1

Investigating possible causal relations among physical, chemical and biological variables across regions in the Gulf of Maine

Knut L. Seip

a)   Oslo and Akershus University College for applied sciences, Engineering department,

Pilestredet 35, Oslo, Norway. E-mail: knut.lehre.seip@iu.hio.no;

Tlf +4722453210; Fax 22453205

Abstract

We examine potential causal relations between ecosystem variables in four regions of the Gulf of Maine under two major assumptions: i) a causal cyclic variable will precede, or lead, its effect variable; e.g., a peak (through) in the causal variable will come before a peak (through) in the effect variable. ii) If physical variables determine regional ecosystem properties, then independent clusters of observations of physical, biological and interaction variables from the same stations will show similar patterns. We use the Leading – lagging, LL- strength method to establish leading strength and potential causality, and we use Principal component analysis, PCA, to establish if regions differ in their ecological characteristics. We found that several relationships for physical and chemical variables were significant, and consistent with "common knowledge" of causal relations. In contrast, relationships that included biological variables differed among regions. In spite of these findings, we found that physical and chemical characteristics of near shore and pelagic regions of the Gulf of Maine translates into unique biological assemblages and unique physical – biological-interactions.

Key words: ecosystem dynamics, plankton, estuaries, causality, succession, habitats, Gulf of Maine   .

Introduction

Finding the rules for plankton assembly and succession has been a challenge both for fresh – and marine water systems. Some authors hold that species succession in ocean ecosystems are almost impossible to predict, e.g., Behrenfeld and Boss (2014) and Banse (2013). Here we first examine if it is possible to establish successional rules for physical, chemical and biological variables in the semi-enclosed waters of the Gulf of Maine, and secondly if there are specific rules for regions within the gulf that have different morphologies, and that are exposed to different physical and chemical forcings. The differences may be important because near shore and open regions may play complementary roles in the oceanic ecosystem, and they may be differently impacted by changes in the global climate (Elliott and Whitfield 2011; Llope et al. 2011; Ji et al. 2013).

A causal relationship requires the cause to come before the effect. If cause and effect variables are represented by cyclic time series, like the time series for a prey and its predator, we would expect the prey to peak before the predator. It does so in the Lotka-Volterra formulation, Lotka (1924). In this study, we will say that the cause is a leading variable to the effect in the sense that its peaks and throughs comes before the peaks and throughs of the effect variable. However, our measure for "leading" extends to the full series, not only to peaks and throughs. A causal variable may not represent a direct cause. The variable itself may be a proxy for the real cause, e.g., light is a proxy for heat transfer to the ocean. It may also be the observed link in a chain of causes and effects, e.g., temperature causing stratification that again limits nutrient transport. In economics, leading, coinciding and lagging indexes are used to predict, or verify, changes in

business cycles (Seip and McNown 2007). The method section gives details.

For temperate fresh water systems there has emerged a conceptual model for the succession of events formulated by Sommer et al. (1986), the so called Plankton Ecology Group, PEG - model. There are indications that there also will be successional traits in marine waters. Conceptual models for marine systems that are similar to the fresh water models have been developed by e.g., Huntsman and Barber (1977, Fig 6), Wilkerson et al. (2006), Song et al. (2011, Fig 2), and Doney et al. (2009). Ecosystem models, like the one in Song et al. (2011) are often depicted with arrows pointing from nutrients to phytoplankton to zooplankton, the N → P → Z model. Arrows will normally indicate that phytoplankton take up nutrients (e.g., phosphorus and nitrogen) and zooplankton graze on phytoplankton. Leading and lagging relationships will differ in the mechanisms that are candidate causal mechanisms for the succession. Unidirectional food- consumer relations may be important for some variables. For other LL- relationships, physical and chemical characteristics of a species may determine its position along successional and eutrophic gradients, Seip and Reynolds (1995).

Current hypotheses.

The Sverderup (1953) "critical depth" hypothesis links temperature, stratification and phytoplankton biomass. Chiswell (2011) details the theory and argue that blooms occur when surface chlorophyll is contained in density layers that corresponds to temperature differences of 0.1ºC or less, getting sufficient light and nutrients, and are then destroyed by lack of nutrients, grazing by zooplankton, deepening of the mixed layer, overturn events, or lateral movements of water masses (Behrenfeld and Boss 2014). For freshwater we would assume

that the first zooplankton bloom occur just after the first spring phytoplankton bloom, SPB, creating the so called "clear water phase", and that it is then followed by a fall phytoplankton bloom, FPB, (Sommer et al. 1986). All of these mechanisms would require some variables to lead other variables, but the consistency and duration of leading - lagging sequences may vary.

Observations.

The sequence SPB and FPB is observed in many gulf ecosystems, e.g., the Gulf of Maine (Song et al. 2010) and the Adriatic Sea, (Mozetic et al. 2012). It is reported for the subtropical waters off east New Zealand (Chiswell 2011) and Marshall and Peters (1989) observed it for lakes. Kahru et al. (2011) observed for the Arctic that early decrease in ice concentrations allowed phytoplankton bloom maximums to occur earlier. Light is a proxy for heat transfer to the ocean (e.g. as $W\ m^{-2}$) and will normally be a leading variable to sea surface temperature, SST. This "general knowledge" will be used as a benchmark for our method.

Firstly, we hypothesize that we will obtain significant leading – lagging, LL- relations corresponding to the succession of events that can be predicted from conventional ecosystem theory, e.g., phytoplankton abundance will come before zooplankton abundance. In particular, we hypothesizes that we will obtain the sequence: nutrients – phytoplankton –zooplankton, $N \rightarrow P \rightarrow Z$. However, findings by Tømte et al. (1998), Hsieh et al. (2005), and Behrenfeld et al. (2013) show that subtle disruptions in food web equilibrium and dynamic chaos is present at higher trophic levels and may, together with stochastic events, destroy successional patterns. In addition, the normal monthly sampling frequency may be too low to allow conclusions.

Secondly, we hypothesize that the ecosystems of different regions in the Gulf of Maine will develop differently, that is, based on morphological, physical and chemical characteristics of the regions, presence and absence of species groups will be different, and the species groups will interact differently. However, water masses shift between regions and stochastic events may be more important than regularities in the forcing functions. Contradicting views can be found in Brooks (2009), Ji et al. (2013) and Anderson et al. (2014). Studies by Wong et al. (2007) and Valesini et al. (2010) as well as studies summarized in Valesini et al. address whether regions can be classified based on physical and chemical criteria so that biological characteristics can be predicted from the classification. However, these studies do not include interactions between variables. We will accept our second hypothesis if sites that are different in morphology and physical and chemical characteristics are also significantly different in biological characteristics and in the way variables interact.

Capturing relations between cyclic time series.

To test our hypotheses, we make comparisons in terms of physical and chemical variables, in terms of biological variables and in terms of interaction between variables. We measure interactions by comparing two measures, one that express associations between paired time series and one that expresses the "before" and "after" relationships between paired cyclic variables. For cyclic time series the two measures corresponds to measures of pro – cyclisity / counter – cyclisity (that is positive and negative associations respectively) and leading and lagging, LL- relationships. The technique for comparing LL- relationships between variables is new to ecology. The method can be applied to very short time series, n > 3, it identifies outlier events in otherwise regular LL- relationships, and it will detect breakpoints where e.g., a persistent leading relationship changes into a

persistent lagging relationship. We believe this is an important task for the study of changes in match- and mis-matches between species that interact with each other.

In the rest of the paper we first present location and materials, thereafter we give an outline of our two major methods: the LL-strength method that contribute to the possible identification of causal relationship and the novel application of principal component analysis, PCA, that help distinguish ecosystem characteristics. Lastly, we present and discuss the results.

Location and materials

The Gulf of Maine experiences a tidal range that exceeds 3 m, leading to complex and vigorous circulation patterns (Brooks 2009). The study sites are located in the western Gulf of Maine and stretches from the Merrimac River in the south to Kennebec River in the north, Figure 1.

The area stretches out about 75 km offshore (coordinates for the farthest offshore station is $42^o85'$,$-69^o86'$). The stations can be divided into two series, a transect going from the near shore and out to deep waters of Wilkinson Basin (the WB stations) and a coastal transect along the shore (the CT stations). Station depths along the WB transect ranged from 20 m near shore to 270 m offshore, and the CT stations ranged from 20 m to 100 m. A particular station, CT4, was located about 2000 m west of the mouth of the Kennebec River. The river has a flow volume in the range 1000 to 6000 $m^3.s^{-1}$ and turn to its right after leaving the estuary. It is at least an order of magnitude larger than any local source of freshwater to the Bay, (Janzen et al. 2005). The CT4 station is well within the influence zone of the Kennebec river as indicated by salinity profiles around the mouth (Salisbury et al. 2008). The other coastal stations may be affected by alongshore coastally trapped buoyant plumes,

Franks and Anderson (1992 Figures 2-6), but probably less frequently and with less impact.

