Using empirical dynamic modeling to assess relationships between atmospheric trace gases and eukaryotic phytoplankton populations in coastal Southern California

Jesse M. Wilson[1], Melissa L. Carter[1], Jens Mühle[1], Jeff S. Bowman[1,2]

[1]Scripps Institution of Oceanography, UCSD

[2]Center for Microbiome Innovation, UCSD

Contact: jmwilson@ucsd.edu

Keywords: trace gas, phytoplankton, time-series, causality testing

Abstract

Many different atmospheric trace gases have been directly and indirectly linked to biological sources and sinks. Here we assess how atmospheric mixing ratios of a range of halocarbons (CH_3Br, CH_2Br_2, $CHBr_3$, CH_3Cl, $CHCl_3$, and CH_3I) and COS are causally connected to naturally occurring marine eukaryotic phytoplankton in coastal Southern California. We use a self-organizing map to characterize the abiotic environment and empirical dynamic modeling with convergent cross mapping to identify causal interactions between multiple *in situ* 8-year time-series, sampled at the Ellen Browning Scripps Pier at Scripps Institution of Oceanography. Our work supports previous findings that halocarbon production is found in a variety of marine phytoplankton taxa and suggests that local phytoplankton may have the ability to affect changes in the mixing ratios of halocarbons in nearshore environments. There were notable links between changes in CH_3I and several different diatom taxa and between changes in $CHCl_3$ and a group of phytoplankton during specific ecosystem states. Our results suggest that both seasonal and non-seasonal shifts in eukaryotic phytoplankton structure contribute to small fluctuations in atmospheric halocarbon mixing ratios that exhibit strong seasonality and may occasionally play a larger role in atmospheric mixing ratios of halocarbons that display reduced seasonality.

Introduction

Phytoplankton can affect the atmospheric mixing ratios of a range of trace gases directly through production and indirectly by contributing biomass that is broken down during senescence and death. It is known that phytoplankton can produce halogenated volatile organic compounds (VOCs), a class of trace gases that are of interest due to their effects on tropospheric chemistry and their ability to destroy ozone when transported to the stratosphere (WMO, 2018). Specifically, a variety of marine phytoplankton and macrophytes have been shown to emit methyl bromide (CH_3Br), dibromomethane (CH_2Br_2), bromoform ($CHBr_3$), methyl chloride (CH_3Cl), chloroform ($CHCl_3$), and methyl iodide (CH_3I) in the lab (Manley and Dastoor, 1988; Scarratt and Moore, 1996, 1998; Moore, 2003; Colomb *et al.*, 2008; Brownell *et al.*, 2010; Paul and Pohnert, 2011; Johnson *et al.*, 2015; Lim *et al.*, 2018) and at offshore sites (Baker *et al.*, 2000; Moore and Tokarczyk, 1993; Moore, 2003). While *in situ* studies relating specific nearshore coastal marine phytoplankton to atmospheric mixing ratios of these trace gases are notably lacking, previous work in the open ocean has linked CH_3Br, CH_2Br_2, $CHBr_3$, CH_3Cl, $CHCl_3$, and CH_3I to chlorophyll *a* concentration and phytoplankton type (Arnold *et al.*, 2010).

The inorganic compound carbonyl Sulfide (COS) is another trace gas that is linked to phytoplankton. COS is the most abundant sulfur-containing gas in the atmosphere and can indirectly affect climate by contributing to the stratospheric sulfate pool through interactions with UV light and free radicals (Carpenter and Reimann *et al.*, 2014; Whelan *et al.*, 2018). COS from the marine environment is produced by a combination of sources, including salt marsh vegetation (Commane *et al.*, 2013; Whelan *et al.*, 2013) and the photochemical oxidation of dissolved organic sulfur compounds (Andreae and Ferek, 1992), mainly dimethylsulfide and the organic building block carbon disulfide (Andreae and Crutzen, 1997; Kettle *et al.*, 2002). Thus,

47 COS production is indirectly related to phytoplankton activities and it has also been observed to
48 be consumed at low levels by marine algae (Blezinger *et al.*, 2000).

49 The aforementioned studies either took place in laboratory settings or are based on
50 relatively few data points during fairly short periods of fieldwork, limiting the projection of these
51 data to other settings or the wider ecosystem. The lack of long-term monitoring has made it
52 difficult to establish clear links between natural populations of phytoplankton and atmospheric
53 trace gases. Long-standing time-series provide a unique framework in which to analyze and
54 statistically relate many different variables in an open system and simultaneously allow for
55 baseline observations, an understanding of natural cycles (*e.g.*, seasonality, upwelling, *etc.*), and
56 an understanding of anthropogenic change. By studying variables that are not removed from
57 natural ecological interactions or environmental perturbations, conclusions regarding
58 connections between different variables are more accurate.

59 We took advantage of several long-standing time-series collected from the Ellen
60 Browning Scripps Pier (Scripps Pier), located within the Southern California Bight ecosystem, to
61 establish causal links between naturally occurring and ecologically active (changing due to
62 natural abiotic and biotic pressures) eukaryotic phytoplankton populations and atmospheric trace
63 gases over 8 years. Our overarching goal was to assess how atmospheric mixing ratios of
64 halocarbons and COS may be affected by natural populations of ecologically active marine
65 eukaryotic phytoplankton through time. We did not seek to explain all of the variation observed
66 in any of the trace gas mixing ratios, nor in the phytoplankton populations. Rather, we sought to
67 assess whether changes in groups of specific phytoplankton led to time-lagged changes in trace
68 gases at certain times. Many of the time-series displayed strong nonlinear dynamics (meaning
69 that the time-series could be correlated at some points in time and not others) and so we utilized

1) a self-organizing map (SOM, Kohonen, 2001) to segment and characterize the abiotic portion of the ecosystem state (*e.g.* water temperature), and 2) empirical dynamic modeling (EDM) with convergent cross mapping (CCM) to assess causal links between individual phytoplankton and trace gases. This latter method has been used for a wide range of applications including finding drivers of global influenza cases (Deyle *et al.*, 2016) and identifying controls on fish recruitment (Deyle *et al.*, 2018). EDM with CCM has also been used to predict coastal algal blooms in Southern California (McGowan *et al.*, 2017). EDM with CCM is particularly useful to our purposes because it can show causality between nonlinear time-series that may correlate during certain system states (all the ecological and abiotic factors that describe a regime or ecosystem state) but not in others, and can recreate system states without input for all the variables naturally acting on a time-series (Sugihara *et al.*, 2012). Put another way, ecological and environmental interactions may only be present during some time-periods and CCM allowed us to assess causal links between individual phytoplankton taxa and trace gases without data describing every variable present in our system (*e.g.*, terrestrial biomass production of a trace gas, atmospheric transport, *etc.*). However, the link between phytoplankton and trace gases may not be direct (*e.g.*, trace gases could be indirectly linked to phytoplankton community structure through the microbial degradation of senescent or dead phytoplankton). Therefore, we used CCM to show which populations may have a direct *or* indirect link to atmospheric trace gas mixing ratios. Finally, while a SOM does not show causality, it can characterize different ecosystem states for the variables that were used and thus allowed us to assess if there were certain abiotic ecosystem states during which certain phytoplankton and traces gases were causally linked.

We had two broad expectations going into this study. First, because many eukaryotic phytoplankton and trace gases follow seasonal cycles, we expected that annually seasonal

93 phytoplankton would causally link to trace gases that followed corresponding seasonal cycles.
94 Second, we anticipated causal links between stochastic phytoplankton and trace gases with
95 reduced seasonality. Regardless of our expectations, establishing connections between
96 phytoplankton and trace gases (and the ecosystem states that those links exist within) is the first
97 step to assess how direct and/or indirect phytoplankton production of these trace gases are
98 affected by long-term sustained ecosystem shifts or changes in specific seasonal factors.

99 **Methods**

100 All data were collected off the Scripps Pier at Scripps Institution of Oceanography (SIO;
101 32.8663° N, 117.2546° W). SIO is located in La Jolla, CA within the often-studied Southern
102 California Bight ecosystem (Di Lorenzo, 2003; Tai and Palenik, 2009; McGowan *et al.*, 2017).
103 Local conditions are often influenced by the adjacent California Current ecosystem, with
104 offshore upwelling typically occurring in summer months (Bakun, 1973).

105 The Advanced Global Atmospheric Gases Experiment (AGAGE) measures a range of
106 trace gases, including many halocarbons and COS, at multiple sites all around the world.
107 Meanwhile, the Southern California Coastal Ocean Observing System (SCCOOS) measures
108 harmful algal species in southern California, with additional taxa enumerated at SIO as part of
109 the McGowan Plankton and Chlorophyll Program. We leveraged additional environmental
110 variables (water temperature, air temperature, wind speed, wind direction, and atmospheric
111 pressure) that are collected by the National Oceanic and Atmospheric Administration (NOAA)
112 on and near the Scripps Pier in order to assess how variability in the marine environment affects
113 atmospheric mixing ratios of halocarbons and COS in the boundary layer.

The methods and scope of the AGAGE project are covered in Prinn *et al.* (2018). SIO is the central calibration and development laboratory of the global AGAGE network (Mühle *et al.*, 2007). Two "Medusa" cryogenic pre-concentration systems with gas chromatographic separation and mass spectrometric detection are operated at SIO which sample ambient air from the Scripps pier at a height of 13 m above sea surface. We selected trace gases from the available dataset that were potentially produced directly or indirectly by eukaryotic phytoplankton taxa based on existing literature (CH_3Br, CH_2Br_2, $CHBr_3$, CH_3Cl, $CHCl_3$, CH_3I, and COS). Each Medusa sampled ambient air every 2 hours for 20 minutes, with the resulting concentration representing an average over these 20 minutes. We took the mean for each time-point for the two Medusa systems, and then took a daily mean of all time-points for each gas. Each air measurement is bracketed by a working standard measurement, which in turn is compared once a week (n=4 times) with a higher-level standard calibrated at SIO against secondary calibration standards. Working and higher-level standards are comprised of compressed ambient air samples (RIX industries, USA) stored in high-pressure gas cylinders (Essex Industries, USA), for details see Prinn *et al.* (2000, 2018). Mixing ratios are reported as part per trillion (ppt) dry-air mole fractions on the following SIO and NOAA primary calibration scales: CH_3Cl and CH_3Br on SIO-05; CH_3I on NOAA-2004; $CHCl_3$ on SIO-98; CH_2Br_2 and $CHBr_3$ on NOAA-2003; and COS on a provisional transfer of the NOAA scale to SIO. Precisions were estimated from working standard measurements for each day and were typically: ~0.2% for CH_3Cl; ~0.6% for CH_3Br and COS; ~1% for $CHBr_3$; and ~1.5% for CH_3I, $CHCl_3$ and CH_2Br_2.

Select eukaryotic phytoplankton, chlorophyll *a*, and phaeophytin were measured weekly by the Scripps Pier node of SCCOOS (http://sccoos.org/harmful-algal-bloom). Select eukaryotic phytoplankton species were enumerated as part of the McGowan Plankton Chlorophyll Program,

as described in Hatch *et al.* (2013). Briefly, photosynthetic pigments were obtained by filtering seawater across a 0.7 µm glass fiber filter and then extracted by soaking in 10 ml of 90% acetone for 24 hours. A calibrated Turner 10 AU fluorometer was used to determine concentrations. Abundances of diatoms and dinoflagellates were enumerated by settling 10-50 ml of seawater preserved with 4% formaldehyde and then identifying and counting cells through a phase-contrast, inverted light microscope at 200x amplification to the lowest taxonomic group (species or genus) feasible. To maintain consistency in the groups of plankton counted over the 8-year time-series, we combined some species to form genus level groups when consistent species-levels were not feasible over the entire time-series. We limited the number of species and genera used in this analysis to those that were present at least 5% of the time. In total, we assessed 27 different groups of diatoms and 19 different groups of dinoflagellates. These include total counts for diatoms and dinoflagellates, and two subgroups for the genus *Pseudo-nitzschia* based on valve width (with *Pseudo-nitzschia delicatissima* having a width of less than 3 µm and *Pseudo-nitzschia seriata* having a width of greater than 3 µm). Genus- and species-level counts for the dinoflagellate genera *Ceratium* and *Dinophysis* were used to assess the sensitivity of our analysis to different levels of classification. These genera were selected due to the high confidence of identification.

Environmental data (water temperature, air temperature, wind speed, wind direction, and air pressure) were obtained from the NOAA National Data Buoy Center on 21 August 2019 and reflect data collected for Station LJAC1 – 9410230 in La Jolla, CA (https://www.ndbc.noaa.gov/station_history.php?station=ljac1). Air temperature was taken 16.5 m above mean sea level, the anemometer height was 17.5 m above mean sea level, and the barometer height was 20.6 m above mean sea level. All data were averaged for a daily value,

with wind direction averaged using the 'circular' package version 0.4-93 in R (Lund *et al.*, 2017). Season was represented by a sinusoid with an annual wavelength and amplitude from −1 to 1, with the maxima occurring in early July and minima occurring in early January. The weekly change in season and the weekly change in daily averaged water temperature, air temperature, wind speed, and air pressure were incorporated into EDM analysis, while all weekly incremented daily averaged environmental variables (including wind direction) were used to train the SOM (see below). Weekly incremented daily average values were used to match the phytoplankton sampling interval.

Averaging resulted in a time-series with a resolution of one week that ran from 1 January 2011 through 31 December 2018. This resolution allowed for the identification of significant trace gas production by ecologically active phytoplankton while also filtering out much of the short-term diel variation that is observed in typical trace gas mixing ratios. It is important to recognize that our analysis was not designed to explain all variation in the investigated trace gases. Rather, we assessed whether changes in phytoplankton composition were important over time-scales greater than a week (up to a season). Additionally, we did not expect phytoplankton to be a sole contributor at any time, but to potentially explain a portion of the variation in atmospheric trace gas mixing ratios.

