surface Temperature, Turbidity, and Chlorophyll a. *Frontiers in Marine Science, 4*(190).

Song, H., Ji, R., Stock, C., & Wang, Z. (2010). Phenology of phytoplankton blooms in the Nova Scotian Shelf–Gulf of Maine region: remote sensing and modeling analysis. *Journal of Plankton Research, 32*(11), 1485-1499.

Sullivan, B. E., Prahl, F. G., Small, L. F., & Covert, P. A. (2001). Seasonality of phytoplankton production in the Columbia River: A natural or anthropogenic pattern? *Geochimica et Cosmochimica Acta, 65*(7), 1125-1139.

Thompson, Borns (1985). Late Pleistocene history of north-eastern New England and adjacent Quebec [Canada]: Borns, H.W.Jr., Pierre LaSalle and W.B. Thompson (eds.), Spec. Pap. geol. Soc. Am., 197:159pp; 14 papers. Inst. for Quat. Stud., Univ. of Maine, Orono, ME 04469, USA. (1985). *Deep Sea Research Part B. Oceanographic Literature Review, 32*(9), 749.

Townsend, D. W. (1998). Sources and cycling of nitrogen in the Gulf of Maine. Journal of Marine Systems, 16(3), 283-295.

Townsend, D. W., Pettigrew, N. R., & Thomas, A. C. (2005). On the nature of Alexandrium fundyense blooms in the Gulf of Maine. *Deep Sea Research Part II: Topical Studies in Oceanography, 52*(19), 2603-2630.

Townsend, D. W., Rebuck, N. D., Thomas, M. A., Karp-Boss, L., & Gettings, R. M. (2010). A changing nutrient regime in the Gulf of Maine. *Continental Shelf Research, 30*(7), 820-832.

Turner, R. E., Rabalais, N. N., Justic, D., & Dortch, Q. (2003). Global patterns of dissolved N, P and Si in large rivers. *Biogeochemistry, 64*(3), 297-317.

Tyrrell, T. (2019). Redfield Ratio☆. In J. K. Cochran, H. J. Bokuniewicz, & P. L. Yager (Eds.), *Encyclopedia of Ocean Sciences (Third Edition)* (pp. 461-472). Oxford: Academic Press.

Wolovick, M. J. (2009). Density-driven subtidal circulation in eastern Casco Bay and its effects on Alexandrium blooms. Bowdoin College Honors Thesis, 70 pp.

Xenopoulos, M. A., Downing, J. A., Kumar, M. D., Menden-Deuer, S., & Voss, M. (2017). Headwaters to oceans: Ecological and biogeochemical contrasts across the aquatic continuum. *Limnology and Oceanography*, *62*(S1), S3-S14.

Zhang, M., Zhang, Y., Yang, Z., Wei, L., Yang, W., Chen, C., & Kong, F. (2016). Spatial and seasonal shifts in bloom-forming cyanobacteria in Lake Chaohu: Patterns and driving factors. Phycological Research, 64(1), 44-55.

Appendix

Table 1. Main stem Kennebec River station sub-watershed bedrock and land use area coverage and percent coverage. The bedrock section describes each rock type observed within each sub-watershed and its respective area and percent coverage. The land use section describes each land use type within each sub-watershed and its respective area and percent coverage. Bedrock and land use data was obtained from USGS and was processed in ArcGIS.

Bedrock				**Land Use**		
Sub-Watershed	**Rock Type**	**Area Coverage (km^2)**	**Percent Coverage (%)**	**Land Use Type**	**Area Coverage (km^2)**	**Percent Coverage (%)**
K01	mudstone	347.95	40.62	Forested	449.51	52.49
	sandstone	166.73	19.46	Open Fresh Water	314.63	36.74
	gabbro	116.71	13.62	Swamp	52.15	6.09
	felsic volcanic rock	56.54	6.6	Floodplain	13.55	1.58
	melange	52.26	6.1	Harvested Forest	10.44	1.22
	limestone	27.52	3.21	Barren land	4.46	0.52
	greenstone	23.22	2.71	Cultivated Land	2.19	0.26
	diorite	17.48	2.04	Wetland	1.15	0.13
	quartzite	13.75	1.6			
	metasedimentary rock	12.45	1.45			
	rhyolite	8.95	1.05			
	granite	8.54	1			
	slate	4.37	0.51			
	schist	0.16	0.02			
K03	quartzite	23.52	34.21	Forested	77.14	81.89
	mudstone	14.62	21.26	Swamp	8.41	8.92
	melange	10.98	15.97	Harvested Forest	3.54	3.75
	sandstone	10.77	15.67	Open Fresh Water	2.33	2.47
	metasedimentary rock	6.97	10.13	Floodplain	2.33	2.47
	phyllite	1.06	1.54	Developed	0.87	0.92
	gabbro	0.44	0.64	Barren Land	0.59	0.63
	mafic volcanic rocks	0.4	0.59	Wetland	0.09	0.1
K05	quartzite	51.15	78.99	Forested	43.87	67.8
	schist	8.18	12.63	Developed	6.42	9.93