During the period January 2005 to July 2008, samples of physical, chemical and biological variables were taken at 29 stations in the Gulf of Maine. Figure 2 shows observations from station WB3 normalized to unit standard deviation and shifted 3 units relative to each other for clarity. Satellite observations give information on surface chlorophyll concentrations,, e.g. Behrenfeld (2010), and surface samples may show stronger seasonal changes than depth averaged time series (Chiswell 2011 Fig 4). The physical variables were daily averages of sea surface temperature, T, ^{o}C, light, L, as daily Photosynthetic Active Radiation, PAR, $\mu E\ m^{-2}s^{-1}$, wind, W; as the cube of the wind speed, U3 m^3s^{-3}. The chemical variables were salinity as Practical Salinity Units, PSU, nitrogen as the sum of nitrite NO_2^- and nitrate $NO_3^{-,}$ designated NOx (mg m^{-3}), and Orthophosphate PO_4 (mg m^{-3}). The bio-logical samples were measured as chl-a (mg m^{-3}), C, and as the fractions of diatoms, flagellates and cyanobacteria derived from HPLC pigment concentrations and CHEMTAX (Mackey et al. 1996). The fractions were multiplied by chl-a to get an expression of the biomass of each species group. All samples were surface samples, taken down to 1 or 2 meters depending upon data availability. There were observed 21 species of zooplankton, the most abundant being *Calanus finmarchicus* and *Oithona similis* (ind. m^{-3}, unfortunately, neither mass nor length measurements were taken for the zooplankton). Samples were taken from 0 to 20 m depth to include the effects of vertical migratory behavior. As a proxy for zooplankton abundance, we used the sum of all sampled individuals.
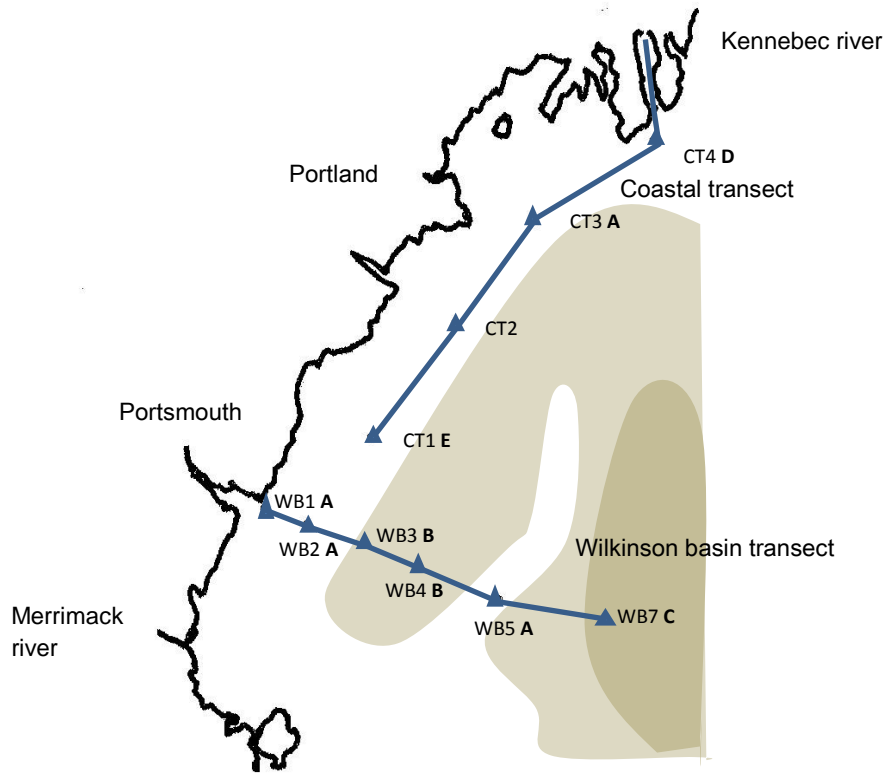
Figure 1 Gulf of Maine with coastal transect (CT-stations) and Wilkinson basin transect (WB – stations). Darker shades show increasing depths. Letters in bold identify sites that are similar in terms of 13 equally weighted morphological, chemical and biological characteristics. The station CT2 did not have sufficient number of observations to be included, see text.

The sampling frequency in this study was about once a month, occasionally twice, at each of nine stations during the summer half year from April to September, and less frequently during the winter half year. During some winter months, no observations were taken. Two hundred and eighty two (282) samples were taken including all variables. All data available from GoMOOS (2010) now NERACOOS (2013). Details of sampling and sample preparation is given in Moore (2008). We use i) the 2005-2009 data set for temperature, light, wind, salinity, nitrogen, phosphorus, phytoplankton and zooplankton for the study of possible causal relationships (hypothesis 1) and ii) the complete data set 2005-2009 to group the stations into regions that may show distinct ecological characteristics (hypothesis 2)..
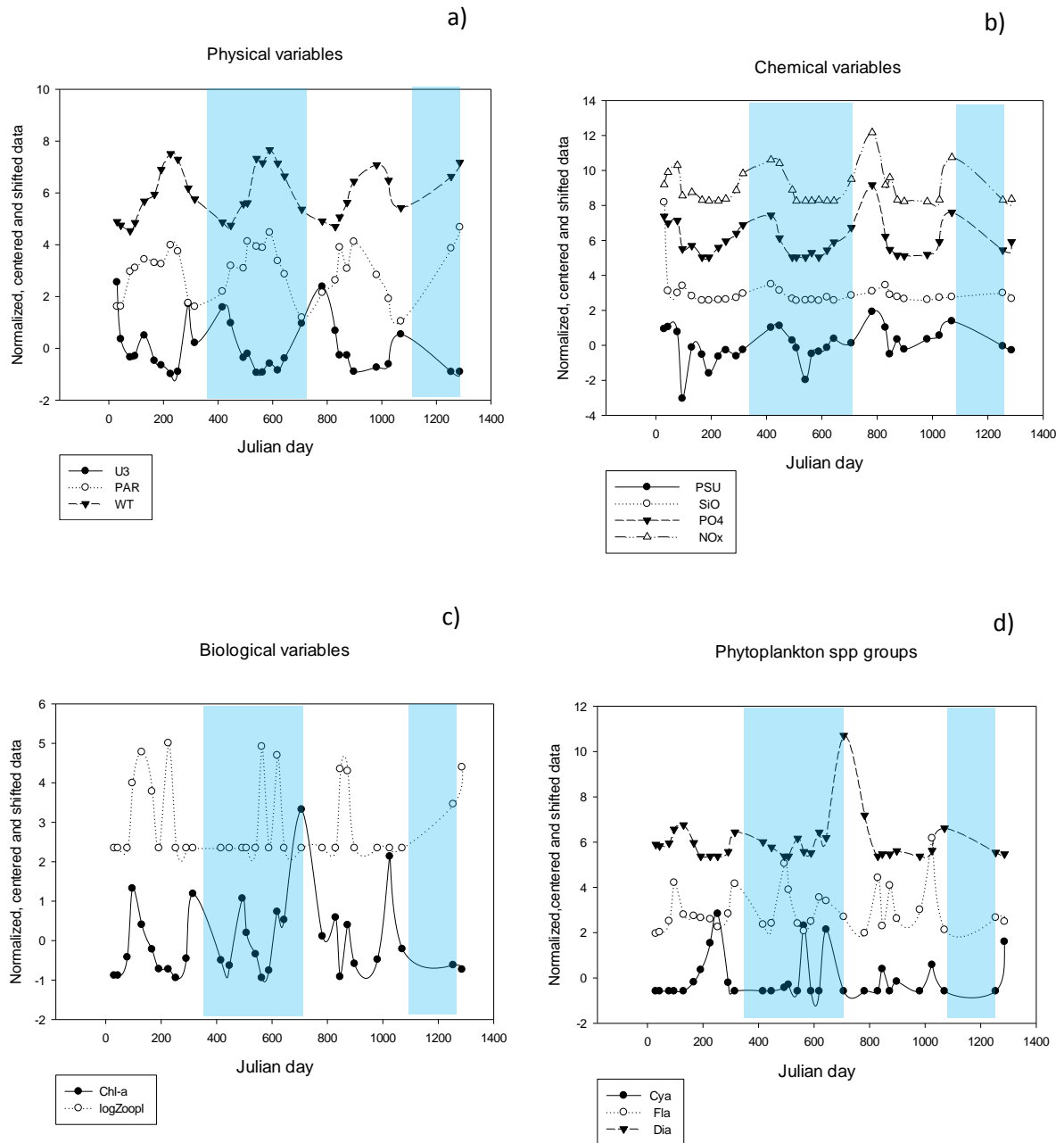
6

Figure 2 Observations of physical and chemical and biological data from the ocean habitat B (station WB3) during the period January 2005 to July 2008 in the Gulf of Maine. The x-axis shows days after the first sample Jan 30, 2005. All data were normalized to unit standard deviation, but shifted 3 units relative to each other for clarity. a) Physical variables: WT = surface water temperature, PAR = Light, U3 = wind; b) Chemical variables: PSU = salinity; SiO = silica, PO$_4$ = orthophosphate, NO$_x$ = sum of nitrite and nitrate; c) Biological variables: Chl-a = phytoplankton as Chl-a, Log Zoopl = The logarithm of zooplankton counts. d) Phytoplankton species groups: Dia = diatoms, Fla = Flagellates, Cya = Blue-greens.   The blue shaded areas enclose data for one year.

Methods

We first give an outline of the data pretreatment. Thereafter, we describe specific features of each method that are relevant for our application.

Pretreatment of the data

All numerical calculations were applied to time series that were normalized to unit standard deviation. This eliminates any effects of measuring units, and is also required for calculating the strength of the leading and lagging relationships that will be explained below.

Smoothing of the time series emphasizes different features of the series. Here we smooth the observed series to remove noise. Removing noise may interfere with small changes that in some cases can be established as real because they are drivers for effects that appears later, e.g, nutrient pulses that causes a subsequent increase in chl-a (Mozetic et al. 2012). We therefore smoothed the observations only lightly in the present study. We used the LOWESS algorithm with 2nd to 4th order interpolation and $1/4^{th}$, $1/6^{th}$ and $1/10^{th}$ of the series length as moving average period, SigmaPlot©.

Grouping observations.

Since we wanted to examine regional differences among ecosystems in the Gulf of Maine, we merged data sets from several stations where the data were "similar" in a certain sense. To do this, we first applied Principal component analysis, PCA, Camo A/S ©, to the full data set (282 samples and 14 variables) including the morphological variables (depth, distance from land), physical variables, chemical variables and biological variables. Since PCA was applied to normalized data, each variable had the same weight, and we used cross validation to establish model confidence. The score plot of the PCA gives coordinates for stations that are similar in a least square sense. By applying PCA we avoid problems with coo-linearity among the variables.

We then applied a hierarchical clustering analysis, Sysstat ©, to the scores on the 1st and the 2nd principal components of a PCA. Most of the variance in the data are explained by the two first principal components, PC_1 and PC_2. Including more components increases the risk of explaining noise in the data. Those stations that were separated at a low level were merged as one region and those that were separated on a high level (2/3 of the distance separation scale) were defined as separate regions. Assessment of significance was made in the PCA plots. We found seven clusters of stations based on non-biological and biological characteristics at the sampling sites. Five of these clusters (A to E) had sufficient number of samples to continue the analysis. (> 20 samples gives about one or two samples per month during an extended summer half year). The resulting clusters are shown as regions A to E in Table 1. The Table also shows average values and their standard deviation for the variables at the resulting regions. Note that the ratio of the standard deviations for temperature and salinity observations is about 5:1. Since the variables contribute to the density gradients in the water with a ratio of about 1 to 5, these two variables contribute about equally to the density gradients in the water. The map in Figure 1 show how the sites are distributed graphically. For some LL- studies, we used observations from single stations to avoid averaging effects.