Following Bowman *et al.* (2018), a SOM was used to reduce the multidimensional environmental dataset consisting of water temperature, air temperature, wind speed, wind direction, and air pressure to a single categorical variable ('kohonen' package in R version 3.0.8; Wehrens and Kruisselbrink, 2018) termed "environmental mode" (EM). Prior to assignment, the ranges of all environmental data were scaled by dividing each variable by the root-mean-square of its vector using the 'scale' function in 'base' R version 3.5.1 (R Core Team, 2018). During

SOM training we varied the SOM map units between grids sized 5x5 to 8x8. After visual inspection of the SOMs, we settled on a 6x6 toroidal grid based on the distribution of the number of samples (387 dates which represent timepoints when all data parameters were collected) (Fig. S1). K-means clustering was used to segment map units into EMs, with the final value for k selected based on a within-clusters sum of squares scree plot, and experimentally varying k around the perceived optimum before settling on the final segments. The resulting EMs were related to the constituent environmental variables and trace gas mixing ratios using analysis of variance (ANOVA) and Tukey's Honest Significant Difference test (TukeyHSD) ('stats' package in R version 3.5.1; R Core Team, 2018) to assess whether the means for each variable were statistically different for various EMs.

We used the R package 'rEDM' v. 0.7.2 (Ye *et al.*, 2016) to conduct EDM with CCM. This allowed us to relate 1) changes in environmental variables and season to changes in individual trace gas mixing ratios, 2) changes in individual marine phytoplankton to changes in individual trace gas mixing ratios, and 3) changes in season to changes in individual phytoplankton. EDM can show causality between nonlinear time-series that may correlate during certain system states but not in others (Sugihara *et al.*, 2012). Briefly, in EDM nonlinear relationships are established by empirically modeling the state space of a time-series at any given moment using that time-series' own lagged components as embedding dimensions (as substitutions for the external variables that are affecting it). Time-series are identified as nonlinear using the S-map forecasting method (Sugihara *et al.*, 1994). Once attractor reconstruction (time-series modeled in a state space) is achieved for multiple time-series, the attractor associated with one time-series can be used to predict another. Causality is demonstrated when one attractor accurately predicts another (the predicted time-series is

embedded in the predictor time-series and thus the predicted time-series causes the predictor) and the cross-map ability (ρ) between the two time-series should increase as the size of the attractor increases if there is causality (Sugihara *et al.*, 2012). This process is the CCM component of EDM.

All data were scaled prior to analysis and we evaluated autocorrelation with the 'acf' function in the R 'stats' package. Due to autocorrelation among many variables in our time-series, we took the first difference (the difference between sequential time points) of each time-series that went into CCM (*i.e.* the weekly change in season and weekly change in daily averaged water temperature, air temperature, wind speed, air pressure, each trace gas, and each phytoplankton taxon's abundance). This technique has been shown to be helpful for identifying meaningful attractor-based predictions rather than false correlations between highly autocorrelated time-series (Deyle *et al.*, 2018). Put another way, it ensures that we were relating changes in one time-series to changes in another. This furthered our purpose of relating the most ecologically active phytoplankton (those most impacted by growth, mortality, or transport) to changes in trace gases. After taking the first difference we identified a τ of 2 weeks (where τ is the lag used to create the various embedding dimensions for attractor reconstruction) as appropriate because that value was sufficiently low to be ecologically meaningful (allowing the previous time point to relate to the following time point), yet sufficiently large to prevent auto-correlation for trace gas time-series.

Multiple prediction horizons were used to assess possible time-delayed causations. The prediction horizon (tp) was set at 0, −1, −2, and −3 weeks for all CCM relationships related to changes in atmospheric trace gases and at −1 week for relating change in season to changes in phytoplankton taxa (since the variation in season was represented by a sinusoidal wave there was

no reason to test multiple time delays for phytoplankton). When tp is negative, changes in one time-series are being used to predict past changes of another; when tp is 0, concurrent changes are being predicted. We tested a range of prediction horizons to capture the effect of slow-growing phytoplankton on and indirect relationships of phytoplankton with trace gases. The optimal embedding dimension (Table S1) for each time-series was found via simplex projection by allowing the embedding dimension to vary between 1 and 10, then identifying the embedding dimension that maximized ρ. The optimal embedding dimension is the number of dimensions that the reconstructed attractor needs to unfold, and that produces the highest forecast skill. The embedding dimensions varied depending on the features of each time-series (*e.g.*, *Cylindrotheca closterium* was abundant and displayed strong seasonality but also possessed many smaller fluctuations, translating to changes in *C. closterium* requiring 8 embedding dimensions to unfold; Table S1). For simplex projection, τ was set to 2 weeks (see reasoning above) and tp was set to the default of 1 week (1 timepoint in the future). We tested for nonlinearity for each differenced time-series using S-maps (with a τ of 2 and tp of 1, similar to simplex projection) and all time-series tested at this level of sampling displayed nonlinearity except for season, COS, total dinoflagellates, and the dinoflagellate *Akashiwo sanguinea*, thus justifying the use of EDM with CCM to relate the various time-series. Library sizes for differenced time-series related by CCM ranged from 1 to 411 in steps of 10 (the total number of data points for each time-series was 418 before the first difference was taken), with 100 iterations at each library size. We allowed complete overlap between the library (predictor) time-series and forecasted (predicted) time-series, forcing the automatically enabled "leave-one-out-cross-validation" to occur. The optimal embedding dimension for the variable that represented the library time-series (as opposed to the forecasted time-series) was selected when running various CCMs.

CCM can differentiate between causal relationships and synchrony or natural seasonality because it is run in both directions (*i.e.* with each time-series run as both the predictor and predicted time-series). When two variables exhibit synchrony, then the direction of causation is obscured and bidirectional forcing will be observed (Sugihara et al, 2017). Bidirectional forcing is when both variables appear to affect one another, but each may actually be responding to a third variable (Sugihara *et al.*, 2012). Thus, we tested for bidirectional forcing of CCM relationships two ways. First, we used Welch's two-sample t-test to assess whether the mean of the maximum ρ (cross-map skill or the correlation coefficient between predictions and observations) values for each CCM pairing (*i.e.* where the predictor and predicted variables are reversed) were significantly different ($p < 0.05$). Because 100 iterations were run at each library size the n was 100. This test considers both predictor and predicted roles for a given set of variables and we restricted our analysis to ρ values at or greater than 0.1 for when a trace gas represented the library (predictor) variable. When a trace gas was used to predict or forecast another time-series, the signature of the other time-series was embedded within the time-series of the trace gas, thus the trace gas was significantly affected by the response variable. Second, we assessed if the largest library size had a significantly larger ρ than the smallest library size for each CCM test. Because ρ should increase as the size of the attractor increases (Sugihara *et al.*, 2012), assessing how frequently ρ at the largest library size is larger than ρ at the smallest library size allows us to assess significance ($p < 0.20$) using Eq. (1):

$$(1)\ \mathrm{p} = 1 - \Sigma\frac{\rho_{small} < \rho_{large}}{n}$$

For this significance test we present results for CCMs in both directions (with each trace gas as both the predictor and predicted variable) when the maximum ρ was at or above 0.1. This

allowed us to determine if ρ increased with increasing library size and if there was bidirectional forcing.

The ability of changes in multiple phytoplankton taxa to forecast changes in each trace gas was tested against the forecast skill of the trace gas's own lagged components using the 'block_lnlp' command in the 'rEDM' package. This test allows generalized forecasting with mixed embeddings or multivariate attractor reconstruction, and served as a check on whether the phytoplankton that CCM found to be important actually increased our ability to accurately forecast changes in atmospheric trace gases. Used this way it can confirm the causal relationships between phytoplankton and trace gases (as we were able to model all other necessary external variables through a trace gas's own lagged components). We employed the L2-norm type and simplex projection, and set tp 1 week in the future. We did not include environmental data and only used the lowest significant lag when the change in a taxon was significant at multiple lags. Additionally, we used the current trace gas time-series to predict the change in that trace gas one time-step in the future. When the optimal embedding dimension of a trace gas called for more variables than significant phytoplankton, we also used a trace gas's own lagged times-series (at the lowest possible lags) to reach the optimal embedding dimension. For CH_3Br and CH_3I we went above the optimal embedding dimension in order to utilize all the phytoplankton that were found to be significant. For CH_3I, which had an optimal embedding dimension of one, we also tried each of the significant phytoplankton taxa individually. We did not include total diatom counts for CH_3Br because the significant diatoms were already accounted for. Average phaeophytin concentration was an important variable for $CHBr_3$ and $CHCl_3$ and so we completed generalized forecasting both with and without average phaeophytin.

All data management and analyses took place in the R Studio statistical environment (R Studio Team, 2018).

Results

Variation in trace gases, environmental variables, and eukaryotic phytoplankton

All of the observed trace gases displayed significant variation over the time-series and were sorted into three groups based on seasonality: 1) visible annual seasonality, 2) visible annual seasonality with an aseasonal feature, and 3) multiple-seasonality (with reliable peaks in multiple seasons) to no seasonality. CH_3Cl, CH_3I, and COS each displayed strong seasonality with maxima in one season: winter/spring for CH_3Cl and summer for CH_3I and COS (Fig. 1D, F, G). CH_2Br_2 and $CHBr_3$ also displayed seasonality, typically peaking in late summer, but showed extremely high mixing ratios in 2014 that dwarfed the seasonal signal (Fig. 1B, C). Finally, $CHCl_3$ displayed multi-maxima-per-year seasonality while CH_3Br displayed no visible seasonality, with local maxima potentially occurring in all seasons throughout the time-series (Fig. 1A, E).

There were also clear seasonal signals in the environmental data. Water temperature ranged from 12.70 ºC to 24.32 ºC and peaked in the summer, but possessed interannual variation in the yearly maxima and minima with a small increase in mean temperature over the course of the time-series (Fig. 2A). 2015 had the highest recorded temperature for this time-series, while 2013 had the lowest recorded temperature. Air temperature closely tracked water temperature ($r^2 = 0.59$, $p < 0.01$; Fig. 2B) and ranged from 8.83 ºC to 27.58 ºC, also peaking in the summer of 2015. Wind speed tended to peak in the spring/summer, with the lowest values observed in late fall and early winter, but showed wide variation in all seasons (Fig. 2C). Wind direction was

318 highly variable but typically originated from the west (ocean) in the spring and summer (on-
319 shore winds) and from the east (land) during the winter (off-shore winds; Fig. 2D). Air pressure
320 had strong seasonality and peaked in the winter (Fig. 2E).

321 Peak abundance of both dinoflagellates and diatoms tended to occur in summer, though
322 the timing of the absolute peak varied from year to year (Fig. 3A, B). Some taxa were found
323 more frequently in non-summer months but tended to have lower maximum abundances than
324 taxa that peaked in the summer. Additionally, some taxa, such as *C. closterium*, were frequently
325 found in the time-series while others, such as *Dinophysis acuminata*, were found much less
326 frequently (Fig. 3C, D). Further variation in seasonality and frequency is demonstrated by the
327 phytoplankton time-series in the Supplementary Material (Fig. S3-S8).

328 *Relationships between environmental variables and atmospheric trace gases*

329 Characterization of the above suite of environmental variables using a SOM and k-means
330 clustering identified seven EMs (Fig. 4). Some of these related well to atmospheric trace gases
331 (Fig. 5). CH_3Cl and $CHCl_3$ each tended to peak during EM6 (Fig. 5D, E), which possessed a
332 wide range of temperatures and air pressures (Fig. 4B, C, F), but had low wind speeds from the
333 southeast (Fig. 4D, E) that did not occur in the peak of summer (Fig. 4A). CH_3I was highest
334 during EM2 and EM3 (Fig. 5F), which was characterized by warmer temperatures and low wind
335 speeds (Fig. 4B, C, D) and was often observed during spring, summer, and fall (Fig. 4A). COS
336 was lowest during EM6 and EM7 (Fig. 5G), which was characterized by low easterly wind
337 speeds (Fig. 4D, E) and thus COS was highest during spring, summer, and fall (Fig. 4A).
338 However, CH_2Br_2 and $CHBr_3$ did not show any significant variation across the different EMs
339 (Fig. 5B, C) and CH_3Br showed very little variation across the different EMs (Fig. 5A).

Variables affecting trace gases with visible annual seasonality (CH_3Cl, CH_3I, and COS)

Changes in CH_3Cl, CH_3I, and COS were affected by the sinusoidal wave used to represent season at all prediction horizons tested (ranging from 0 to −3 weeks) and by changes in certain environmental variables and certain diatom and/or dinoflagellate taxa at various prediction horizons (Fig. 6). It is noteworthy that several of the CCM relationships in which changes in environmental variables were shown to affect annually-seasonal trace gases were not significant (Fig. 6) and tended to show bidirectional forcing, even when significantly different (Fig. S2). Meanwhile, CCM relationships in which changes in phytoplankton affected changes in CH_3Cl, CH_3I, and COS were always significant (Fig. 6) and did not demonstrate bidirectional forcing.

Numerous phytoplankton taxa had a causal relationship with seasonally abundant trace gases. Changes in CH_3Cl were significantly affected by changes in three diatom genera/species (*Licmophora*, *Odontella*, and *Plagiogramma vanheurickii*), changes in CH_3I were affected by changes in four diatom genera/species (*Asterionellopsis*, *C. closterium*, *Odontella*, and *Proboscia alata*), and changes in COS were affected by changes in three dinoflagellate genera/species (*Dinophysis*, *Cochlodinium*, and *Scrippsiella trochiodea*) (Fig. 6B, C). These results spanned all four prediction horizons tested (ranging from a delay of 0 weeks to prediction to −3 weeks to prediction), with changes in *Odontella* spp. affecting changes in CH_3Cl at multiple prediction horizons. We found that utilizing changes in these phytoplankton at their lowest significant lags (vs. only using a trace gas's own lagged components) increased the forecast skill of generalized forecasting with mixed embeddings for changes in all three of these trace gases (Table 1). Since changes in CH_3I had an optimal embedding dimension of one (and we had to go over the optimal embedding dimension to assess all relevant phytoplankton at once) we also assessed the

363 generalized forecasting skill of changes in each significant phytoplankton at their optimal lags
364 individually. Changes in each *Asterionellopsis* spp. and *Odontella* spp. forecasted changes in
365 CH_3I better than when only utilizing changes in current CH_3I (Table 1). Additionally, changes in
366 *C. closterium* and *Proboscia alata* were each almost as good at forecasting changes in CH_3I as
367 when only using changes in current CH_3I (Table 1).