	shale	5.12	7.9	Cultivated Land	3.42	5.29
	sandstone	0.3	0.47	Open Fresh Water	2.93	4.53
				Floodplain	2.81	4.35
				Swamp	2.78	4.3
				Harvested Forest	2.13	3.29
				Barren land	0.29	0.45
				Wetland	0	0.01
K07	sandstone	64.32	96.99	Forested	38.83	58.57
	schist	1.5	2.26	Cultivated Land	7.87	11.87
	marble	0.32	0.48	Developed	6.86	10.35
	metasedimentary rock	0.11	0.16	Floodplain	4.87	7.35
	quartzite	0.07	0.1	Swamp	4.09	6.16
				Open Fresh Water	2.22	3.35
				Harvested Forest	1.45	2.18
				Barren Land	0.06	0.1
				Wetland	0.02	0.03
K10	slate	56.87	37.28	Forested	71.63	46.97
	sandstone	50.35	33	Developed	24.65	16.16
	quartz monzonite	18.29	11.99	Cultivated Land	23.71	15.55
	calc-silicate rock	15.34	10.05	Swamp	14.93	9.79
	mudstone	11.71	7.67	Floodplain	8.82	5.79
	marble	0.02	0.01	Open Fresh Water	5.61	3.68
				Harvested	1.74	1.14
				Barren Land	1.14	0.75
				Salt Marsh/Estuary	0.02	0.01
				Wetland	0.01	0.01
K11, K12, K13	slate	116.65	78.33	Forested	65.13	43.74
	granite	15.78	10.6	Salt Water	35.66	23.95
	mafic metavolcanic rock	7.58	5.09	Developed	14.74	9.9
	metavolcanic rock	4.46	2.99	Salt Marsh/Estuary	11.44	7.68
	metasedimentary rock	3.27	2.19	Swamp	11.34	7.61
	granodiorite	0.86	0.58	Floodplain	3.98	2.67
	limestone	0.33	0.22	Cultivated Land	3.18	2.13
				Harvested Forest	2.56	1.72
				Open Fresh Water	0.49	0.33
				Barren Land	0.34	0.23
				Wetland	0.02	0.01

Table 2. Mean ± standard deviation N:P and N:Si nutrient ratios for all mainstem KR stations and HS (station 14) from 2011-2013.

Station	N:P	N:Si
	Mean±SD	Mean±SD
K01	115.6±94.4	0.09±0.06
K03	139.7±168.4	0.06±0.03
K05	95.1±116.2	0.05±0.05
K07	17.7±9	0.04±0.03
K10	22.5±9.6	0.04±0.02
K11	32.4±25.9	0.05±0.06
K12	21.4±15.4	0.06±0.05
K13	6.2±2.7	0.04±0.02
14 (HS)	0.3±0.2	0.05±0.03

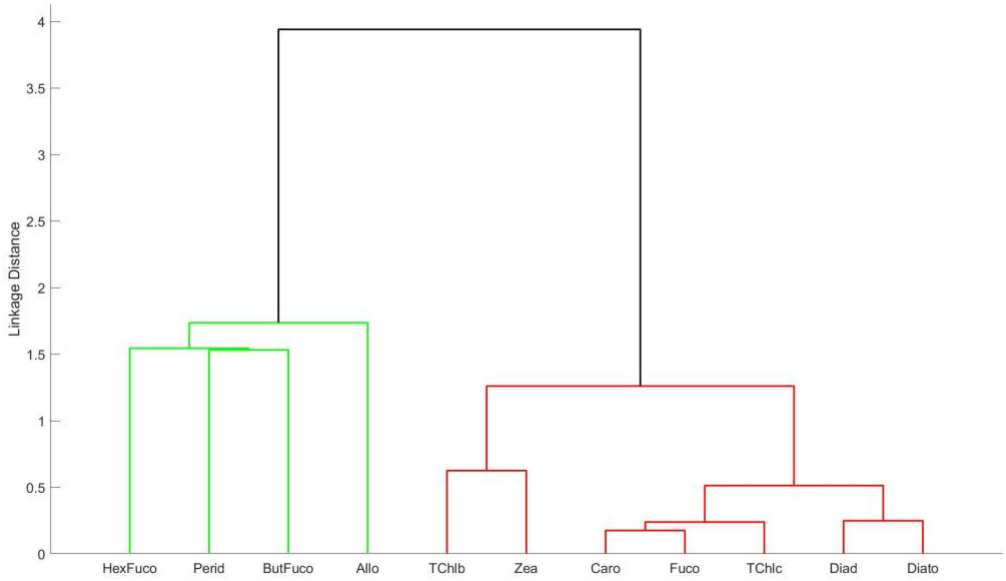


Figure 1. Hierarchical cluster analysis of downstream KR stations K11, K12, and K13 pigments using a correlation distance of 1-R (R=Pearson's correlation coefficient between pigments) and Ward's linkage method. Each cluster is color coded based on the dominant taxonomic value.