The PCA produces a score plot and a loading plot for each of the sets. The score plot shows how observations are related and the loading plot shows how variables are related. Variables that are positioned at similar positions in the loading plot and the score plot will characterize each other. Also, variables that are positioned opposite along a line through origin will contribute to the characterization, but with numerically low values. Variables that are connected to the origin at right angles to each other are either unrelated or having a leading –

lagging relationship to each other. If interactions between pairs of species follow Lotka –Volterra dynamics (Lotka 1924) the association will result in a regression coefficient of R ≈ 0, rather than a negative coefficient (Holmengen and Seip 2009). The score plot we obtain in this study shows how observations are related. Since we are interested in the relationship between observations within each region, we calculated the average value of the scores (PC_1 and PC_2 values) for all observation that belonged to regions A, B, C, etc. Thus, we got a new score plot that shows the region centers.

The leading-lagging, LL –strength method for identifying possible causal relationships.
With our method, the sequence of time series can be diagnosed for LL- relationships in the phase plot (Draftsman plot) for paired series. For example, if we compare the time series for light, PAR, and water temperature, WT, at the ocean station, Figure 2 a, we see that the shift between PAR and WT are less than ¼ of a cycle length, λ, relative to each other. For perfect sines a shift by ¼ λ means that synoptic values of the two series, when regressed, will obtain close to zero explained variance, r^2, in spite of their relationship to each other. Our method calculates LL- relations as moving averages over paired time series (n = 3). Visually, LL- relations are easiest distinguished by comparing peaks and throughs between the two series.

Table 1 Habitat characteristics

A to E are cluster of observations identified in the study. (±) identify standard deviation of the numbers.

Sites	A	B	C	D	E
Characteristics	Shallow water	Deep water	Ocean water	River mouth	Coastal water
Depth, m	48 ± 15	123 ± 20	259 ± 2	28 ± 3	67± 4
Stations	WB1-2,WB5, CT3	WB3-4	WB7	CT4	CT1
#samples	115	62	24	23	29
Distance from land. km	16 ± 15	22 ± 6	63	4.5	13
NO_X, N (mg.m^{-3})	2.79 ± 3.60	3.29 ± 4.05	3.34 ± 4.14	3.61 ± 3.53	3.01 ± 3.34
PO_4, P (mg.m^{-3})	0.37 ± 0.29	0.39 ± 0.32	0.33 ± 0.27	0.41 ± 0.29	0.36 ± 0.27
Temp. T °C	11.29 ± 5.67	11.16 ± 5.74	10.61 ± 5.26	12.06 ± 5.43	11.23 ± 5.96
Light, L, µE m^{-2}s^{-1}	34.19 ± 13.68	33.15 ± 12.93	35.42 ± 12.42	35.74 ± 13.19	34.29 ± 14.41
Salinity, S (PSU),	31.14 ± 1.19	31.46 ± 0.88	32.12 ± 0.70	29.53 ± 1.10	31.25 ± 0.96
Wind, W, U^3. m^{3}s^{-3}	338 ± 296	360 ± 297	366 ± 311	259 ± 259	345 ± 305
Chl-a, C, mgm^{-3}	1.44 ± 1.14	1.16 ± 0.82	0.87 ± 0.63	2.59 ± 1.82	1.64 ± 2.31
Diatoms, Dia (mg.m^{-3})	0.48 ± 0.77	0.37 ± 0.52	0.17 ± 0.20	1.30 ± 1.31	0.77 ± 2.21
Flagellates, Fla (mg.m^{-3})	0.93 ± 0.91	0.76 ± 0.63	0.67 ± 0.57	1.28 ± 0.84	0.85 ± 1.07
Cyanobacteria, Cya (mg.m^{-3})	0.02 ± 0.04	0.02 ± 0.04	0.02 ± 0.02	0.01 ± 0.01	0.03 ± 0.06
Zooplankton, Z (ind.m^{-3})	832.42±1611	972.06±2248	605.22±1740	797.29±1718	855.38±1447

To facilitate description of the method, we first think of the peak - through sequence of the observations as a pair of "observed" sines that both have wavelength $\lambda = 2\pi$ but that are shifted in time relative to each other.

(1) $\quad y_1 = \sin(\omega t) ; y_2 = \sin(\omega t + \tau\ )$

We can then describe the relationship between the two sines by two parameters. The first i) is a rotational direction, V, between sequential trajectories through points i-1, i, i+1 in the phase plot for the sines, and the second, ii) is the slope, or the β – coefficient, of the scatter plots of the "observations". Four examples are shown in Figure 3a. Further explanations are given as Online resource 1.

To give the rotational direction in the phase plot a numerical expression, we calculate the angle, V, between two sequential trajectories, or vectors $\bar{v}_1$ and $\bar{v}_2$ formed by three sequential points i-1, i and i+1 in the phase plots. We use the equation from Seip and McNown (2007):

(2)

$$V = sign(\bar{v}_1 \times \bar{v}_2) \cdot Arc\cos\left(\frac{\bar{v}_1 \cdot \bar{v}_2}{|\bar{v}_1| \cdot |\bar{v}_2|}\right)$$

With two minimal series pasted as (A1, A2, A3) and (B1, B2, B3) in an Excel spread sheet, the angle, V, is calculated by pasting the following Excel expression into C2: =SIGN((A2-A1)*(B3-B2)-(B2-B1)*(A3-A2))*ACOS(((A2-A1)*(A3-A2) + (B2-B1)*(B3-B2))/(SQRT((A2-A1)^2+(B2-B1)^2)*SQRT((A3-A2)^2+(B3-B2)^2))). An Excel version of Eq (2) is shown in Online Resource 1

The angles, V, between two consecutive vectors range between -180° and + 180°. To avoid dominance of a few large angles, we express the

leading-lagging, LL – strength, LLᵒ, of a paired time series by the proportion of positive rotations (counter-clock-wise rotations by convention) relative to the total number of rotations. We then normalize the measure to range from -1 to + 1. By this convention, counter clock-wise rotations are positive and clock-wise rotations are negative:

(3)

$$LL^o = n_{pos}/(n_{pos} + n_{neg}), \ \ LL = 2 \times LL^o\text{-}1$$

The measure LL- strength captures two aspects of the cause-effect relationship between paired variables. It obtains a high / low value when one variable is consistently leading or lagging another. A consistent value requires the two series to change cycle lengths in concert. We believe that concerted cycle lengths are a supporting factor for a causal relationship between variables.

The angle V (in radians) also gives an expression for how fast processes are since cycle length, CL, and the angle V are inversely related through Eq. (4):

$$CL = N/\left|\sum_{2}^{N-1} V_{i-1,i,i+1} \frac{1}{2\pi}\right|$$

(4)

Short cycle times would correspond to fast processes. When we apply the strength measure to the time series in Figure 3b, it is seen from the phase plot in Figure 3c that trajectories rotate largely clock-wise, consistent with the LL-relation PAR → SST (temperature, SST, on x – axis and light, PAR, on y-axis).

Several studies show that a causal agent, e.g., nutrients for phytoplankton (Fussmann et al. 2005), mink prey for muskrat predators (Holmengen and Seip 2009) peak before the effect variable. With the conventional

nomenclature, nutrients / prey on x- axis and grazer / predator on y-axis, strong counter clock-wise rotation can be interpreted as a strong food - consumer pattern. The LL- strength of the paired time series in Figure 3b is - 0.46. The rotational angles for successive triplets of observations as one move forward with time from 2005-2007 can be depicted as in Figure 3d. By smoothing the series one will see trends as well as outliers. Trends that cross the zero line may identify breakpoints in the LL-relationships. Observation no 19 in the series is marked with an "A" in Figure 3 b, c and d. It shows up as an exception in an otherwise largely clock-wise rotation pattern. "A" corresponds to the observation on May 23, 2007 and it would give reasons to search for errors in the observations.

Time shifts between variables.

In addition to the LL-strength we calculate the regression coefficients between pairs of variables. For perfect sines the regression coefficient will directly express the shift in time between the paired variables, c.f., the Lissajous equation: Merino (2003) and Wikipedia (2013). A strong positive correlation, β – coefficient $\approx$ 1.0, shows that the shift is short and that the two variables are peaking at about the same time. This may indicate that a third, external factor is modulating the rise and fall of the variables. If the distance τ is close to $\pi/4$ the regression coefficient is close to zero, and in combination with a high LL- value, suggest that a food consumer relationship is present, (Fussmann et al. 2005; Holmengen and Seip 2009). If the distance between the peaks is larger than $\pi/4$ but less than 3 $\pi/4$, around $\pi/2$, the regression

coefficient is negative. This suggests that there is a competitive component to the relationship between the two variables. In our nomenclature, it is the rotational direction in the phase plots and the β – coefficients in combination that suggest a tight coupling between consumers and their food (and the prerequisite for top down control.)

Estimations of significance
For the LL- strength measure, we use Monte-Carlo simulations to find confidence bounds. The 5% confidence bounds for paired, uniformly distributed, random series of 30 observations give a rotational direction of 0.5 ± 0.13. Length of the observed time series in the present study is 23 to 115. The confidence interval were determined by increasing the number of samples geometrically from 10 to 160. We found the estimated asymptotic value for the confidence interval to be 0.1. With the transformed values for conservative LL- strength, Eq. (3), a significant positive rotation gives LL > 0.23 and a significant negative rotation gives LL $< - 0.23$.