368 Several of the phytoplankton causally linked to seasonal trace gases themselves displayed
369 seasonality. Changes in *Asterionellopsis* spp., *C. closterium*, *Cochlodinium* spp., *Limophora*
370 spp., and *P. vanheurickii* all exceeded our threshold of $\rho > 0.1$ (Fig. 7). Seasonal highs in CH_3Cl,
371 CH_3I, and COS often did not coincide with the highest abundances of these similarly seasonal
372 phytoplankton. Rather, the highest abundances of all relevant phytoplankton (at the lowest
373 significant lags) tended to coincide with moderate peaks in these annually-seasonal trace gases
374 and vice versa (Fig. S3, S4, S5). Similarly, the weekly change in the mixing ratios of CH_3Cl and
375 COS and the change in the sum of significant phytoplankton for each trace gas (at the lowest
376 significant lags) were only synced during some states, with synching at least once a year for
377 CH_3Cl and during a phytoplankton bloom in 2017 for COS (Fig. S3E, S5E). However, the
378 weekly change in the mixing ratio of CH_3I had a more direct correlation with the change in the
379 sum of significant phytoplankton (at the lowest significant lags), primarily during EM2 and EM7
380 (Fig. S4F, G, H).

381 *Variables affecting annually-seasonal trace gases with a non-seasonal component (CH_2Br_2 and*
382 *$CHBr_3$)*

383 Despite the strong aseasonal peaks in 2014, changes in CH_2Br_2 and $CHBr_3$ were each
384 shown to be affected by the change in the sinusoidal wave that represented season at all
385 prediction horizons and changes in several environmental variables (Fig. 6A). However, there

386 was bidirectional forcing between changes in CH_2Br_2 and $CHBr_3$ and changes in season and
387 environmental variables at many prediction horizons (Fig. S2A). Similar to the seasonal trace
388 gases discussed above, changes in phytoplankton that were shown to affect changes in CH_2Br_2
389 and $CHBr_3$ did not display bidirectional forcing. Changes in average phaeophytin concentration
390 also affected changes in $CHBr_3$ (Fig. 6A).

391 For phytoplankton, changes in two dinoflagellate species (*Ceratium furca* and *Ceratium*
392 *fusus*) significantly affected changes in CH_2Br_2 and changes in one diatom genus (*Rhizosolenia*)
393 significantly affected changes in $CHBr_3$ (Fig. 6B, C). We found that utilizing changes in these
394 phytoplankton and phaeophytin at their lowest significant lags (vs. only using a trace gas's own
395 lagged components) did not increase the forecast skill of generalized forecasting with mixed
396 embeddings for changes in either gas (Table 1).

397 Of the phytoplankton causally linked to CH_2Br_2 and $CHBr_3$, changes in *C. fusus*
398 displayed seasonality (Fig. 7). Both *C. fusus* and C. *furca* (when lagged) had peaks that
399 overlapped with moderate peaks of CH_2Br_2, and *C. fusus* had a peak that coincided with a large
400 peak immediately following the 2014 spike in CH_2Br_2 (Fig. S6A, C). Changes in both *C. furca*
401 and C. *fusus* synced with changes in weekly incremented CH_2Br_2 at several points throughout the
402 study, despite anticorrelation at other points, such as late 2011, late 2012, mid 2017, and mid
403 2018 (Fig. S6B, D). Meanwhile, when lagged, *Rhizosolenia* spp. had population peaks that
404 precipitated many of the local peaks in $CHBr_3$, with syncing at least once per year between
405 weekly changes in the trace gas and phytoplankton (when *Rhizosolenia* spp. were manually
406 lagged), and included peaks during the 2014 spike (Fig. S6F, G). Finally, the average
407 phaeophytin concentration tended to be inversely related to the $CHBr_3$ mixing ratio, with
408 extremely low phaeophytin concentrations observed during the 2014 spike in $CHBr_3$ and the

highest phaeophytin concentrations observed during a period in 2018 when $CHBr_3$ was low (Fig. S6E).

Variables affecting trace gases that displayed multiple-seasonality to no seasonality (CH_3Br and $CHCl_3$)

While neither CH_3Br nor $CHCl_3$ displayed visible annual seasonality, CCM revealed that changes in season affected changes in $CHCl_3$ at all prediction horizons tested (Fig. 6). Changes in both trace gases were also affected by changes in several environmental variables (though many relationships were bidirectional) and changes in $CHCl_3$ were affected by changes in phaeophytin at all prediction horizons tested (Fig. 6A, 7A).

For phytoplankton, changes in two diatom genera (*Chaeotoceros* and *Gonyaulax*), the sum of total diatoms, and three dinoflagellate genera/species (*Gonyaulax*, *Prorocentrum*, and *Scrippsiella trochoidea*) affected changes in CH_3Br, while changes in four diatoms genera/species (*Navicula*, *Odontella*., *Proboscia alata*, and *Thalassiosira*) and two dinoflagellate genera/species (*Dinophysis acuminate* and *Ceratium*) affected changes in $CHCl_3$, with certain taxa affecting each trace gas at multiple prediction horizons (Fig. 6B, C). We found that utilizing changes in these significant phytoplankton at their lowest significant lags (vs. only using a trace gas's own lagged components) increased the forecast skill of generalized forecasting with mixed embeddings for changes in $CHCl_3$ but not for CH_3Br (Table 1). Since changes in CH_3Br had an optimal embedding dimension of three (and we had to go over the optimal embedding dimension to assess all relevant phytoplankton at once), we also assessed the generalized forecasting skill of changes in current CH_3Br with the change in the two most abundant phytoplankton genera CCM selected (*Prorocentrum* and *Chaetoceros*) at their lowest significant lags and observed an increase in forecast skill. Generalized forecasting of changes in $CHCl_3$ was further improved by

including the change in average phaeophytin concentration in addition to changes in relevant phytoplankton abundance.

Of the phytoplankton causally linked to CH_3Br and $CHCl_3$, changes in *Prorocentrum* spp. (for CH_3Br) and *Ceratium* spp. and *Thalassiosira* spp. (for $CHCl_3$) displayed seasonality (Fig. 7). Similar to the other trace gases, phytoplankton abundance (when lagged) often peaked during smaller local peaks of CH_3Br and $CHCl_3$ (Fig. S7, S8). However, there were several instances in which large peaks between phytoplankton and CH_3Br or $CHCl_3$ overlapped and other times when trace gas peaks were left unexplained by phytoplankton abundance (Fig. S7G, S8H). This was again demonstrated by changes in weekly incremented CH_3Br or $CHCl_3$ and the sum of the important phytoplankton at the lowest significant lags, which had many instances that correlated throughout the time-series (Fig. S7H, S8I). There were also many instances in which the change in CH_3Br was anticorrelated with the change in causally-linked phytoplankton (Fig. S8I).

Discussion

Through the use of EDM with CCM we were able to show that changes in several species, genera, and types of eukaryotic phytoplankton led to changes in select trace gases in a nearshore Southern California coastal environment. Though ρ values for various phytoplankton affecting trace gases were low, our findings demonstrate that phytoplankton community composition has an impact on atmospheric chemistry in the marine boundary layer. Phytoplankton influence trace gas mixing ratios through direct production, indirectly by producing an intermediate, or by stimulating production by another organism. Put another way, CCM did not show us that any single phytoplankton taxon was solely responsible for changes in any of the trace gases measured. Rather, the assemblage and abundance of groups of phytoplankton contributed to changes in atmospheric trace gases, some regularly and some

455 sporadically. Therefore, the changes in the sum of several different phytoplankton were one of
456 many factors that affected changes in trace gases in coastal Southern California over this time-
457 series, and the important phytoplankton taxa varied depending on the trace gas in question.
458 Finally, we saw that phytoplankton affected atmospheric trace gases at all prediction horizons
459 tested (*i.e.* time lags). While a lag of 3 weeks may seem long, Sœmundsdóttir and Matrai (1998)
460 observed a lag of one to three weeks for the peak CH_3Br signal in phytoplankton cultures grown
461 under optimal conditions, demonstrating the need for time-delayed ecological studies relating
462 phytoplankton to environmental parameters.

463 EDM with CCM has strengths and drawbacks. The strength of EDM with CCM is that it
464 is able to show causation in non-linear systems that experience different system states (so
465 variables may be related at certain times but not others). While we were not able to measure or
466 find measurements for all the variables that could have affected trace gases in the open system of
467 coastal Southern California, EDM with CCM does not require that all variables are measured.
468 Rather, a time-series' own lagged components can be substituted for variables that affected it.
469 Limitations of the method are that it can be difficult to differentiate between direct versus
470 indirect causation unless data collection is extremely frequent and ρ values are very high. Such
471 values are unlikely for the present study as many other factors affect the flux of trace gases from
472 the water (or land) to the atmosphere. Additionally, CCM does not easily translate into the
473 amount of a trace gas that was produced by a specific taxon with the resolution and open nature
474 of the available data. This was further masked by the fact that we related the changes in
475 phytoplankton taxa to changes in trace gases. This allowed us to eliminate autocorrelation for
476 trace gases and phytoplankton that exhibited similar seasonality and to zone in on phytoplankton
477 whose populations were actively growing and decreasing, but made it so that comparison of raw

478 phytoplankton abundance did not directly relate to raw trace gas mixing ratios. Finally, while
479 environmental variables changed as phytoplankton populations changed, the strength of the ρ
480 values and the lack of bidirectional forcing for phytoplankton helped to separate when change in
481 a trace gas was purely due to an environmental variable versus affected by phytoplankton.

482 We approached the data with two broad expectations. First, that eukaryotic
483 phytoplankton with annual seasonality may contribute to changes in mixing ratios of halocarbons
484 and/or COS that also display annual seasonality clearly defined within the same season. Second,
485 that eukaryotic phytoplankton exhibiting abundance patterns other than seasonal are more likely
486 to contribute to the mixing ratios of halocarbons and/or COS that display less clear seasonality.
487 Our results suggested the presence of three scenarios in which we tested our expectations: (1)
488 trace gases that displayed visible annual seasonality (CH_3Cl, CH_3I, and COS); (2) trace gases
489 with visible annual seasonality with a strong aseasonal feature (CH_2Br_2 and $CHBr_3$); and (3)
490 trace gases that displayed multiple-seasonality (with reliable peaks in multiple seasons) or no
491 seasonality (CH_3Br and $CHCl_3$). Within these three categories, eukaryotic phytoplankton with
492 and without seasonality (assessed via CCM) affected trace gases in each of the above categories,
493 suggesting that both cyclical and stochastic eukaryotic phytoplankton contribute to traces gases
494 with annual seasonality, multiple-seasonality, and non-seasonality.

495 In the case of CH_3Cl, our first expectation of annually seasonal trace gases being affected
496 by annually seasonal phytoplankton was demonstrated partially true, as changes in both seasonal
497 and non-seasonal phytoplankton affected changes in CH_3Cl (Fig. S3). Consistent with our
498 findings, spikes in coastal atmospheric CH_3Cl mixing ratios have previously been associated
499 with lower wind speeds during winter months (Yokouchi *et al.*, 2000). As the largest natural
500 contributor to atmospheric chlorine, CH_3Cl has numerous potential sources in Southern

501 California, including terrestrial plants, salt marshes (Rhew *et al.*, 2000; Rhew *et al.*, 2002), and
502 the coastal and open ocean (WMO, 2018). The different potential sources cannot be separated
503 out as CH_3Cl has a fairly long atmospheric life-time (~0.9 years; WMO, 2018), but potential
504 oceanic sources include a range of eukaryotic phytoplankton that have been observed to emit
505 CH_3Cl (Tait and Moore, 1995; Scarratt and Moore, 1996; Colomb *et al.*, 2008; Lim *et al.*, 2018).
506 Our CCM results reinforce these findings. However, the phytoplankton that were observed to
507 affect atmospheric CH_3Cl were associated with local peaks in CH_3Cl (rather than the prolonged
508 seasonal peak), and again did not all display strong seasonality themselves (Fig. S3). Further,
509 while we did not observe any of the phytoplankton species that have been shown to emit CH_3Cl
510 in past laboratory cultures (Scarratt and Moore, 1996; Scarratt and Moore, 1998; Colomb *et al.*,
511 2008; Lim *et al.*, 2018), the number of taxa in these studies indicates that emission of CH_3Cl is
512 found in a range of phytoplankton. Taken together these data suggest that smaller fluctuations in
513 CH_3Cl, rather than the large seasonal signal, are potentially linked to growth of a range of
514 eukaryotic phytoplankton in coastal Southern California, while the large seasonal cycle is due to
515 other factors.

516 CH_3I exhibited a complex relationship with seasonally abundant phytoplankton and our
517 expectation about seasonally active phytoplankton affecting it may be true. CH_3I is a key species
518 in the biogeochemical cycling of iodine. Because it has a relatively short lifetime in the
519 atmosphere (<14 days; WMO, 2018), the ocean—which is supersaturated in CH_3I—is thought to
520 be a major source (Moore, 2003), potentially making it easier to identify the link between
521 changes in local phytoplankton and CH_3I mixing ratios (as the sources are likely closer to the
522 pier). Consistent with previous work, we observed increased atmospheric CH_3I mixing ratios in
523 summer and autumn months (Yokouchi *et al.*, 2011). The major seasonal peaks in CH_3I

524 correlated with EM2 and EM3 and, to a lesser extent, EM6 and EM7. EM2 and EM3 each had
525 warmer temperatures with low wind speeds originating from western and southern directions,
526 supporting an oceanic source. Photochemical production of CH_3I has been reported in the
527 laboratory (Moore and Zafiriou, 1994) and the open ocean (Happell and Wallace, 1996), making
528 it a potential seasonal source in coastal Southern California. Nonetheless, CH_3I production has
529 been observed for cultured phytoplankton originating from coastal areas (Scarratt and Moore,
530 1999), co-cultures of phytoplankton and bacteria (Manley and de la Cuesta, 1997), macroalgae
531 (Manley *et al.*, 1992), and macroalgae and associated bacteria (Manley and Dastoor, 1988). CH_3I
532 has also been linked to phytoplankton abundance in the open ocean using back trajectories
533 (Arnold *et al.*, 2010).