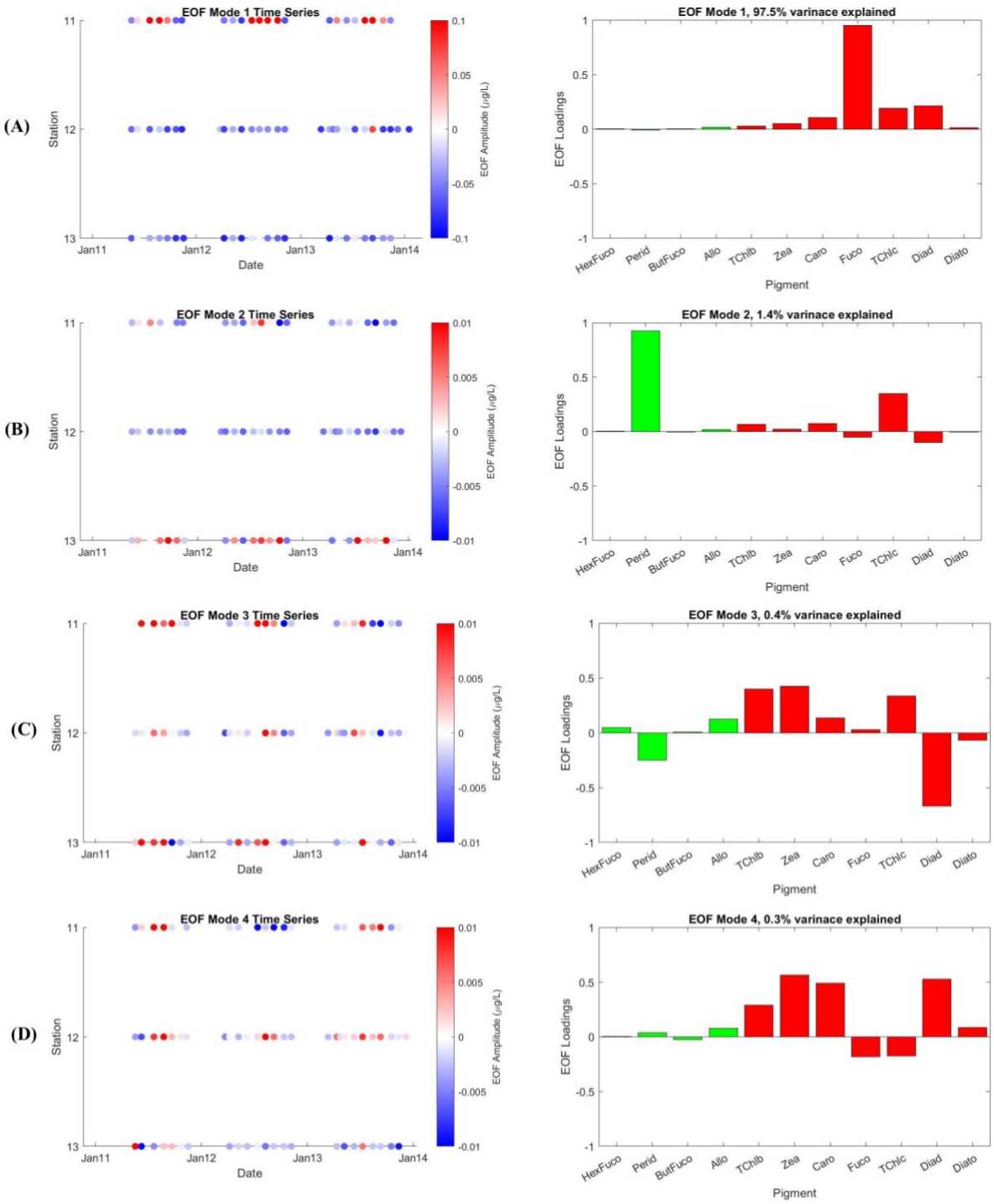


Figure 2. Pigment Empirical Orthogonal Function loadings of modes 1-4 (A-D) (right) and respective time series (left) for the Kennebec River downstream stations K11, K12, and K13 from 2011-2013. Bar plots are color coded to match the cluster analysis for downstream stations (Figure 1, Appendix). The mode number and percent variance explained by each mode are displayed above each bar plot. Time series plots are color coded by EOF amplitude.