To estimate confidence intervals for clusters in the score and loading plot, we applied Monte Carlo simulations by adding "sites" and "variables" where the values for the sites and variables were random numbers. A set of sites or variables characterized by random numbers will cluster close to the origin of the PCA plot, and standard errors along the two principal components for the distribution of the random "sites" and "variables" were used to estimate the 95% confidence interval for points in the plots. However, by adding random numbers the original data set is somewhat contaminated, so the estimates give only a guidance.
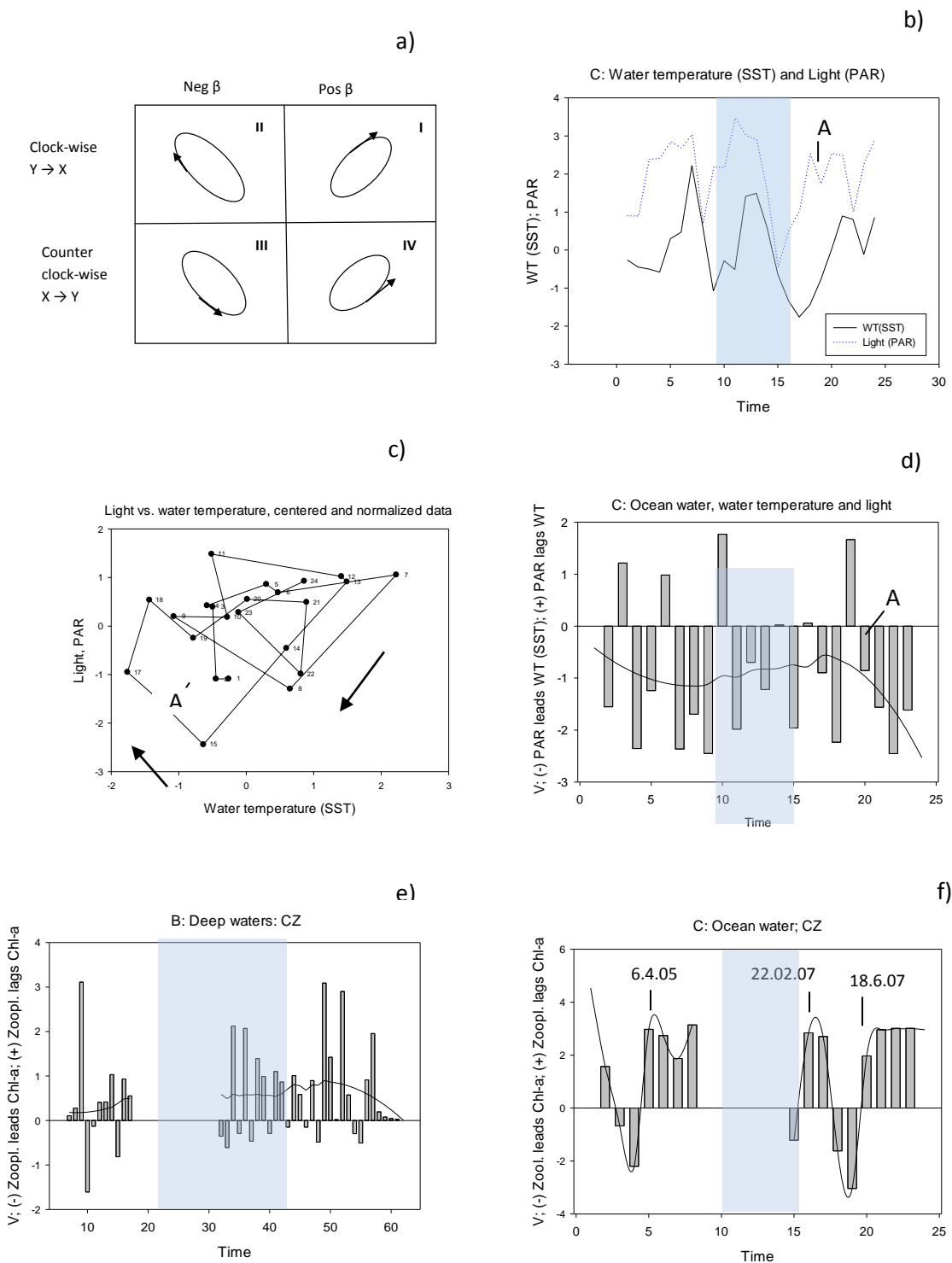
Figure 3 Calculating angles in phase portraits. a) Cyclic time series x and y plotted in phase plots on the x-axis and the y-axis respectively. The two upper panels show clock-wise rotations corresponding to y as a leading variable to x, that is, y will peak before x and represent a possible causal factor for x. The two lower panels show the opposite situation.   b) The time series for water temperature, WT (SST) and Light, PAR, 2005-2007 in the ocean habitat, C c) Phase plot of the time series in (b). Arrows indicate

dominating rotational direction. d) Histogram for the angles between successive vectors in (c). Note that the x- and y-scales in Figure c are not quite equal, distorting the angles somewhat. The LL- strength for the time series in (b) is LL= - 0.46. (27 % positive rotations). For the unsmoothed series, LL = - 0.36. Thin line show smoothed approximation to angles. Shaded are indicate observations during the middle year 2006. The letter "A" in the figure shows corresponding values for observations, phase plot and angles. The particular observation cause an exception to the otherwise clock-wise rotation. e) Histogram for angles, V, in the phase plot for chl-a (x-axis) versus log zooplankton (y-axis) in region B (deep waters.) f) Histogram for angles, V, in the phase plot for chl-a (x-axis) versus log zooplankton (y- axis) in region C (ocean waters)

We also used Monte Carlo simulations to estimate the probability that one site out of 5 should be outside 2×0.05 - confidence interval distant from its nearest neighbor site during 3 trials. The probability was $p = 0.037 < 0.05$. Thus, if one site or cluster is 2 confidence interval distant from another cluster, the two clusters probably characterize two different systems.

We examined whether the clustering of group of stations would be similar if the clustering was based on i) morphological, physical and chemical data; ii) time series for biological data and iii) the interaction between the species. For the first two analyses the matrices consist of the relevant variables as columns and the 282 samples as rows, all series normalized to unit standard deviation. For the last analysis, the matrix for the PCA analysis has the LL-strength and the β – coefficients as columns $(2 \times 19$ columns) and the regions A to E as rows (5 rows), Tables 2a and b were transposed.

Results

We first present generic results for the significant relationships among variables at all sites. Secondly, we examine how the LL-relationships and associations between variables apply to regions in the Bay area. Lastly, we present results for the separate regions and show that two regions distinguish themselves from the rest.

Generic leading- lagging, LL- relationships

For the whole material nine regression coefficients and six LL-strength relationships were significant at all sites. Our analysis showed that relationships between physical and chemical variables were more consistent across sites than the relationships between biological variables. Three of 9 physical and chemical LL-relationships and all of the β - coefficients were consistent across sites (no 1-3, 5 - 6, 8, 14, 16-17). For pairs that include biological variables 3 of 12 LL - relationships and nil of 12 β – coefficients were consistent across sites. (no 7, 9-13, 15, 18-21.)

For the LL-strength to be significant across all regions in Table 2 a, the confidence intervals for the average of all five regions had to be either below - 0.23 or above + 0.23. The significant relationships are marked with (°). For the β - coefficients the confidence interval should not overlap zero. In the discussion that follows, we only report relationships that are significant at the 0.05 level for each region. We quote the LL-strength values as $LL_{XY} = [-1,+1]$. A positive number on the right hand side shows that the y-axis variable is *lagging* the x-axis variable and a negative number shows that the y-axis variable is *leading* the x-axis variable. For significant LL-relationships we will also write $X \rightarrow Y$ for X leading Y or $Y \rightarrow X$ for Y leading X.

Table 2 Interaction parameters for sites A to E

a)   LL- strength. LL values below - 0.23 shows that there are significantly more clock-wise rotations (negative rotations) and the y-variable is a leading variable to the x-variable. LL-values above +0.23 shows that there are significantly most counter clock-wise rotations (positive rotations) and the y-variable is a lagging variable to the x-variable. "*" show that the value is significant at 5% level. (°) shows that the average values are significant at the 5% level. The LL-strength values for NOx and PO4 versus temperature that were similar for groups C, D and E were double-checked. "Acro" is acronyms for the pairs listed as x-axis and y- axis variable, e.g., the first pair, NOx and PO4, has the acronym NP. "SW" is shallow water, "DW" is Deep water, "OW" is ocean water, "RM" is River mouth and "CW" is coastal water.

No	x-var	y-var	Acro	A (SW)	B (DW)	C (OW)	D (RM)	E (CW)	aver
1	NOx	PO4	NP	-0.1	-0.2	-0.1	0.08	-0.46	-0.16
2	NOx	temp	NT	-0.12	0	-0.36*	-0.46*	-0.26*	**-0.24°**
3	NOx	wind	NW	-0.24	0.04	-0.18	0.1	-0.08	-0.07
4	NOx	Chl-a	NC	-0.3*	-0.3*	-0.08	-0.18	-0.36*	**-0.24°**
5	PO4	Temp	PT	-0.02	-0.14	-0.36*	-0.46*	-0.26*	**-0.25°**
6	PO4	Wind	PW	-0.14	0	0.18	0.1	0.54*	0.14
7	PO4	Chl-a	PC	-0.14	-0.4*	0.54*	-0.54*	-0.28*	-0.16
8	Temp	Wind	TW	-0.16	0.1	0.28*	0.46*	0	0.14
9	Temp	Chl-a	TC	0.06	0.1	0	0	0	0.03
10	Temp	Zoopl	TZ	-0.06	-0.5	0	-0.24	-0.12	-0.18
11	Wind	Chl-a	WC	-0.22	0.04	0.18	-0.36*	-0.08	-0.09
12	Wind	Zoop	WZ	-0.24	-0.02	0.12	0.26*	-0.06	0.01
13	Chl-a	Zoopl	CZ	-0.26*	0.38*	0.0	- 0.26*	-0.26*	0.08
14	temp	light	TL	-0.1	-0.1	-0.54*	-0.54*	-0.54*	**-0.36°**
15	light	Chl-a	LC	0.04	0.24	0.28*	0.1	0.1	0.15
16	Salinity	NOx	SN	- 0.16	-0.23*	-0.27	0.05	-0.33*	-0.19
17	Salinity	Temp	ST	0.07	0.03	-0.36*	0.05	-0.19	-0.08
18	Salinity	Chl-a	SC	0.00	-0.17	-0.09	0.00	-0.11	-0.07
19	Dia	Fla	DF	-0.14	-0.10	0.26*	-0.34*	-0.04	-0.07
20	Dia	Cya	DC	-0.34*	-0.40*	-0.82*	0.04	-0.50*	**-0.40°**
21	Fla	Cya	FC	-0.16	-0.34*	-0.36*	0.22	-0.56*	**-0.24°**

Physical and chemical variables. There are 3 physical and chemical pairs of variables that both show significant leading and lagging relationships and significant regression relations across all regions. Temperature and light are positively correlated at all sites (R > 0.522) and light (as a proxy for heath transfer) is a leading variable to temperature, L ≈ T, L → T (no 14; R = 0.5; LL_{TL} = - 0.36). This result corresponds to our benchmark for accepting the method and it gives confidence to the data set. Both nutrients, NO_x and PO_4, were positively related to strong winds (no 3,6, R > 0.47) and negatively related to temperature over all sampling stations (no 2, 5, R < - 0.58). Temperature was a leading variable to NO_x and PO_4 (T → NO_x, T → PO_4; LL_{NT} = - 0.24; LL_{PT} = - 0.25).