534 Of the four diatom taxa that were shown to affect atmospheric mixing ratios of CH_3I in
535 our study, only *C. closterium*, which exhibited a strong change in abundance with season but was
536 always present, likely contributed significantly to the seasonal cycle of CH_3I. This contribution is
537 expected primarily when changes in relevant phytoplankton and CH_3I synced during EM2, when
538 winds originated from oceanic directions, and EM7, which had low-speed winds (Fig. S4B, F, G,
539 H). Other phytoplankton were also causally associated with smaller fluctuations (Fig. S4). It
540 should be noted that it can be difficult to discern between direct production of a trace gas by a
541 given phytoplankton taxon versus indirection production (*e.g.*, production by heterotrophic
542 bacteria responding to a phytoplankton bloom). Though this distinction is beyond the scope of
543 this study, the extremely transient nature of several of the phytoplankton that affected changes in
544 CH_3I support the likelihood that heterotrophic bacterial decomposition of organic matter (that
545 these phytoplankton contributed to during certain system states) was likely a source of CH_3I in
546 this system.

COS did not meet our first expectation of annually seasonal trace gases being affected by annually seasonal phytoplankton. COS is the one trace gas we assessed that is not a halogenated compound. While it is just one of many sulfur species commonly found in seawater (Liss *et al.*, 1997) it is of global importance because it is the longest-lived of all atmospheric sulfur compounds (~2 years; WMO, 2018), thus a fraction of it is transported into the stratosphere and contributes to the sulfate pool, affecting radiative transfer (Andreae and Ferek, 1992; Andreae and Crutzen, 1997; Whelan *et al.*, 2018). Oceanic COS emissions tend to peak in summer while terrestrial sources peak in the winter (Ulshöfer *et al.*, 1995; Weiss *et al.*, 1995), with uptake by terrestrial vegetation highest in summer months in the northern hemisphere (Kettle *et al.*, 2002). Our results suggest that oceanic sources were dominant at our study site as we observed peaks in the summer, and COS mixing ratios were lowest during EM6 and EM7, when winds came from terrestrial directions.

Despite oceanic origins, phytoplankton did not drive most of the change in atmospheric COS during our time-series as most variation did not correlate with phytoplankton or changes in phytoplankton. Notably, the three causally-linked dinoflagellates were extremely rare in 2014 and 2015 (Fig. S5). Because the dominant seasonal cycle was still present during these two years, we can assume that alternate sources of organic sulfur compounds drive COS production in this system. COS has been observed to be produced photochemically in seawater from dimethylsulfide and carbon disulfide (Ferek and Andreae, 1984; Andreae and Ferek, 1992; Andreae and Crutzen, 1997; Kettle *et al.*, 2002) which seems like a more likely source. However, there were notable exceptions (*e.g.*, the summer of 2017 and 2018), which are why changes in these phytoplankton were linked to changes in COS. Of the phytoplankton connected to COS, *Cochlodinium* is a harmful algal bloom-forming genus that has been associated with red tides

570 (Curtiss *et al.*, 2008) and, similar to several taxa that affected CH_3Cl and CH_3I, *Cochlodinium*
571 spp. was found very infrequently in the dataset.

572 The second scenario suggested by our dataset had to do with trace gases that displayed
573 annual seasonality but had a strong aseasonal feature that dwarfed the seasonal cycle. CH_2Br_2
574 and $CHBr_3$ each displayed a strong annual seasonal cycle but had a large aseasonal peak in 2014.
575 Specifically, we expected that the seasonal signal in these trace gases would be due to
576 phytoplankton with similar seasonality, while the aseasonal peak would either be due to a similar
577 aseasonal spike in an otherwise seasonal phytoplankton or due to stochastically present
578 phytoplankton. For our data both of these appeared false. Previous research has not linked the
579 production of either CH_2Br_2 or $CHBr_3$ to the abundances of individual phytoplankton in the
580 environment (Quack *et al.*, 2007; Kurihara *et al.*, 2012), but both trace gases have been
581 correlated with phytoplankton via marker pigments (Quack *et al.*, 2007; Roy, 2010) and linked to
582 phytoplankton abundances using gas back trajectories (Arnold et al, 2010). However, we did not
583 find evidence for a major phytoplankton derived origin for either trace gas with the relatively
584 few phytoplankton that were shown to affect CH_2Br_2 and $CHBr_3$ in our dataset. Only *C. fusus*
585 exhibited seasonality, and none of the observed phytoplankton had an abnormally large
586 population spike the coincided with the start of the aseasonal peak in either CH_2Br_2 or $CHBr_3$
587 (Fig. 6B, C, S6). As we saw with CH_3I, it can be easier to identify significant links with
588 phytoplankton when the lifetime of the trace gas is short; however, the lifetimes of CH_2Br_2 and
589 $CHBr_3$ likely did not impede identification of links with phytoplankton as CH_2Br_2 has a long
590 lifetime (~150 days; WMO, 2018) while $CHBr_3$ has a short lifetime (~16 days). We also note
591 that it is likely that some CH_2Br_2 was formed from $CHBr_3$, as the peak in CH_2Br_2 occurred
592 slightly after the peak in $CHBr_3$. $CHBr_3$ can be broken down to CH_2Br_2 either abiotically in the

593 presence of light (Betterton *et al.*, 1995) or via heterotrophic transformation (Ichikawa *et al.*,
594 2015; Kataoka *et al.*, 2019).

595 An alternate source of CH_2Br_2 and $CHBr_3$ may have been the direct production by
596 macroalgae, or transformation of alga-derived photosynthate by heterotrophic microorganisms.
597 Production of CH_2Br_2 and $CHBr_3$ has been observed in laboratory cultures of macroalgae
598 (Manley *et al.*, 1992; Nightingale *et al.*, 1995; Goodwin *et al.*, 1997) and environmental analysis
599 has found higher concentrations of brominated halocarbons in the coastal ocean, with the highest
600 concentrations typically found near macroalgae (Carpenter *et al.*, 2009). However,
601 phytoplankton in open ocean upwelling zones are potentially significant contributors (Quack *et*
602 *al.*, 2007; Raimund *et al.*, 2010) and it should be noted that multiple EMs had winds originating
603 from offshore areas, including summertime when offshore upwelling peaks in Southern
604 California (Bakun, 1973). Specifically, 13 of the 14 highest mixing ratios for $CHBr_3$ in 2014
605 occurred during EM2 and EM3, which each had winds originating from oceanic directions.
606 However, these two EMs also had the lowest wind speeds (and originated from slightly different
607 oceanic directions), further suggesting that $CHBr_3$ and CH_2Br_2 were each likely produced close
608 to SIO as the aseasonal peak persisted despite switching between EM2 and EM3. And while
609 moderate to extremely high mixing ratios of CH_2Br_2 and $CHBr_3$ co-occurred with peaks and
610 changes in key taxa at certain time points, most phytoplankton coincided with lower mixing
611 ratios or smaller local peaks of CH_2Br_2 and $CHBr_3$.

612 The strongest evidence that the aseasonal spike in CH_2Br_2 and $CHBr_3$ was linked to
613 macroalgae was that it coincided with a peak in kelp coverage for the San Diego coastal region
614 (Table 2; MBC, 2017) in 2013 and 2014. While we do not have enough data to assess why
615 macroalgae were more abundant in these years, summer is when light levels and offshore

upwelling in Southern California are highest (Bakun, 1973) and macroalgae tend to be most abundant. Further, light levels are preferential for VOC production (Bondu *et al.*, 2008; Carpenter *et al.*, 2000) and $CHBr_3$ is a known signaling molecule, with previous work suggesting that macroalgae produce it as a defense mechanism against microorganisms and herbivores (McConnell and Fenical, 1977; Gschwend *et al.*, 1985). The ecosystem also experienced a prolonged warm-water perturbation termed the 'blob' in the boreal winter of 2013/2014, during which a wide-spread positive sea surface heat anomaly existed in the eastern North Pacific (Leising *et al.*, 2015). This anomaly may have been a factor in either macroalgae or heterotrophic production of CH_2Br_2 and $CHBr_3$ as the peaks occurred in the summer immediately following this event, as kelp was decreasing and relatively few phytoplankton were observed. Finally, we note that *C. closterium* also had a large population spike in 2014 that coincided with the start of the CH_2Br_2 or $CHBr_3$ peaks (Fig. S4A). It is not possible to accurately model the occurrence of a singular event, but further observations of CH_2Br_2 or $CHBr_3$ aseasonal spikes may allow us to resolve a causal relationship between *C. closterium* and/or macroalgae with $CHBr_3$.

The last scenario suggested by our dataset had to do with trace gases that displayed multiple-seasonality or no seasonality. Both CH_3Br and $CHCl_3$ fell into this category, with CH_3Br showing no seasonality and $CHCl_3$ possessing a seasonal signal masked by multi-year peaks. CH_3Br is a major source of bromine to the atmosphere, supplying more bromine to the stratosphere than $CHBr_3$ and CH_2Br_2 combined. The coastal and open ocean are thought to be important sources of CH_3Br (Singh and Kanakidou, 1993; Lobert *et al.*, 1995; King *et al.*, 2000). In addition to anthropogenic sources (Moore, 2003), production of CH_3Br has been associated with many types of phytoplankton in laboratory studies (Scarratt and Moore, 1996; Sœmundsdóttir and Matrai, 1998) and observed in natural marine habitats in Southern California

639 (Rhew *et al.*, 2000). Due to the relatively long lifetime (~0.8 years; WMO, 2018), multiple
640 source regions may influence the mixing ratio measured at the Scripps pier, potentially making it
641 difficult to identity correlations with local phytoplankton populations. Meanwhile, $CHCl_3$, like
642 other halocarbons discussed here, is of concern due to its ability to destroy ozone when it reaches
643 the stratosphere and, with 90% of emissions estimated to originate naturally, offshore seawater is
644 thought to be the largest source (Khalil *et al.*, 1999; Khalil and Rasmussen, 1999; Keene *et al.*,
645 1999; McCulloch, 2003). While this trace gas has a relatively long atmospheric lifetime (~183
646 days; WMO, 2018), many different macroalgae have been observed to release $CHCl_3$
647 (Nightingale *et al.*, 1995; Baker *et al.*, 2001), with additional studies linking it to phytoplankton
648 in laboratory studies (Scarratt and Moore, 1999; Colomb *et al.*, 2008), the open ocean (Arnold *et*
649 *al.*, 2010), and to phytoplankton pigments (Roy, 2010).

650 Our expectation regarding stochastic eukaryotic phytoplankton contributing to trace gas
651 mixing ratios that lack clear annual seasonality may be true during some system states, as many
652 significant phytoplankton did not exhibit seasonality. Unlike some of the previous trace gases
653 discussed, several of the significant phytoplankton taxa were present during major atmospheric
654 peaks in in CH_3Br and $CHCl_3$ (Fig. S7, S8). CH_3Br also had many instances when the change in
655 CH_3Br was anticorrelated with the change in the abundance of important phytoplankton (*e.g.*,
656 winter of 2012 and 2014; Fig. S7H), indicating that the death and decay of these phytoplankton
657 were likely contributing to the production of CH_3Br. However, none of the statistically
658 significant phytoplankton were associated with all major peaks or changes (either directly or
659 anticorrelated) in CH_3Br or $CHCl_3$. Further, there were several strong peaks and changes in both
660 trace gases that were not accompanied by large peaks in any of the statistically significant diatom
661 or dinoflagellate taxa, indicating there are other important sources in this system. Many of the

662 CH_3Br and $CHCl_3$ peaks that were not accompanied by peaks in phytoplankton represent periods
663 when the system was in EM6 and EM7, when wind direction was easterly (overland), and so
664 atmospheric mixing ratios likely had strong terrestrial influences.

665 $CHCl_3$, in particular, was a good example of both local and non-local sources being
666 important factors to local mixing ratios. $CHCl_3$ is a relatively stable molecule with residence
667 times ranging from approximately 3 months to 3 years in the atmosphere (WMO, 2018), which
668 may be why there was no strong annual seasonality in its atmospheric mixing ratio. This also
669 means that much of the observed changes in atmospheric $CHCl_3$ may have been from non-local
670 oceanic sources, perhaps stemming from industrial uses of $CHCl_3$ when the system was in EM6
671 and EM7. Nonetheless, EM2 and EM3, which had low wind speeds originating from oceanic
672 directions, also possessed many peaks in $CHCl_3$, and the greatest positive change in $CHCl_3$ was
673 seen when the system was in EM2 toward the end of 2017 (Fig. S8I), reinforcing the importance
674 of local oceanic sources. Thus, changes in general photosynthetic activity and degradation of
675 chlorophyll *a* (as indicated by the fact that average phaeophytin was found to be important and
676 because the change in relevant phytoplankton correlated with the change in $CHCl_3$ multiple
677 times each year) appear to affect periodic production of $CHCl_3$ in coastal Southern California
678 (Fig. S8A, I). Further, generalized forecasting with mixed embeddings utilizing changes in
679 significant phytoplankton and phaeophytin had the highest overall ρ value and was higher than
680 the ρ value only utilizing $CHCl_3$'s own lagged components (Table 1). All of these factors
681 indicate that phytoplankton and their activities are important to atmospheric $CHCl_3$.

682 **Conclusion**

683 This work demonstrated that naturally occurring, ecologically active eukaryotic
684 phytoplankton affect a range of trace gases in coastal Southern California on scales of 0 to 3

685 weeks during different times of the year. We utilized long-standing *in situ* observations to
686 capture the effects of naturally occurring environmental or ecological interactions, such as
687 predation and competition, and applied causality testing to assess directionality of relationships.
688 The method successfully identified some causal interactions between coastal marine
689 phytoplankton and atmospheric trace gases. Overall, our work supports the understanding that
690 many different phytoplankton directly or indirectly produce CH_3Br, CH_2Br_2, $CHBr_3$, CH_3Cl,
691 $CHCl_3$, CH_3I and COS. In some cases, these trace gases may affect phytoplankton populations as
692 well, a facet that was beyond the scope of this analysis but warrants further study. We also
693 showed that, while strong seasonal signals in these atmospheric trace gases are not due to local
694 eukaryotic phytoplankton, these phytoplankton can affect small changes in atmospheric trace
695 gases with strong seasonality and may, at times, play a significant role in the production of trace
696 gases that do not exhibit clear seasonality. The links observed between changes in CH_3I and
697 some phytoplankton taxa may have been due to the short atmospheric lifetime of this gas (which
698 potentially increases the observable influence of local sources) while links between $CHCl_3$ and
699 phytoplankton during specific system states were observed in spite of its relatively long
700 atmospheric lifetime. By identifying specific phytoplankton taxa that affect trace gases, this work
701 serves as a springboard for targeted laboratory studies that seek to characterize the mechanisms
702 of trace gas release and future *in situ* studies that use correlative and/or causality testing in order
703 to relate trace gases and phytoplankton taxa. In certain cases, we also highlighted the need for
704 measuring heterotrophic activities and dissolved organic matter to fully understand biology's
705 role in the production of atmospheric trace gases in coastal marine environments.