Table 2 continued)

b) Regression coefficients. Since all series were normalized to unit standard deviation slopes and regression coefficients are identical.

No	x-var	y-var	Acro	A (SW)	B (DW)	C (OW)	D (RM)	E (CW)	average
1	NOx	PO$_4$	NPs	0.734*	0.876*	0.790*	0.634*	0.795*	**0.766 º**
2	NOx	temp	NTs	-0.654*	-0.707*	-0.704*	-0.907*	-0.734*	**-0.741 º**
3	NOx	wind	NWs	0.712*	0.757*	0.656*	0.793*	0.679*	**0.719 º**
4	NOx	Chl-a	NCs	-0.202	-0.167	-0.353	-0.592*	0.094	-0.244
5	PO$_4$	Temp	PTs	-0.579*	-0.587*	-0.584*	-0.584*	-0.651*	**-0.597 º**
6	PO$_4$	Wind	PWs	0.719*	0.779*	0.685*	0.469*	0.632*	**0.657 º**
7	PO$_4$	Chl-a	PCs	-0.142	-0.096	-0.208	-0.088	0.245	-0.058
8	Temp	Wind	TWs	-0.736*	-0.666*	-0.611*	-0.759*	-0.750*	**-0.704 º**
9	Temp	Chl-a	TCs	-0.142	-0.164	0.066	0.534*	-0.323	-0.006
10	Temp	Zoopl	TZs	0.011	0.132	-0.381	-0.136	0.101	-0.055
11	Wind	Chl-a	WCs	-0.047	-0.092	-0.087	-0.617*	0.178	-0.133
12	Wind	Zoop	WZs	-0.172	0.031	-0.110	0.055	0.052	-0.0294
13	Chl-a	Zoopl	CZs	0.050	-0.035	-0.050	-0.402*	0.322	-0.0023
14	temp	light	TLs	0.6104*	0.545*	0.522*	0.647*	0.672*	**0.600 º**
15	light	chla	LCs	-0.098	-0.196	-0.006	0.373	-0.230	-0.031
16	Salinity	NOx	SN	0.509*	0.647*	0.672*	0.027	0.594*	0.490º
17	Salinity	Temp	ST	-0.433*	-0.499*	-0.738*	0.076	-0.627*	**- 0.444 º**
18	Salinity	Chl-a	SA	0.032	-0.004	0.038	0.092	0.155	0.063
19	Dia	Fla	DFs	-0.088	0.006	0.105	0.54	-0.228	0.067
20	Dia	Cya	DCs	-0.394	-0.325	-0.123	0.33	-0.278	-0.158
21	Fla	Cya	FCs	0.072	0.006	0.136	-0.15	0.010	0.015

Plankton variables.

Chl-a was generally a *leading* variable to NO$_x$, C → NO$_x$, (no 4, LL$_{NC}$ average = - 0.24; significantly at sites A, B and E). At the ocean region, LL was not significant. Chl-a was also generally a *leading* variable to PO$_4$, (no 7, C → PO$_4$), except at the ocean region were it was a significant *lagging* variable, (LL$_{PC}$ = 0.54.)

Cyanophytea generally leads diatoms and flagellates, (no 20, 21 ; LL < - 0.24), however, there are generally no positive or negative association between them.

Zooplankton abundance are largely unrelated to Chl-a, except at the river mouth region where they are counter cyclic (no 13, CZs = - 0.40). Zooplankton would both lead and lag chl-a (no 13). We found no significant pattern showing that zooplankton would lag chlorophyll when chl-a concentrations were high, (p > 0.1). We compare running LL- relationships for chl-a vs. zooplankton at the deep-water region, B (stations WB3 and 4), and at ocean station, WB7 in Figure 3 e and f.
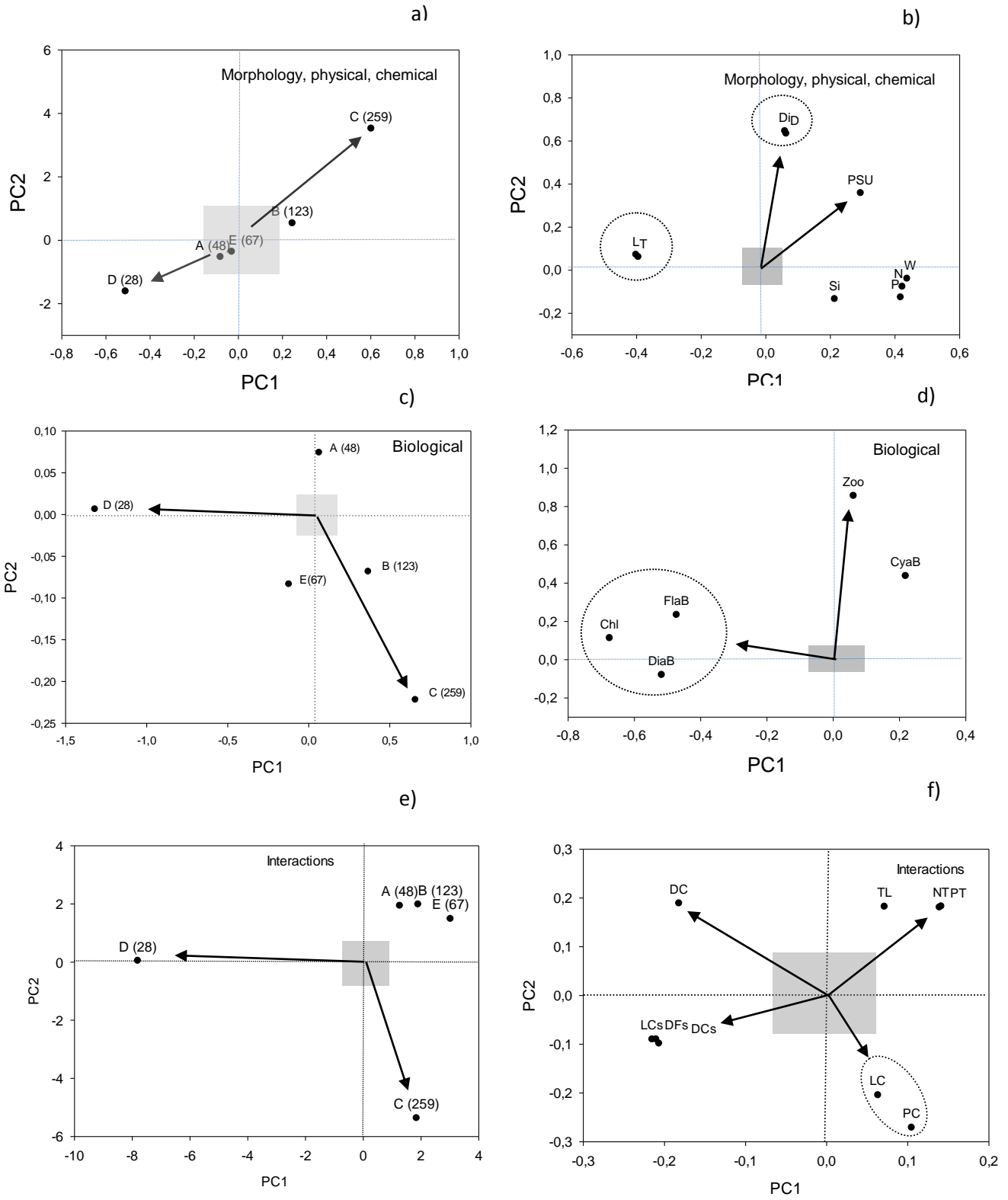
Figure 4

a)

b)

c)

d)

e)

f)

Figure 4 Clustering of groups of stations. Left panels show score plots. Right panels show loading plots. Shaded squares indicate size of 5% confidence interval. a) Clustering based on morphological, physical and chemical characteristics; b) Clustering based on biological characteristics. c) Clustering based on pair wise interaction characteristics. Letters designates groups of stations as in Table 1. Acronyms for upper right panel are L= light, T = temperature, Di = distance from land, D = depth, PSU = salinity, Si = silicon, W = wind, N = nitrogen, P = phosphorus. See also Table 2. Acronyms for middle right panel are Chl = Chlorophyll –a, Zoo = zooplankton, CyaB = biomass of cyanobacteria, FlaB = biomass of flagellates, DiaB = biomass of diatoms. Acronyms for lower right panel consist of two letters. The first letter indicates x-variable. The second letter the y-variable. NT is thus the acronym for the nitrogen - temperature LL-strength; a high value for NT shows that T is a lagging variable to N and a low value shows that T is a leading variable to N. XYs indicates regression β - coefficient for the pair X and Y. L = light, P = phosphorus, C = chlorophyll –a , D = diatoms, F = flagellates. Table 2 a, b shows full list of acronyms. Parentheses enclose average station depths. For clarity, we show only PCA loadings that are significant for the distribution of station groups.

Clustering sites and characterizing regions

Five clusters had sufficient number of observations (< 20) so that we could interpret the results in terms of regional properties. A preliminary observation is that the morphological, physical and chemical characteristics of a region – within a bay area - will affect the biomass and distribution of the species as well as the way in which they interact. Chl- a and its volatility (as standard deviation, SD) decrease with depth and distance from land, (n is here number of stations, not number of regions.)

(5)   Depth (m) = $3.18\times$ Distance (km), $R^2 = 0.86$, $p < 0.003$, $n = 9$

(6)   Chl-a (mgm^{-3}) =$5.77 \times$ Depth (m) $^{0.336}$, $R^2 = 0.71$, $p < 0.05$, $n = 9$

(7)   SD $_{Chl-a}$ (mg.m^{-3}) = $4.71\times$Depth (m) $^{-0,34}$, $R^2 = 0.42$, $p < 0.05$, $n = 9$

Analysis of regions A to E.

For regions A to E we applied PCA to data that characterize the morphological, physical and chemical relationships. Figure 4 a) and b) show the results. To help recognize the regions we have added the average depth of the stations in the legends. The results for the analysis of biological variables are shown in Figure 4c) and d) and the results for interactions are shown in Figure 4 e) and f). The explained variances for

PC$_1$ range from 43 % to 45 % and for PC$_2$ from 20 % to 24 %. The 5% significance levels for the score plots a, c, d, cover about $10 - 20$ % of the range of their axes. Significance levels for the loading plots corresponds approximately to 0.06, 0.1 and 0.6 units on the PC axes on the graphs 4 b, d and f respectively. The overall results show that the sites C (259 m depth) and the site D, (28 m depth and at the mouth of Kennebec River) distinguish themselves from the sites A, B and E. To identify distinguishing characteristics, we exclude relationships that are common for almost all sites and we only include relationships that are significant at 5% level at the sites where they are candidate explanatory variables. We refer to the PCA plots, but the actual data for the PCA matrix can be read off in Table 2.

Regions A, B and E consist of the observation sites that are at medium depths (48 m to 123 m). They are along the coast or on an underwater N-S ridge about 45 km out in the sea. Since the position of these sites in the PCA plot is fairly close to the origin, they will have average values of all variables. We therefore examine the regions C (ocean water) and D (shallow water and close to river mouth) more thoroughly.

Ocean water region, C. The site is represented by one station (WB7) at deep water (259 m). The morphological, physical and chemical characteristics shows that region C is characterized – of course - by large depths, but

also by relatively high values of salinity, PSU (Fig 4a, b). At this site there is a low density of zooplankton, (Fig 4 c, d, arrows pointing in approximately opposite directions from the origin.) Chl-a is a lagging variable to light and PO_4 (no 15,7; L → C; PO_4 → C, LL_{LC} = 0.28, LL_{PC} = 0.54). Cyanobacteria is a leading variable to diatoms, (no 20, 21, Cya → D, $LL_{DC(ya)}$ = -0.82). Zooplankton is not related to chl-a on the average, (no 13, C → Z, LL_{CZ} = 0.0), but there are periods where zooplankton lags chl-a, Fig 4 f.

River mouth region, D, 28 m. The region is represented by one station, CT4, at the mouth of Kennebec River; it is at the other extreme of the deep-water site C, showing relatively high values of light and temperature, Figure 4a, and b. At this station chl-a is high and diatoms and flagellates are the most abundant phytoplankton species groups. Diatoms and flagellates, as well as diatoms and cyanobacteria, are associated (no 19, R = 0.54, no 20, R = 0.33 respectively). There is also a close association between light and chl-a. (no 15, R = 0.37). At the river mouth region light is a strong leading variable to temperature (no 14, LL_{TL} = - 0.54) and temperature is a strong leading variable to nutrients (no 2, LL_{NT} = -0.46; no 5, LL_{PT} = -0.46.)

Discussion
Firstly, we discuss the relationship between chemical and biological variables that are significant across all regions. Secondly, we discuss the ecological differences we found among regions that were in the center of the basin versus those that were close to the mouth of the Kennebec River or far from the coast. We find that clusters of stations based on morphological and pysico-chemical attributes seem to carry on to unique biological characteristics as well as how variables interact.

Generic relationships across all sites
In a seasonal environment species successsion are related both to optimal growth conditions for a particular species with respect to physical and chemical variables, water movements and how species interact. The three set of forces will act together and produce time series that may be complex (Seip and Pleym 2000). In this study we try to capture i) co-movements and counter-cyclic movements by regression analysis and ii) possible causal relationships by a leading - lagging, LL- strength measure.

We obtain two contrasting results. On one side we identify 3 leading- lagging relationships that are part of ecological "common knowledge", on the other side we do not detect important LL-relationships that follow from general ecosystem theory and their mathematical formulations, e.g., nutrients comes before its grazers, the prey before its predator, (Lotka 1924).

Physical and chemical variables. Temperature is a leading variable to nutrients, T → NO_x, T → PO_4, (LL_{NT} =- 0.24; LL_{PT} = - 0.25) and temperature and nutrients are inversely related T = - NO_x; T = - PO_4 (no 2, 5; R< - 0.58). This result suggest that temperature is a contributing causal factor for the decrease in nutrients with increasing temperature, supporting findings that increasing stratification may limit nutrient supply from deeper waters (Boyce et al. 2010). However, increasing temperature also leads to higher algal growth rate (temperatures in the range $10°C$- $20°C$) and thus enhanced depletion of nutrients. Hu et al also identified a similar clock-wise cycling to the one we found. (2008 their Fig 12 a, NO on x- axis) for Georges Bank suggesting that temperature in general is a leading variable to nutrients. We found a significant positive relationships between wind and nutrients (No 3, 6, R > 0.66), but no LL-relationships. Temperature and wind is negatively correlated (WT = - U3; R < - 0.67). We believe our result support the "general" knowledge that wind acts in the opposite direction to temperature with respect to stratification, e.g., Chiswell (2011). Given that increasing temperature limits supply of new nutrients from deeper water, we would anticipate that increases in chl-a would lead to a lack of nutrients. This is also what we generally observe for NO_x (No 4; LL_{NC} = - 0.24), but not significant for the ocean region and the near river region.

Biological variables. In contrast to our first hypothesis, there is a lack of consistent LL-

relationships between pairs of variables that include biological variables. A similar result was observed by Mozetic et al. (2012) who concluded that they failed to find tight couplings in time and magnitude between nutrients and chl-a. However, our result may provide circumstantial evidence that support conjectures about species succession and phytoplankton zooplankton interaction. Firstly, we find that cyanobacteria precedes diatoms, Cya → D, consistent with the observation that experimental blooms began with small prokaryotic organisms and then become dominated by diatoms, Behrenfeld and Boss (2014) on experimental blooms. Secondly, but also circumstantial, is the changing LL- relationship observed for chla- and zooplankton in Figure 4 f for the ocean region. In the years, 2005, 2007 (zooplankton data lacking for 2006) the year ends with zooplankton lagging chl-a during a rather short cyclic sequence, consistent with observations by Behrenfeld and Boss (2014) on experimental blooms that the blooms were stopped by rapid grazing pressure. (In Figure 4 e large values on the y-axis ($\approx \pm$ 3), show large angles for rotating phase plot trajectories, corresponding to short cycle times, Eq (4)). Behrenfeld and Boss (2014) add the requirement of a deepening of the mixed layer depth, probably corresponding to a rebalance between vertical mixing of nutrient across the stratified layer and the dilution effects. The lack of prevalent LL- relationships across regions can either be because there are no such relationships (in spite of theoretical predictions that they might exist, c.f. the introduction), because LL- relationships change with region, or change during successional stages, making the time series too sparse to obtain significant results.

Differences between regions in coastal waters

In contrast to the recent studies on region classification, e.g., Valesini et al. (2010, spatial data) we include biological variables, and we examine interactions between variables. To our knowledge, we use PCA in a novel way when we examine the significance of distances between clusters in PCA plots.

We demonstrate differences in ecosystem characteristics among regions within the semi closed system of the Gulf of Maine. The region close to the mouth of a relatively large river appeared to be significantly different from other regions and different from regions further out. The sites at shallow water depths had higher chl-a concentrations than regions further out, in agreement with findings by Ji et al. (2007 Fig. 4), and also a larger volatility in chl-a in spite of similar volatilities in temperature and nutrients. A high volatility in biomass is predicted to signal dynamic chaos (Scheffer et al. 2001) and support the finding that interaction patterns at shallow waters are qualitatively different from those at deep waters e.g., as in Malin et al. (2005) and Hu et al.(2008).

River mouth region. Areas influenced by river flow tend to be more affected by alternating upwelling events and calm periods (Brooks 2009). During calm periods, stratification can occur rapidly because of higher temperatures or lower salinity levels. Temperature and salinity are negatively correlated at all sites except at the river mouth site D.

Ocean region. Our ocean region, C, has about double the chl-a concentration of other ocean regions reported in the literature, (0.87 mgm^{-3} vs. 0.4 mgm^{-3} reported by Behrenfeld (2010). It is interesting that Boyce et al.(2011) exempt observations from inshore areas to calculate global phytoplankton decline over the past century.

Our *first* hypothesis was only partially supported; we found significant, and potentially causative, leading – lagging, LL- relationships for pairs of physical and chemical variables across regions, but not for LL- relationships that include biological variables. However, some findings were consistent with recent theories for species succession, but data were too sparse to establish that the findings were significant. Our *second* hypothesis were supported, regions in the Gulf of Maine show distinguishable traits and morphological, physical and chemical characteristics of a site appears to translate into biological characteristics and into the way paired variables interact with each other. We believe that our result support the development of

predictive regional simulation models for plankton blooms, e.g., as in Wong et al. (2007), Ji et al. (2013), Chiswell (2011), and Behrenfeld and Boss (2014) for ecosystem responses to climate changes. However, to validate the models, sampling frequency should probably be higher.

Conclusion

For the Gulf of Maine, we found probable causative leading – lagging, LL- relationships between physical and chemical variables that were consistent across regions, but LL-relations that include biological variables varied among regions, or were not significant. We found that the physical and chemical differences among sites within the Gulf of Maine translate into differences in how the plankton ecosystem functions and we found circumstantial evidence for differences among regions suggesting that more frequent sampling could give significant results.

We do not address issues of trophic mismatch (Edwards and Richardson 2004; Lewandowska and Sommer 2010; Head et al. 2013) that may affect production at higher trophic levels. However, the LL- strength method gives modelers a new tool for validating simulations results with observations.

References

Anderson, D. M., B. A. Keafer, J. L. Kleindinst, D. J. McGillicuddy, J. L. Martin, K. Norton, C. H. Pilskaln, J. L. Smith, C. R. Sherwood and B. Butman. 2014. Alexandrium fundyense cysts in the Gulf of Maine: Long-term time series of abundance and distribution, and linkages to past and future blooms. Deep-Sea Research Part Ii-Topical Studies in Oceanography 103: 6-26.