706

707 **Acknowledgements**

The authors wish to specifically acknowledge Kristi Seech and Mary Hilbern for their efforts on phytoplankton cell identification and enumeration, James Fumo for data integration, John A. McGowan for support and leadership over the decades with the McGowan Plankton and Chlorophyll Program (Funding provided by private donors and the MacArthur Foundation) and Southern California Coastal Ocean Observing Harmful Algal Bloom Monitoring Program (NOAA NA16NOS0120022, NA11NOS120029, and NA17RJ1231). The authors would also like to thank Chase James and Jooil Kim for their advice and comments throughout the analysis and Emelia Chamberlain and Sarah Abboud for reading drafts of the manuscript. The trace gas measurements at the SIO pier and the SIO calibration activities are part of the Advanced Global Atmospheric Gases Experiment (AGAGE) and are supported by the National Aeronautics and Space Administration (NAG5-4023, NNX07AE87G, NNX07AF09G, NNX11AF15G, and NNX11AF16G). We thank Chris M. Harth and Roland Schmidt for their critical support of the AGAGE operations at SIO. JSB was funded by the Simons Foundation Early Career Marine Microbial Investigator program.

723 References

724 Andreae, M.O., Crutzen, P.J., 1997. Atmospheric aerosols: Biogeochemical sources and role in
725 atmospheric chemistry. Science 276, 1052–1058.
726 https://doi.org/10.1126/science.276.5315.1052

727 Andreae, M.O., Ferek, R.J., 1992. Photochemical production of carbonyl sulfide in seawater and
728 its emission to the atmosphere. Global Biogeochem. Cycles 6, 175–183.
729 https://doi.org/10.1029/91GB02809

730 Arnold, S.R., Spracklen, D. V., Gebhardt, S., Custer, T., Williams, J., Peeken, I., Alvain, S.,
731 2010. Relationships between atmospheric organic compounds and air-mass exposure to
732 marine biology. Environ. Chem. 7, 232. https://doi.org/10.1071/EN09144

733 Baker, A.R., Turner, S.M., Broadgate, W.J., Thompson, A., McFiggans, G.B., Vesperini, O.,
734 Nightingal, P.D., Liss, P.S., Jickells, T.D., 2000. Distribution and sea-air fluxes of biogenic
735 trace gases in the eastern Atlantic Ocean. Global Biogeochem. Cycles 14, 871–886.
736 https://doi.org/10.1029/1999GB001219

737 Baker, J.M., Sturges, W.T., Sugier, J., Sunnenberg, G., Lovett, A.A., Reeves, C.E., Nightingale,
738 P.D., Penkett, S.A., 2001. Emissions of CH_3Br, organochlorines, and organoiodines from
739 temperate macroalgae. Chemosph.-Glob. Chang. Sci. 3, 93–106.
740 https://doi.org/10.1016/S1465-9972(00)00021-0

741 Bakun, A., 1973. Coastal upwelling indices, west coast of North America, 1946-71. U.S. Dep.
742 Commer. NOAA Tech. Rep. NMFS SSRF-671, 103 p.

743 Betterton, E.A., Arnold, R.G., Kuhler, R.J., Santo, G.A., 1995. Reductive dehalogenation of
744 bromoform in aqueous solution. Environ. Health Perspect. 103, 89–91.
745 https://doi.org/10.1289/ehp.95103s489

746 Blezinger, S., Wilhelm, C., Kesselmeier, J., 2000. Enzymatic consumption of carbonyl sulfide
747 (COS) by marine algae. Biogeochemistry 48, 185–197.
748 https://doi.org/10.1023/A:1006134213995

749 Bondu, S., Cocquempot, B., Deslandes, E., Morin, P., 2008. Effects of salt and light stress on the
750 release of volatile halogenated organic compounds by Solieria chordalis: A laboratory
751 incubation study. Bot. Mar. 51, 485–492. https://doi.org/10.1515/BOT.2008.056

752 Bowman, J.S., Kavanaugh, M.T., Doney, S.C., Ducklow, H.W., 2018. Recurrent seascape units
753 identify key ecological processes along the western Antarctic Peninsula. Glob. Chang. Biol.
754 24, 3065–3078. https://doi.org/10.1111/gcb.14161

755 Brownell, D.K., Moore, R.M., Cullen, J.J., 2010. Production of methyl halides by
756 *Prochlorococcus* and *Synechococcus*. Global Biogeochem. Cycles 24.
757 https://doi.org/10.1029/2009GB003671

Carpenter, L.J., Jones, C.E., Dunk, R.M., Hornsby, K.E., Woeltjen, J., 2009. Atmospheric Chemistry and Physics Air-sea fluxes of biogenic bromine from the tropical and North Atlantic Ocean, Atmos. Chem. Phys.

Carpenter, L.J., Malin, G., Liss, P.S., Küpper, F.C., 2000. Novel biogenic iodine-containing trihalomethanes and other short-lived halocarbons in the coastal east Atlantic. Global Biogeochem. Cycles 14, 1191–1204. https://doi.org/10.1029/2000GB001257

Carpenter, L.J., Reimann, S., Burkholder, J.B., Clerbaux, C., Hall, B.D., Hossaini, R., Laube, J.C., Yvon-Lewis, S.A., Blake, D.R., Dorf, M., Dutton, G.S., Fraser, P.J., Froidevaux, L., Hendrick, F., Hu, J., Jones, A., Krummel, P.B., Kuijpers, L.J.M., Kurylo, M.J., Liang, Q., Mahieu, E., Mühle, J., O'doherty, S., Ohnishi, K., Orkin, V.L., Pfeilsticker, K., Rigby, M., Simpson, I.J., Yokouchi, Y., 2014. CHAPTER 1 Update on Ozone-Depleting Substances (ODSs) and Other Gases of Interest to the Montreal Protocol UPDATE ON OZONE-DEPLETING SUBSTANCES (ODSs) AND OTHER GASES OF INTEREST TO THE MONTREAL PROTOCOL Contents, Global Ozone Research and Monitoring Project-Report. Montreal Protocol.

Colomb, A., Yassaa, N., Williams, J., Peeken, I., Lochte, K., 2008. Screening volatile organic compounds (VOCs) emissions from five marine phytoplankton species by head space gas chromatography/mass spectrometry (HS-GC/MS). J. Environ. Monit. 10, 325–330. https://doi.org/10.1039/b715312k

Commane, R., Herndon, S.C., Zahniser, M.S., Lerner, B.M., McManus, J.B., Munger, J.W., Nelson, D.D., Wofsy, S.C., 2013. Carbonyl sulfide in the planetary boundary layer: Coastal and continental influences. J. Geophys. Res. Atmos. 118, 8001–8009. https://doi.org/10.1002/jgrd.50581

Curtiss, C.C., Langlois, G.W., Busse, L.B., Mazzillo, F., Silver, M.W., 2008. The emergence of Cochlodinium along the California Coast (USA). Harmful Algae 7, 337–346. https://doi.org/10.1016/j.hal.2007.12.012

Deyle, E.R., Maher, M.C., Hernandez, R.D., Basu, S., Sugihara, G., 2016. Global environmental drivers of influenza. Proc. Natl. Acad. Sci. U. S. A. 113, 13081–13086. https://doi.org/10.1073/pnas.1607747113

Deyle, E., Schueller, A.M., Ye, H., Pao, G.M., Sugihara, G., 2018. Ecosystem-based forecasts of recruitment in two menhaden species. Fish and Fisheries. 19, 769–781. https://doi.org/10.1111/faf.12287

Di Lorenzo, E., 2003. Seasonal dynamics of the surface circulation in the Southern California Current System. Deep. Res. Part II Top. Stud. Oceanogr. 50, 2371–2388. https://doi.org/10.1016/S0967-0645(03)00125-5

Ferek, R.J., Andreae, M.O., 1984. Photochemical production of carbonyl sulphide in marine surface waters. Nature 307, 148–150. https://doi.org/10.1038/307148a0

Goodwin, K.D., North, W.J., Lidstrom, M.E., 1997. Production of bromoform and dibromomethane by Giant Kelp: Factors affecting release and comparison to anthropogenic bromine sources. Limnol. Oceanogr. 42, 1725–1734. https://doi.org/10.4319/lo.1997.42.8.1725

Gschwend, P.M., MacFarlane, J.K., Newman, K.A., 1985. Volatile halogenated organic compounds released to seawater from temperate marine macroalgae. Science 227, 1033–1035. https://doi.org/10.1126/science.227.4690.1033

Happell, J.D., Wallace, D.W.R., 1996. Methyl iodide in the Greenland/Norwegian Seas and the tropical Atlantic Ocean: Evidence for photochemical production. Geophys. Res. Lett. 23, 2105–2108. https://doi.org/10.1029/96GL01764

Hatch, M.B.A., Schellenberg, S.A., Carter, M.L., 2013. Ba/Ca variations in the modern intertidal bean clam Donax gouldii: An upwelling proxy? Palaeogeogr. Palaeoclimatol. Palaeoecol. 373, 98–107. https://doi.org/10.1016/j.palaeo.2012.03.006

Ichikawa, K., Kurihara, M., Tamegai, H., Hashimoto, S., 2015. Decomposition of brominated organic halogens by cultures of marine proteobacteria: Phaeobacter, Roseobacter, and Rhodobacter. Mar. Chem. 176, 133–141. https://doi.org/10.1016/j.marchem.2015.09.003

Johnson, T.L., Brahamsha, B., Palenik, B., Mühle, J., 2015. Halomethane production by vanadium-dependent bromoperoxidase in marine *Synechococcus*. Limnol. Oceanogr. 60, 1823–1835. https://doi.org/10.1002/lno.10135

Kataoka, T., Ooki, A., Nomura, D., 2019. Production of Dibromomethane and Changes in the Bacterial Community in Bromoform-Enriched Seawater. Microbes Environ. 34, 215–218. https://doi.org/10.1264/jsme2.ME18027

Keene, W.C., Khalil, M.A.K., Erickson, D.J., McCulloch, A., Graedel, T.E., Lobert, J.M., Aucott, M.L., Gong, S.L., Harper, D.B., Kleiman, G., Midgley, P., Moore, R.M., Seuzaret, C., Sturges, W.T., Benkovitz, C.M., Koropalov, V., Barrie, L.A., Li, Y.F., 1999. Composite global emissions of reactive chlorine from anthropogenic and natural sources: Reactive Chlorine Emissions Inventory. J. Geophys. Res. Atmos. 104, 8429–8440. https://doi.org/10.1029/1998JD100084

Kettle, A.J., Kuhn, U., von Hobe, M., Kesselmeier, J., Andreae, M.O. 2002. Global budget of atmospheric carbonyl sulfide: Temporal and spatial variations of the dominant sources and sinks. J. Geophys. Res. 107, 4658. https://doi.org/10.1029/2002JD002187

Khalil, M.A.K., Moore, R.M., Harper, D.B., Lobert, J.M., Erickson, D.J., Koropalov, V., Sturges, W.T., Keene, W.C., 1999. Natural emissions of chlorine-containing gases: Reactive Chlorine Emissions Inventory. J. Geophys. Res. Atmos. 104, 8333–8346. https://doi.org/10.1029/1998JD100079

Khalil, M.A.K., Rasmussen, R.A., 1999. Atmospheric methyl chloride. Atmos. Environ. 33, 1305–1321. https://doi.org/10.1016/S1352-2310(98)00234-9

King, D.B., Butler, J.H., Montzka, S.A., Yvon-Lewis, S.A., Elkins, J.W., 2000. Implications of methyl bromide supersaturations in the temperate North Atlantic Ocean. J. Geophys. Res. Atmos. 105, 19763–19769. https://doi.org/10.1029/2000JD900251

Kohonen, T., 2001. Self-Organizing Maps, Springer Series in Information Sciences. Springer Berlin Heidelberg, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-56927-2

Kurihara, M., Iseda, M., Ioriya, T., Horimoto, N., Kanda, J., Ishimaru, T., Yamaguchi, Y., Hashimoto, S., 2012. Brominated methane compounds and isoprene in surface seawater of Sagami Bay: Concentrations, fluxes, and relationships with phytoplankton assemblages. Mar. Chem. 134–135, 71–79. https://doi.org/10.1016/j.marchem.2012.04.001

Leising, A.W., Schroeder, I.D., Bograd, S.J., Abell, J., Durazo, R., Gaxiola-Castro, G. *et al.*, 2015. State of the California Current 2014-15: Impacts of the Warm-Water " Blob". California Cooperative Oceanic Fisheries Investigations Reports, 56, 31-68.

Lim, Y.K., Phang, S.M., Sturges, W.T., Malin, G., Rahman, N.B.A., 2018. Emission of short-lived halocarbons by three common tropical marine microalgae during batch culture. J. Appl. Phycol. 30, 341–353. https://doi.org/10.1007/s10811-017-1250-z

Liss, P.S., Hatton, A.D., Malin, G., Nightingale, P.D., Turner, S.M., 1997. Marine sulphur emissions. Philos. Trans. R. Soc. B Biol. Sci. 352, 159–169. https://doi.org/10.1098/rstb.1997.0011

Lobert, J.M., Butler, J.H., Montzka, S.A., Geller, L.S., Myers, R.C., Elkins, J.W., 1995. A net sink for atmospheric CH_3Br in the East Pacific Ocean. Science 267, 1002–1005. https://doi.org/10.1126/science.267.5200.1002

Lund, U., Agostinelli, C., Arai, H., Gagliardi, A., Garcia Portugues, E., Giunchi, D., Irisson, J.-O., Pocernich, M., Rotolo, F., 2017. Package “circular.”

Manley, S.L., Dastoor, M.N., 1988. Methyl iodide (CH_3I) production by kelp and associated microbes. Mar. Biol. 98, 477–482. https://doi.org/10.1007/BF00391538

Manley, S.L., de la Cuesta, J.L., 1997. Methyl iodide production from marine phytoplankton cultures. Limnol. Oceanogr. 42, 142–147. https://doi.org/10.4319/lo.1997.42.1.0142

Manley, S.L., Goodwin, K., North, W.J., 1992. Laboratory production of bromoform, methylene bromide, and methyl iodide by macroalgae and distribution in nearshore southern California waters. Limnol. Oceanogr. 37, 1652–1659. https://doi.org/10.4319/lo.1992.37.8.1652

MBC Applied Environmental Sciences. 2017. Status of the Kelp Beds in 2016: Ventura, Los Angeles, Orange, and San Diego Counties. Prepared for the Central Region Kelp Survey

Consortium and Region Nine Kelp Survey Consortium. 78 p. http://kelp.sccwrp.org/2016/Status_of_the_Kelp_Beds_2016_Ventura_Los%20Angeles_Orange_and_San_Diego_Counties.pdf

McConnell, O., Fenical, W., 1977. Halogen chemistry of the red alga Asparagopsis. Phytochemistry 16, 367–374. https://doi.org/10.1016/0031-9422(77)80067-8

McCulloch, A., 2003. Chloroform in the environment: Occurrence, sources, sinks and effects. Chemosphere. https://doi.org/10.1016/S0045-6535(02)00697-5

McGowan, J.A., Deyle, E.R., Ye, H., Carter, M.L., Perretti, C.T., Seger, K.D., de Verneil, A., Sugihara, G., 2017. Predicting coastal algal blooms in southern California. Ecology 98, 1419–1433. https://doi.org/10.1002/ecy.1804

Moore, R.M., 2003. Marine sources of volatile organohalogens, in: Handbook of Environmental Chemistry. Springer Verlag, pp. 85–101. https://doi.org/10.1007/b10449

Moore, R.M., Tokarczyk, R., 1993. Volatile biogenic halocarbons in the northwest Atlantic. Global Biogeochem. Cycles 7, 195–210. https://doi.org/10.1029/92GB02653

Moore, R.M., Zafiriou, O.C., 1994. Photochemical production of methyl iodide in seawater. J. Geophys. Res. Atmos. 99, 16415–16420. https://doi.org/10.1029/94JD00786@10.1002/(ISSN)2169-8996.IGAC1

Mühle, J., Lueker, T.J., Su, Y., Miller, B.R., Prather, K.A., Weiss, R.F., 2007. Trace gas and particulate emissions from the 2003 southern California wildfires. J. Geophys. Res. 112, D03307. https://doi.org/10.1029/2006JD007350

Nightingale, P.D., Malin, G., Liss, P.S., 1995. Production of chloroform and other low molecular-weight halocarbons by some species of macroalgae. Limnol. Oceanogr. 40, 680–689. https://doi.org/10.4319/lo.1995.40.4.0680

Paul, C., Pohnert, G., 2011. Production and role of volatile halogenated compounds from marine algae. Nat. Prod. Rep. 28, 186–195. https://doi.org/10.1039/c0np00043d

Prinn, R.G., Weiss, R.F., Fraser, P.J., Simmonds, P.G., Cunnold, D.M., Alyea, F.N., O'Doherty, S., Salameh, P., Miller, B.R., Huang, J., Wang, R.H.J., Hartley, D.E., Harth, C., Steele, L.P., Sturrock, G., Midgley, P.M., McCulloch, A., 2000. A history of chemically and radiatively important gases in air deduced from ALE/GAGE/AGAGE. J. Geophys. Res. Atmos. 105, 17751–17792. https://doi.org/10.1029/2000JD900141

Prinn, R.G., Weiss, R.F., Arduini, J., Arnold, T., Langley Dewitt, H., Fraser, P.J., Ganesan, A.L., Gasore, J., Harth, C.M., Hermansen, O., Kim, J., Krummel, P.B., Li, S., Loh, Z.M., Lunder, C.R., Maione, M., Manning, A.J., Miller, B.R., Mitrevski, B., Mühle, J., O'Doherty, S., Park, S., Reimann, S., Rigby, M., Saito, T., Salameh, P.K., Schmidt, R., Simmonds, P.G., Paul Steele, L., Vollmer, M.K., Wang, R.H., Yao, B., Yokouchi, Y., Young, D., Zhou, L.,

2018. History of chemically and radiatively important atmospheric gases from the Advanced Global Atmospheric Gases Experiment (AGAGE). Earth Syst. Sci. Data 10, 985–1018. https://doi.org/10.5194/essd-10-985-2018

Quack, B., Peeken, I., Petrick, G., Nachtigall, K., 2007. Oceanic distribution and sources of bromoform and dibromomethane in the Mauritanian upwelling. J. Geophys. Res. 112, C10006. https://doi.org/10.1029/2006JC003803

R Core Team. 2018. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

R Studio Team. 2018. RStudio: Integrated development for R. RStudio, Inc., Boston, MA URL http://www.rstudio.com/

Raimund, S., Quack, B., Bozec, Y., Vernet, M., Rossi, V., Garç On, V., Morel, Y., Morin, P., 2010. Sources of short-lived bromocarbons in the Iberian upwelling system. Biogeosciences Discuss 7, 8663–8697. https://doi.org/10.5194/bgd-7-8663-2010

Rhew, R.C., Miller, B.R., Bill, M., Goldstein, A.H., Weiss, R.F., 2002. Environmental and biological controls on methyl halide emissions from southern California coastal salt marshes, in: Biogeochemistry. pp. 141–161. https://doi.org/10.1023/A:1019812006560

Rhew, R.C., Miller, B.R., Welss, R.F., 2000. Natural methyl bromide and methyl chloride emissions from coastal salt marshes. Nature 403, 292–295. https://doi.org/10.1038/35002043

Roy, R., 2010. Short-term variability in halocarbons in relation to phytoplankton pigments in coastal waters of the central eastern Arabian Sea. Estuar. Coast. Shelf Sci. 88, 311–321. https://doi.org/10.1016/j.ecss.2010.04.011

Scarratt, M.G., Moore, R.M., 1999. Production of chlorinated hydrocarbons and methyl iodide by the red microalga Porphyridium purpureum. Limnol. Oceanogr. 44, 703–707. https://doi.org/10.4319/lo.1999.44.3.0703

Scarratt, M.G., Moore, R.M., 1998. Production of methyl bromide and methyl chloride in laboratory cultures of marine phytoplankton II. Mar. Chem. 59, 311–320. https://doi.org/10.1016/S0304-4203(97)00092-3

Scarratt, M.G., Moore, R.M., 1996. Production of methyl chloride and methyl bromide in laboratory cultures of marine phytoplankton. Mar. Chem. 54, 263–272. https://doi.org/10.1016/0304-4203(96)00036-9

Singh, H.B., Kanakidou, M., 1993. An investigation of the atmospheric sources and sinks of methyl bromide. Geophys. Res. Lett. 20, 133–136. https://doi.org/10.1029/92GL02634

Sœmundsdóttir, S., Matrai, P.A., 1998. Biological production of methyl bromide by cultures of marine phytoplankton. Limnol. Oceanogr. 43, 81–87. https://doi.org/10.4319/lo.1998.43.1.0081

Sugihara, G., 1994. Nonlinear forecasting for the classification of natural time series. Philos. Trans. R. Soc. London. Ser. A Phys. Eng. Sci. 348, 477–495. https://doi.org/10.1098/rsta.1994.0106

Sugihara, G., Deyle, E.R., Ye, H., 2017. Misconceptions about causation with synchrony and seasonal drivers. Proc. Natl. Acad. Sci. U. S. A. https://doi.org/10.1073/pnas.1700998114

Sugihara, G., May, R., Ye, H., Hsieh, C.H., Deyle, E., Fogarty, M., Munch, S., 2012. Detecting causality in complex ecosystems. Science 338, 496–500. https://doi.org/10.1126/science.1227079

Tai, V., Palenik, B., 2009. Temporal variation of Synechococcus clades at a coastal Pacific Ocean monitoring site. ISME J. 3, 903–915. https://doi.org/10.1038/ismej.2009.35

Tait, V.K., Moore, R.M., 1995. Methyl chloride (CH_3Cl) production in phytoplankton cultures. Limnol. Oceanogr. 40, 189–195. https://doi.org/10.4319/lo.1995.40.1.0189

Team, R., 2018. R: A language and environment for statistical computing; 2015. Vienna, Austria: R Foundation for Statistical Computing.

Ulshöfer, V.S., Uher, G., Andreae, M.O., 1995. Evidence for a winter sink of atmospheric carbonyl sulfide in the northeast Atlantic Ocean. Geophys. Res. Lett. 22, 2601–2604. https://doi.org/10.1029/95GL02656

Wehrens, R., Kruisselbrink, J., 2018. Flexible Self-Organising Maps in kohonen 3.0. J. of Statistical Software 87, 1–25.

Weiss, P.S., Johnson, J.E., Gammon, R.H., Bates, T.S., 1995. Reevaluation of the open ocean source of carbonyl sulfide to the atmosphere. J. Geophys. Res. 100. https://doi.org/10.1029/95jd01926

Whelan, M.E., Min, D.H., Rhew, R.C., 2013. Salt marsh vegetation as a carbonyl sulfide (COS) source to the atmosphere. Atmos. Environ. 73, 131–137. https://doi.org/10.1016/j.atmosenv.2013.02.048

Whelan, M.E., Lennartz, S.T., Gimeno, T.E., Wehr, R., Wohlfahrt, G., Wang, Y., Kooijmans, L.M.J., Hilton, T.W., Belviso, S., Peylin, P., Commane, R., Sun, W., Chen, H., Kuai, L., Mammarella, I., Maseyk, K., Berkelhammer, M., Li, K.F., Yakir, D., Zumkehr, A., Katayama, Y., Oge, J., Spielmann, F.M., Kitz, F., Rastogi, B., Kesselmeier, J., Marshall, J., Erkkila, K.M., Wingate, L., Meredith, L.K., He, W., Bunk, R., Launois, T., Vesala, T., Schmidt, J.A., Fichot, C.G., Seibt, U., Saleska, S., Saltzman, E.S., Montzka, S.A., Berry, J.A., Elliott Campbell, J., 2018. Reviews and syntheses: Carbonyl sulfide as a multi-scale

tracer for carbon and water cycles. Biogeosciences 15, 3625–3657. https://doi.org/10.5194/bg-15-3625-2018

WMO (World Meteorological Organization). 2018. *Scientific Assessment of Ozone Depletion: 2018*, Global Ozone Research and Monitoring Project – Report No. 58, 588 pp., Geneva, Switzerland. https://www.esrl.noaa.gov/csd/assessments/ozone/2018/

Ye, H., Clark, A., Deyle, E., Sugihara, G., 2016. rEDM: an R package for empirical dynamic modeling and convergent cross-mapping.

Yokouchi, Y., Noijiri, Y., Barrie, L.A., Toom-Sauntry, D., Machlda, T., Inuzuka, Y., Akomoto, H., Li, H.J., Fujinuma, Y., Aokl, S., 2000. A strong source of methyl chloride to the atmosphere from tropical coastal land. Nature 403, 295–298. https://doi.org/10.1038/35002049

Yokouchi, Y., Saito, T., Ooki, A., Mukai, H., 2011. Diurnal and seasonal variations of iodocarbons (CH_2ClI, CH_2I_2, CH_3I, and C_2H_5I) in the marine atmosphere. J. Geophys. Res. 116, D06301. https://doi.org/10.1029/2010JD015252

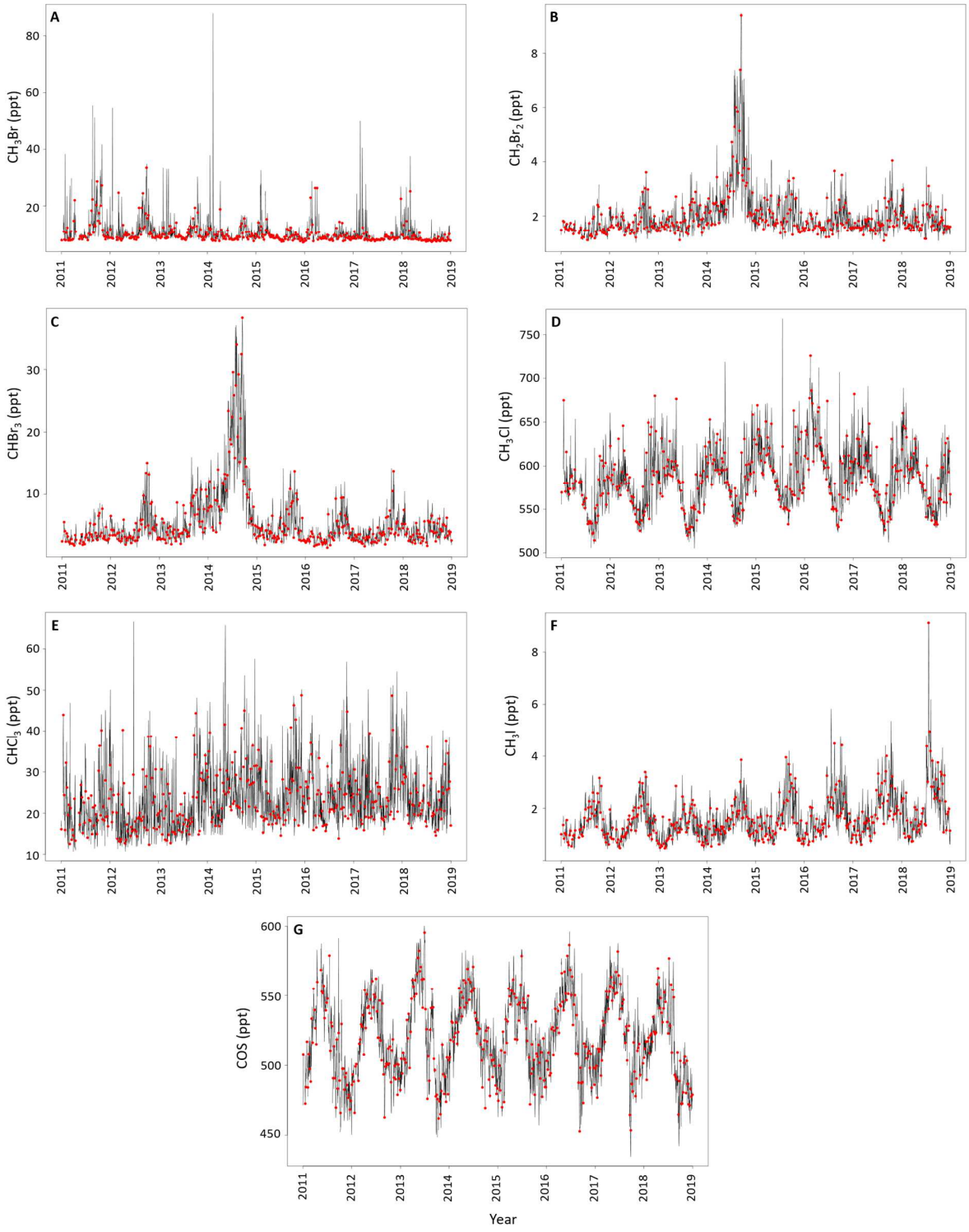

Figure 1) Daily averages of atmospheric gas mixing ratios for the halogenated volatile organic compounds A) CH_3Br, B) CH_2Br_2, C) $CHBr_3$, D) CH_3Cl, E) $CHCl_3$, F) CH_3I, and for the trace gas G) COS from 1 January 2011 to 31 December 2018 measured at the Scripps Pier. Daily

averages are visualized by the black line and dates that went into the weekly incremented models are represented as dots. (COLOR FOR ONLINE AND BLACK AND WHITE FOR PRINT, PLEASE)