Banse, K. 2013. Reflections About Chance in My Career, and on the Top-Down Regulated World. Annual Review of Marine Science, Vol 5 5: 1-19.

Behrenfeld, M. J. 2010. Abandoning Sverdrup's Critical Depth Hypothesis on phytoplankton blooms. Ecology 91(4): 977-989.

Behrenfeld, M. J. and E. S. Boss. 2014. Resurrecting the ecological underpinnings of ocean plankton blooms. Annu. Rev. Mar. Sci. 6: 167-194.

Behrenfeld, M. J., S. C. Doney, I. Lima, E. S. Boss and D. A. Siegel. 2013. Annual cycles of ecological disturbance and recovery underlying the subarctic Atlantic spring plankton bloom. Global Biogeochemical Cycles 27(2): 526-540.

Boyce, D. G., M. R. Lewis and B. Worm. 2010. Global phytoplankton decline over the past century. Nature 466(7306): 591-596.

Boyce, D. G., M. R. Lewis and B. Worm. 2011. Is there a decline in marine phytoplankton? Reply. Nature 472(7342): E8-E9.

Brooks, D. A. 2009. Circulation and dispersion in a cancellate coast: The rivers, bays and estuaries of central Maine. Estuarine Coastal and Shelf Science 83(3): 313-325.

Chiswell, S. M. 2011. Annual cycles and spring blooms in phytoplankton: don't abandon Sverdrup completely. Marine Ecology Progress Series 443: 39-50.

Doney, S. C., I. Lima, J. K. Moore, K. Lindsay, M. J. Behrenfeld, T. K. Westberry, N. Mahowald, D. M. Glover and T. Takahashi. 2009. Skill metrics for confronting global upper ocean ecosystem-biogeochemistry models against field and remote sensing data. Journal of Marine Systems 76(1–2): 95-112.

Edwards, M. and A. J. Richardson. 2004. Impact of climate change on marine pelagic phenology and trophic mismatch. Nature 430(7002): 881-884.

Elliott, M. and A. K. Whitfield. 2011. Challenging paradigms in estuarine ecology and management. Estuarine Coastal and Shelf Science 94: 306-314.

Franks, P. J. S. and D. M. Anderson. 1992. Alongshore Transport of a Toxic Phytoplankton Bloom in a Buoyancy Current - Alexandrium-Tamarense in the Gulf of Maine. Marine Biology 112(1): 153-164.

Fussmann, G. F., S. P. Ellner, N. G. Hairston, L. E. Jones, K. W. Shertzer and T. Yoshida. 2005. Ecological and evolutionary dynamics of experimental plankton communities. Advances in Ecological Research 37: 221-243.

Fussmann, G. F., G. Weithoff and T. Yoshida. 2005. A direct, experimental test of resource vs. consumer dependence. Ecology 86(11): 2924-2930.

GoMOOS. 2010. Gulf of Maine Ocean Observing System. Portland, Maine ME 04101, GoMOOS.

Head, E. J. H., W. Melle, P. Pepin, E. Bagoien and C. Broms. 2013. On the ecology of Calanus finmarchicus in the Subarctic North Atlantic: A comparison of population dynamics and environmental conditions in areas of the Labrador Sea-Labrador/Newfoundland Shelf and Norwegian Sea Atlantic and Coastal Waters. Progress in Oceanography 114: 46-63.

Holmengen, N. and K. L. Seip. 2009. Cycle lengths and phase portrait characteristics as probes for predator–prey interactions: comparing simulations and observed data Can. J. Zool. 87(1): 20-30.

Hsieh, C. H., S. M. Glaser, A. J. Lucas and G. Sugihara. 2005. Distinguishing random environmental fluctuations from ecological catastrophes for the North Pacific Ocean. Nature 435(7040): 336-340.

Hu, S., D. W. Townsend, C. Chen, G. Cowles, R. C. Beardsley, R. Ji and R. W. Houghton. 2008. Tidal pumping and nutrient fluxes on Georges Bank: A process-oriented modeling study. Journal of Marine Systems 74(1-2): 528-544.

Huntsman, S. A. and R. T. Barber. 1977. Primary Production Off Northwest Africa - Relationship to Wind and Nutrient Conditions. Deep-Sea Research 24(1): 25-33.

Janzen, C. D., J. H. Churchill and N. R. Pettigrew. 2005. Observations of exchange between eastern Casco Bay and the western Gulf of Maine. Deep-Sea Research Part Ii-Topical Studies in Oceanography 52(19-21): 2411-2429.

Ji, R. B., C. S. Davis, C. S. Chen, D. W. Townsend, D. G. Mountain and R. C. Beardsley. 2007. Influence of ocean freshening on shelf phytoplankton dynamics. Geophysical Research Letters 34(24).

Ji, R. B., C. Stegert and C. S. Davis. 2013. Sensitivity of copepod populations to bottom-up and top-down forcing: a modeling study in the Gulf of Maine region. Journal of Plankton Research 35(1): 66-79.

Kahru, M., V. Brotas, M. Manzano-Sarabia and B. G. Mitchell. 2011. Are phytoplankton blooms occurring earlier in the Arctic? Global Change Biology 17(4): 1733-1739.

Lewandowska, A. and U. Sommer. 2010. Climate change and the spring bloom: a mesocosm study on the influence of light and temperature on phytoplankton and mesozooplankton. Marine Ecology-Progress Series 405: 101-111.

Llope, M., G. M. Daskalov, T. A. Rouyer, V. Mihneva, K. S. Chan, A. N. Grishin and N. C. Stenseth. 2011. Overfishing of top predators eroded the resilience of the Black Sea system regardless of the climate and anthropogenic conditions. Global Change Biology 17(3): 1251-1265.

Lotka, A. J. 1924. Elements of physical biology. New York, Williams and Wilkins.

Mackey, M. D., D. J. Mackey, H. W. Higgins and S. W. Wright. 1996. CHEMTAX - a program for estimating class abundances from chemical markers: application to HPLC measurements of phytoplankton. Marine Ecology-Progress Series **144**: 265-283.

Mallin, M. A., L. B. Cahoon and M. J. Durako. 2005. Contrasting food-web support bases for adjoining river-influenced and non-river influenced continental shelf ecosystems. Estuarine Coastal and Shelf Science 62(1-2): 55-62.

Marshall, C. T. and R. H. Peters. 1989. General patterns in the seasonal development of chlorophyll a for temperate lakes. Limnology and Oceanography 34: 856-867.

Merino, J. C. 2003. Lissajous figures and Chebyshev polynomials. The College Mathematics journal 34(2): 122-127.

Moore, T., S. 2008. Dynamics of phytoplankton community compsition in the western Gulf of Maine Doctor of philosophy, University of New Hampshire.

Mozetic, P., J. France, T. Kogovsek, I. Talaber and A. Malej. 2012. Plankton trends and community changes in a coastal sea (northern Adriatic): Bottom-up vs. top-down control in relation to environmental drivers. Estuarine Coastal and Shelf Science 115: 138-148.

NERACOOS.2013). "Northeastern regional assocciation of coastal and ocean observing systems." Retrieved December 12, 2013, 2013.

Salisbury, J., M. Green, C. Hunt and J. Campbell. 2008. Coastal acidification by rivers: a threat to shellfish. EOS transaction, AM Geop. Union 89(50): 513.

Scheffer, M., D. Straile, E. H. van Nes and H. Hosper. 2001. Climatic warming causes regime shifts in lake food webs. Limnology and Oceanography 46(7): 1780-1783.

Seip, K. L. and R. McNown. 2007. The timing and accuracy of leading and lagging business cycle indicators: a new approach. International journal of forecasting 22: 277-287.

Seip, K. L. and H. Pleym. 2000. Competition and predation in a seasonal world. Verh. Internat. Verein. Limnol. 27: 823-827.

Seip, K. L. and C. S. Reynolds. 1995. Phytoplankton functional attributes along trophic gradient and season. Limnology and Oceanography, 40: 589-597.

Sommer, U., Z. M. Gliwicz, W. Lampert and A. Duncan. 1986. The PEG-model of seasonal succession of planktonic events in fresh waters. Arch. Hydrobiol. 106: 433-471.

Song, H. J., R. B. Ji, C. Stock, K. Kearney and Z. L. Wang. 2011. Interannual variability in phytoplankton blooms and plankton productivity over the Nova Scotian Shelf and in the Gulf of Maine. Marine Ecology Progress Series 426: 105-U133.

Song, H. J., R. B. Ji, C. Stock and Z. L. Wang. 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.

Sverderup, H. U. 1953. On conditions for the vernal blooming of phytoplankton. J. Cons. Explor. Mer. 18: 287-295.

Tomte, O. T., K. L. Seip and N. Christophersen. 1998. Evidence that loss in predictability increases with weakening of (metabolic) links to physical forcing functions in aquatic ecosystems. Oikos 82(2): 325-332.

Valesini, F. J., M. Hourston, M. D. Wildsmith, N. J. Coen and I. C. Potter. 2010. New quantitative approaches for classifying and predicting local-scale habitats in estuaries. Estuarine Coastal and Shelf Science 86(4): 645-664.

Wikipedia. 2013Lissajous curve DOI: http://en.wikipedia.org/wiki/Lissajous_curve.

Wilkerson, F. P., A. M. Lassiter, R. C. Dugdale, A. Marchi and V. E. Hogue. 2006. The phytoplankton bloom response to wind events and upwelled nutrients during the CoOP WEST study. Deep-Sea Research Part II -Topical Studies in Oceanography 53(25-26): 3023-3048.

Wong, K. T. M., J. H. W. Lee and I. J. Hodgkiss. 2007. A simple model for forecast of coastal algal blooms. Estuarine Coastal and Shelf Science 74(1-2): 175-196.

Supplementary material 1

Method details and suggested Excel formulation for Leading-lagging strength calculations.