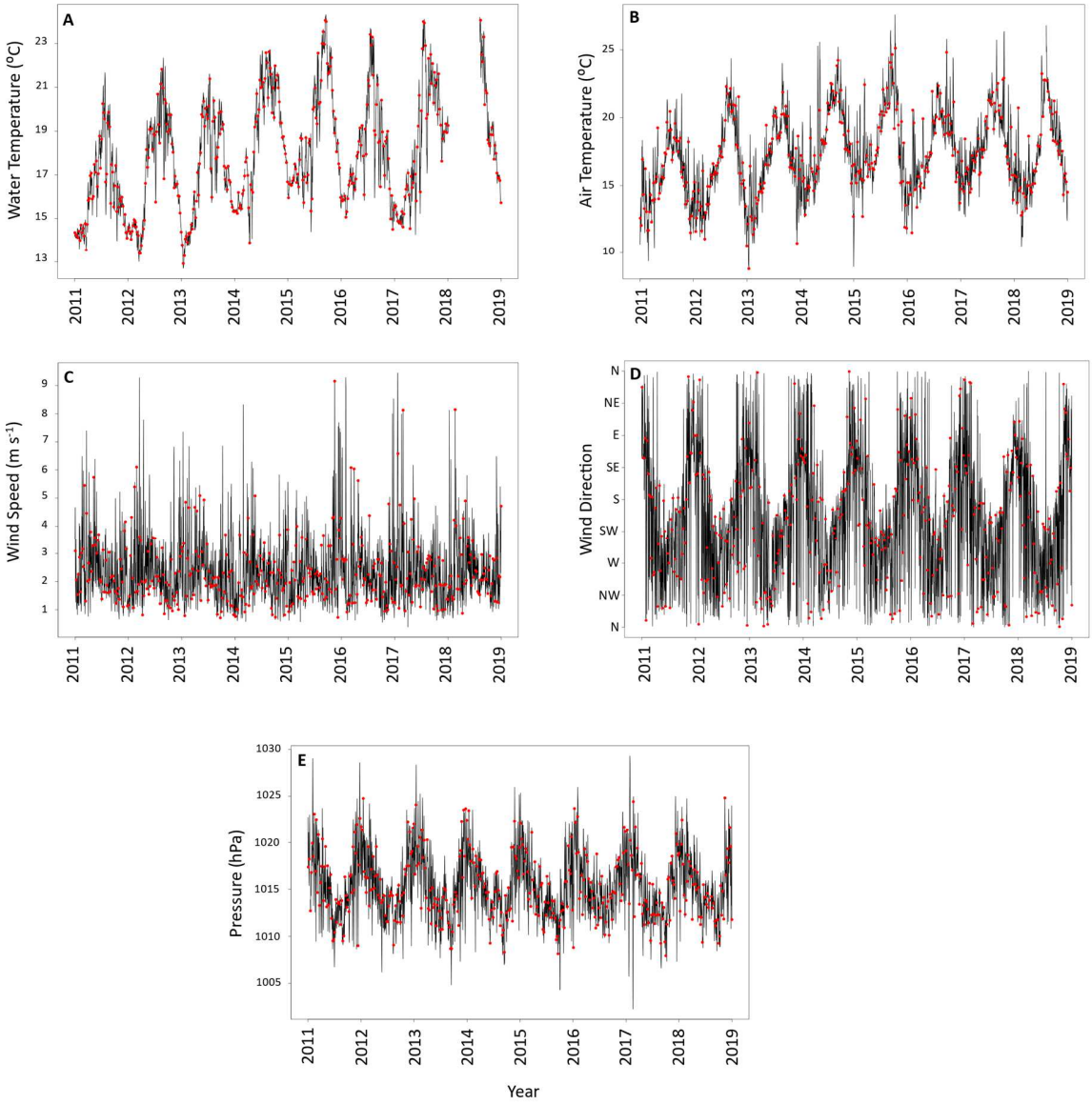

Figure 2) A) Water temperature, B) air temperature, C) wind speed, D) wind direction, and E) air pressure from 1 January 2011 to 31 December 2018 measured at the Scripps Pier. Daily averages are visualized by the black line and dates that went into the weekly incremented models are represented as dots. (COLOR FOR ONLINE AND BLACK AND WHITE FOR PRINT, PLEASE)

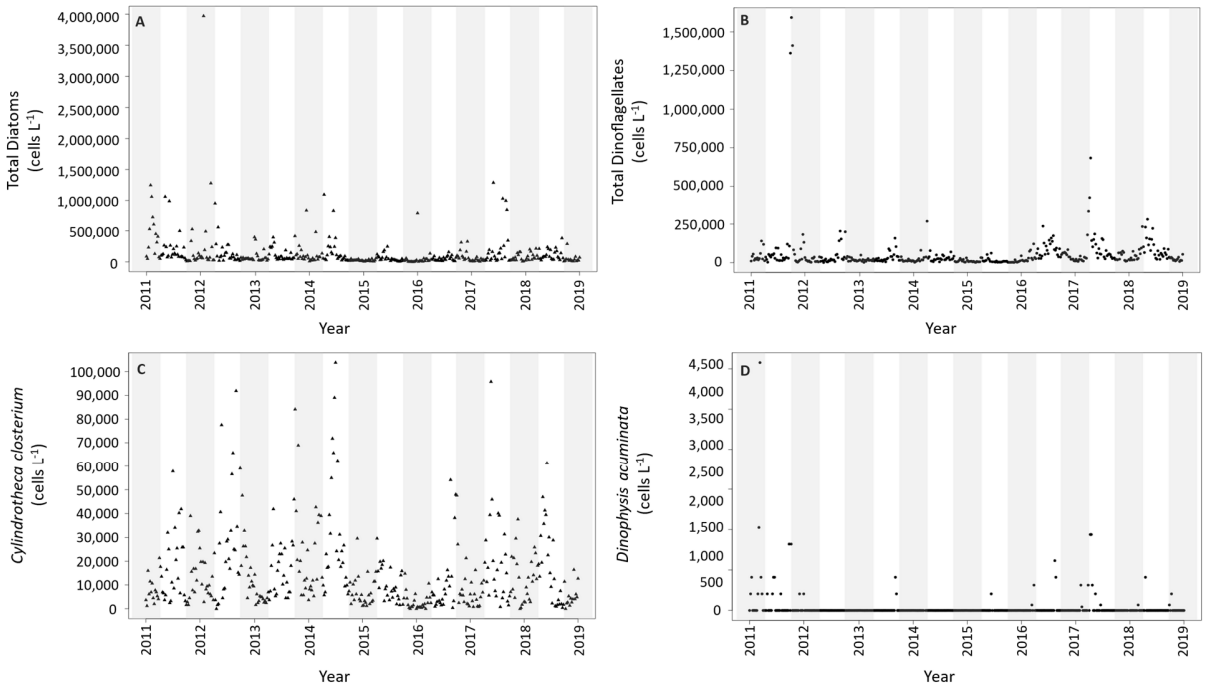

Figure 3) Total cell counts of weekly sampled A) total diatoms, B) total dinoflagellates, C) *Cylindrotheca closterium*, and D) *Dinophysis acuminata* from 1 January 2011 to 31 December 2018 sampled at the Scripps Pier. Triangles represent diatom taxa while circles represent dinoflagellate taxa. Gray bars span from October 1st through April 1st each year. (BLACK AND WHITE, PLEASE)

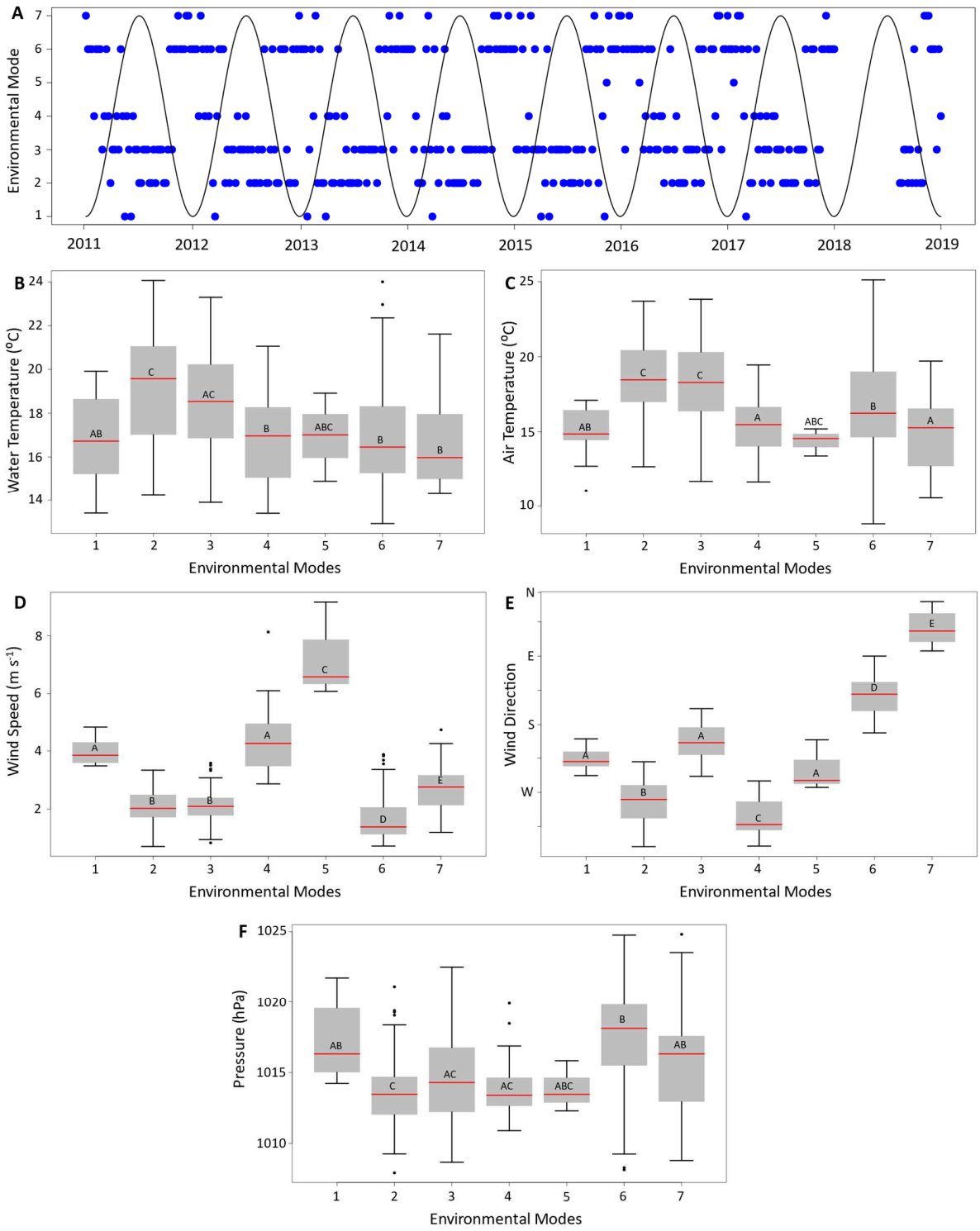

Figure 4) A) Time-series of the different Environmental Modes (EM) shown as dots, selected via a self-organizing map (a neural network algorithm useful for visualizing multidimensional data), over the sinusoidal wave that was used to represent season. Each EM factors in a scaled

(achieved by dividing each variable by the root-mean-square of its vector) version of water temperature, air temperature, wind speed, wind direction, and air pressure resulting in a single categorical variable for each date (with one date per week and a total of 387 weeks represented when all data were collected from 1 January 2011 to 31 December 2018). Box and whisker plots show EMs related to B) water temperature, C) air temperature, D) wind speed, E) wind direction (the direction from which the wind blows), and F) air pressure. Different letters denote significantly different means ($p < 0.05$) for EMs within an environmental parameter, and the same letter indicates that the means between two EMs are not statistically different (as selected via a Tukey HSD Test). For example, EM1 in 4B is statistically similar to all other EMs except for EM2 because all other EMs share either an A and/or B with EM1. For all box and whisker plots: the center line denotes the median, the gray box shows the interquartile range, and the whiskers denote the maximum and minimum values with outliers noted as dots. (BLACK AND WHITE, PLEASE)