Figure 3a in main text. The patterns in Figure 3a correspond to two sine series: sin (2πt) and sin (2πt + τ). The text below gives the values of τ that will give the patterns found in quadrant I, II, III, and IV in the figure. The pattern in quadrant I emerges if the second sine is shifted $\tau \in$ <-2π to -3π/2> relative to the first; the pattern in quadrant II emerge if shift is $\tau \in$<-3π/2, –π>; the pattern in quadrant III emerge if the shift is $\tau \in$<–π , –π/2> and the pattern in quadrant IV emerge if the shift is $\tau \in$<–π/2 , 0>. Two perfect sines that are shifted λ/4 to each other will have a regression coefficient R = 0.
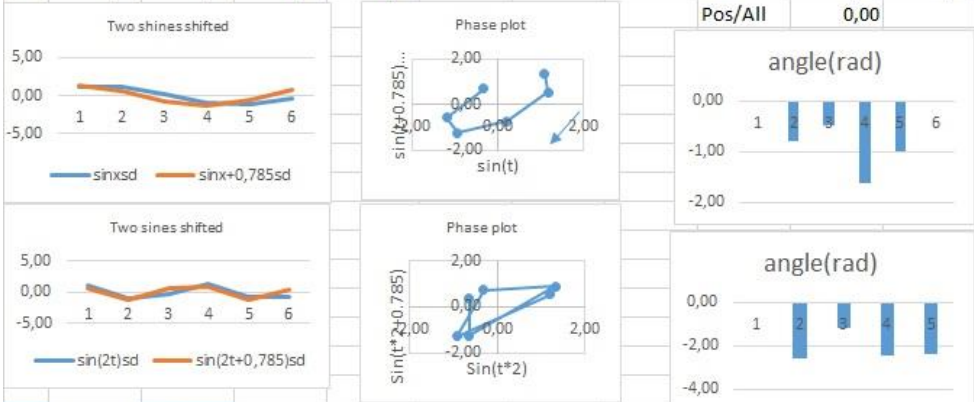
Excel formulation for calculating rotational direction, V, corresponding to the formulae in Eq. (2)

With two series pasted into cells (A1, A2, A3) and (B1, B2, B3) in an Excel spread sheet, the angle, V, is calculated by pasting the following Excel expression into the cell C2: =SIGN((A2-A1)*(B3-B2)-(B2-B1)*(A3-A2))*ACOS(((A2-A1)*(A3-A2) + (B2-B1)*(B3-B2))/(SQRT((A2-A1)^2+(B2-B1)^2)*SQRT((A3-A2)^2+(B3-B2)^2))). The angle, V, corresponds to the expression "angle(rad)" in row 3 and column 6 in the spreadsheet below.

Excel calculations. The upper part shows an example with sampled sine series: sin(t) and sin (t+0.785). The series are centered and normalized to unit standard deviation. The "angle(rad)" column uses the excel expression for V above. The lower part shows part of the uncertainty analysis using the random generator RAND() in Excel.
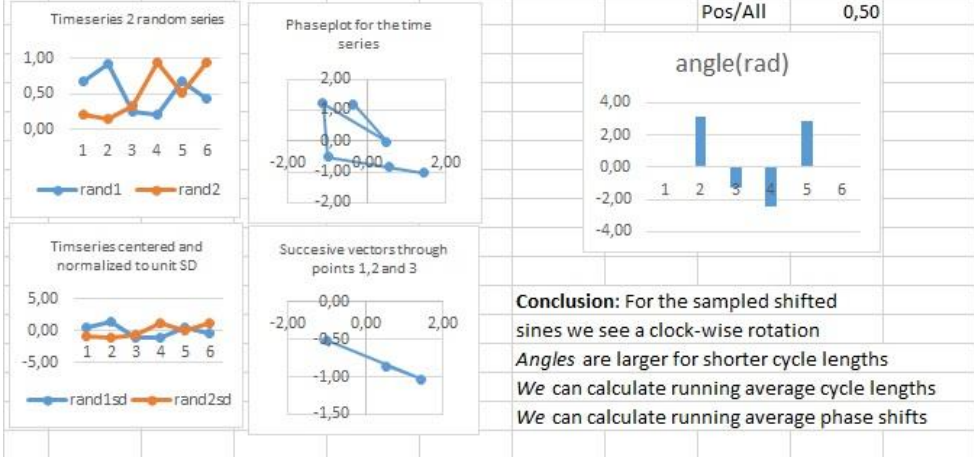
1. Sine example; sin (t), sin(t+0.785): and only figures Sin(2t), sin(2t+0.785)

t	Sine series		Series cen. & norma		Calculation of angles				#Rot	
	Sint	Sin(t+0,78	sinxsd	sinx+0,785	angle(rad	Anglew(d	pos angle	neg angle	#Pos rot	#NegRot
1	0,84	0,98	1,09	1,35						
2	0,91	0,35	1,17	0,52	-0,79	-45,17	0,00	-45,17	0,00	1,00
3	0,14	-0,60	0,20	-0,73	-0,47	-27,15	0,00	-27,15	0,00	1,00
4	-0,76	-1,00	-0,94	-1,26	-1,62	-92,66	0,00	-92,66	0,00	1,00
5	-0,96	-0,48	-1,19	-0,57	-0,99	-56,61	0,00	-56,61	0,00	1,00
6	-0,28	0,48	-0,33	0,69						
aver	-0,02	-0,04	0,00	0,00						
Stde	0,79	0,76	1,00	1,00				sum	0,00	4,00
								Pos/All	0,00	

Two shines shifted

Phase plot

angle(rad)

Two sines shifted

Phase plot

angle(rad)

2. Random example

	Random numbers		Series cen. & norma		Calculation of angles-180o,+180o				#Rot	
	rand1	rand2	rand1sd	rand2sd	angle(rad	Anglew(d	pos angle	neg angle	#Pos rot	#NegRot
1	0,68	0,20	0,54	-0,86						
2	0,93	0,14	1,44	-1,03	3,12	179,04	179,04	0,00	1,00	0,00
3	0,25	0,33	-1,00	-0,52	-1,25	-71,75	0,00	-71,75	0,00	1,00
4	0,22	0,95	-1,12	1,22	-2,40	-137,53	0,00	-137,53	0,00	1,00
5	0,67	0,51	0,49	-0,01	2,84	162,78	162,78	0,00	1,00	0,00
6	0,43	0,94	-0,35	1,20						
aver	0,53	0,51	0,00	0,00						
Stde	0,28	0,36	1,00	1,00				sum	2,00	2,00
								Pos/All	0,50	

Timeseries 2 random series

Phaseplot for the time series

angle(rad)

Timseries centered and normalized to unit SD

Succesive vectors through points 1,2 and 3

Conclusion: For the sampled shifted sines we see a clock-wise rotation
Angles are larger for shorter cycle lengths
We can calculate running average cycle lengths
We can calculate running average phase shifts

Supplementary material 2. The Lissajous curves
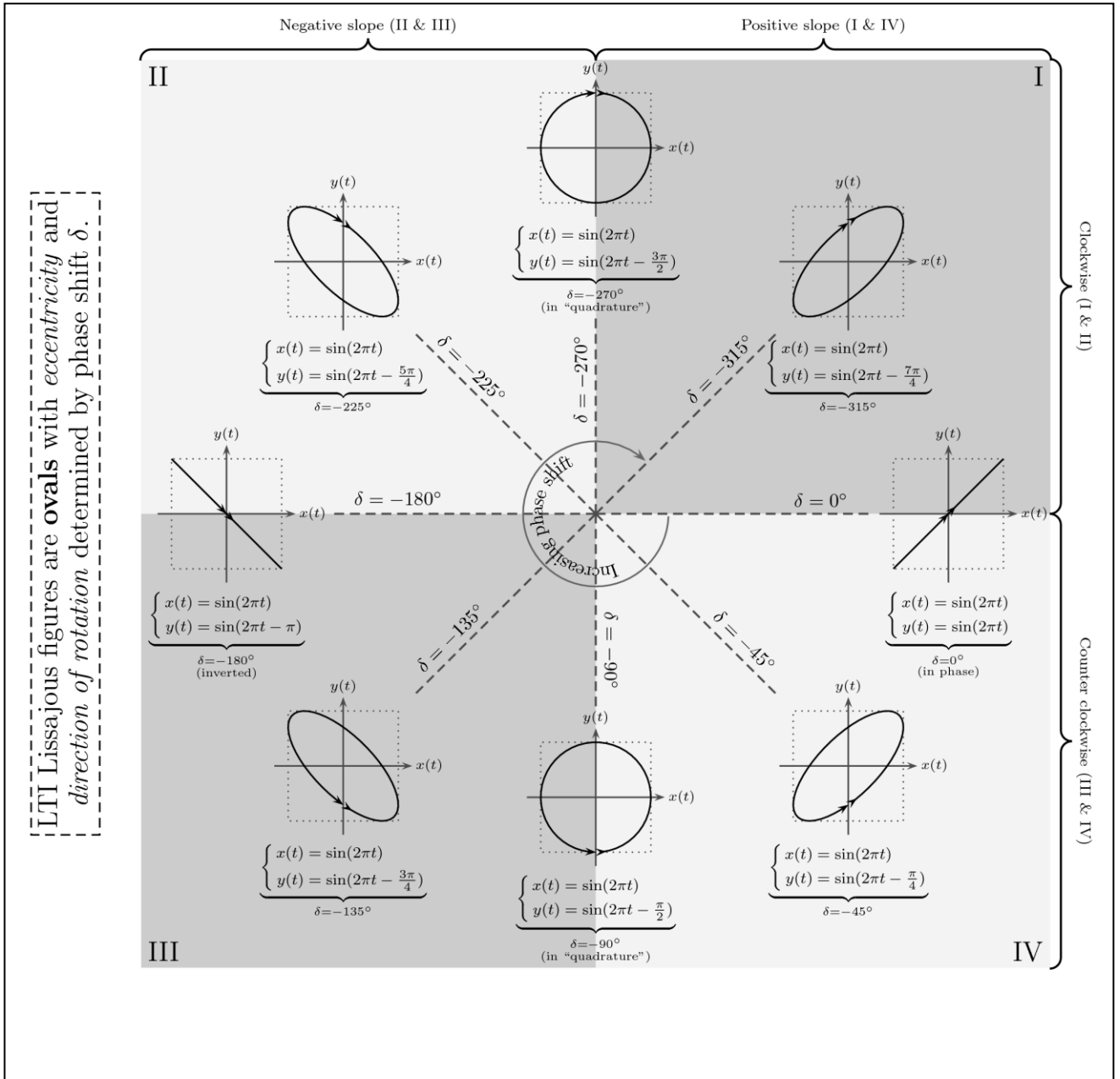
From Wikipedia: "Lissajous curves"