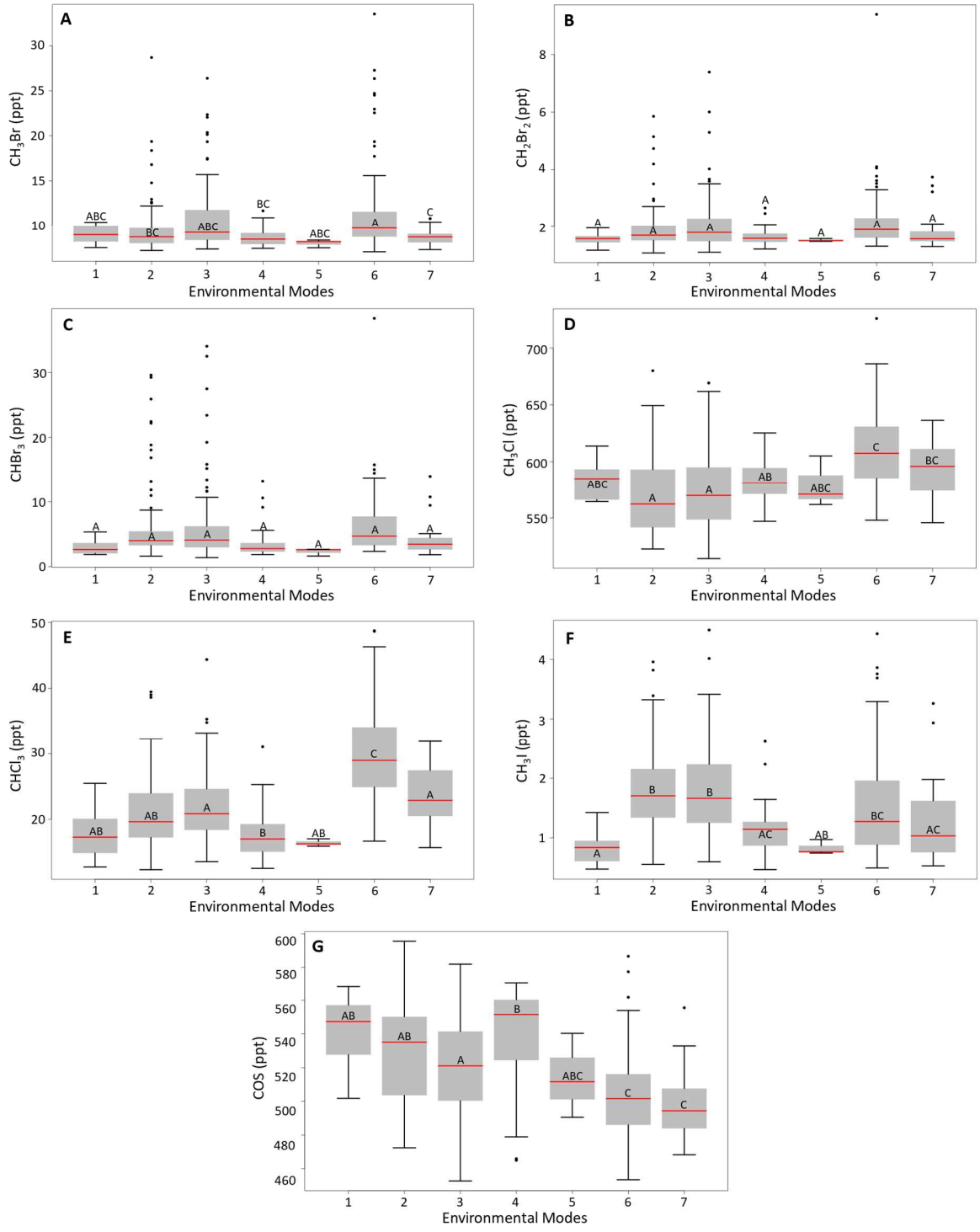

Figure 5) Environmental Modes (EM), selected via a self-organizing map (a neural network algorithm useful for visualizing multidimensional data), identical to those in Fig. 4. Each EM factors in a scaled (achieved by dividing each variable by the root-mean-square of its vector) version of water temperature, air temperature, wind speed, wind direction, and air pressure

resulting in a single categorical variable for each date (with one date per week and a total of 387 weeks represented when all data were collected from 1 January 2011 to 31 December 2018). Box and whisker plots of EMs related to A) CH_3Br, B) CH_2Br_2, C) $CHBr_3$, D) CH_3Cl, E) $CHCl_3$, F), CH_3I, and G) COS. Different letters denote significantly different means ($p < 0.05$) for EMs within a trace gas, and the same letter indicates that the means between two EMs are not statistically different (as selected via a Tukey HSD Test). For example, EM1 in 5A is statistically similar to all other EMs because all other EMs share an A, B, and/or C with EM1. For all box and whisker plots: the center line denotes the median, the gray box shows the interquartile range, and the whiskers denote the maximum and minimum values with outliers noted as dots. (BLACK AND WHITE, PLEASE)

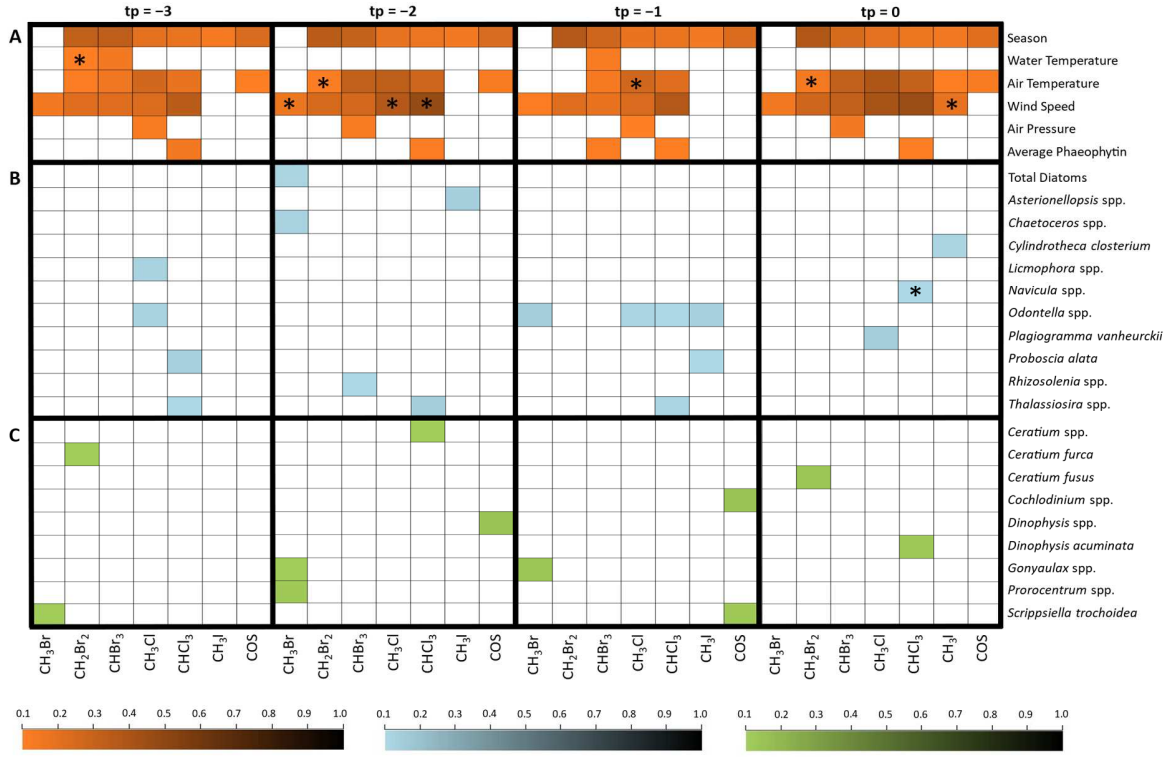

Figure 6) Heatmaps showing how the change in A) environmental parameters, B) abundances of diatom taxa, and C) abundances of dinoflagellate taxa affect changes in the range of halogenated volatile organic compounds and COS. These values were obtained via empirical dynamic modeling with convergent cross mapping and run at prediction horizons (tp) ranging from 0 to −3 weeks. 100 iterations were run at each of 41 library sizes ranging from 1 to 411 for each relationship. The shade (orange for environment, blue for diatoms, and green for dinoflagellates) denotes the maximum cross-map ability (ρ) for when the Y-axis (*e.g.*, a diatom genus or species) affected a trace gas, regardless of what library size that occurred at. Darker shades reflect a higher ρ. Only maximum ρ values above 0.1 are shown and color bars demonstrate the possible shades ranging from 0.1 to 1 (with 1 denoting complete cross-map ability). Squares with an asterisk denote relationships for which the Y-axis affecting the X-axis are not significantly different ($p < 0.05$) from the X-axis affecting the Y-axis according to a Welch's two sample T-test relating the maximum ρ values for each test (*i.e.* there is bidirectional forcing). For example, while *Navicula* spp. affected $CHCl_3$ at tp = 0, $CHCl_3$ also affected *Navicula* spp. at that prediction horizon. (COLOR, PLEASE)

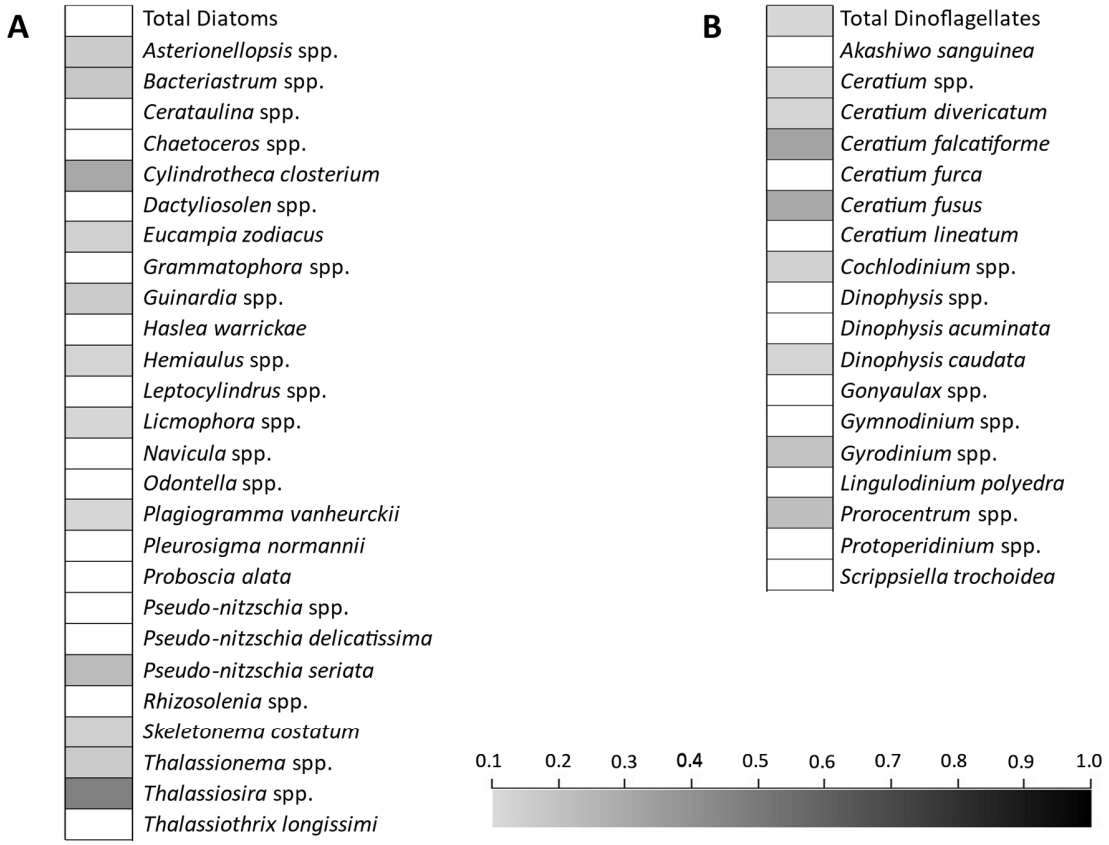

Figure 7) Seasonality for all A) diatom taxa and B) dinoflagellate taxa. These values were obtained via empirical dynamic modeling with convergent cross mapping and run at a prediction horizon (tp) of −1 week. 100 iterations were run at each of 41 library sizes ranging from 1 to 411 for each relationship. The shade denotes the maximum cross-map ability (ρ) for when seasonality affected the taxon, regardless of what library size that occurred at; a darker shade reflects higher ρ (*i.e.* displayed more seasonality). Only maximum ρ values above 0.1 are shown and the color bar demonstrates the possible shades ranging from 0.1 to 1 (with 1 denoting complete cross-map ability). All instances with a ρ above 0.1 were significantly different ($p < 0.05$) from the reverse situation according to a Welch's two sample T-test relating the maximum ρ values for each test. (BLACK AND WHITE, PLEASE)

Table 1: ρ (and p-value) from Generalized Forecasting with Mixed Embeddings

		Predicting the change in a trace gas one week (one time point) into the future using the change in a current trace gas mixing ratio + changes in:							
Trace Gas	Optimal embedding dimension	Only a trace gas's own lags	Phytoplankton + a trace gas's own lags	Phytoplankton + average phaeophytin + a trace gas's own lags	*Prorocentrum* spp. + *Chaetoceros* spp.	*Asterionellopsis* spp.	*Cylindrothecum closterium*	*Odontella* spp.	*Proboscia alata*
CH_3Cl (seasonal)	10	0.29 (0.001)	**0.32 (0.001)**	—	—	—	—	—	—
CH_3I (seasonal)	1	0.31 (0.001)	**0.34 (0.001)**	—	—	0.33 (0.001)	0.28 (0.001)	0.32 (0.001)	0.28 (0.001)
COS (seasonal)	7	0.25 (0.001)	**0.34 (0.001)**	—	—	—	—	—	—
CH_2Br_2 (aseasonal peak)	4	**0.40 (0.001)**	0.38 (0.001)	—	—	—	—	—	—
$CHBr_3$ (aseasonal peak)	8	0.47 (0.001)	0.47 (0.001)	0.44 (0.001)	—	—	—	—	—
CH_3Br	3	0.43 (0.001)	0.42 (0.001)	—	**0.45 (0.001)**	—	—	—	—
$CHCl_3$	9	0.35 (0.001)	0.47 (0.001)	**0.49 (0.001)**	—	—	—	—	—

*Bolded entries have the highest ρ value.

**The number of trace gas lags that were used were what was needed to bring the embedding dimension up to optimal.

***Phytoplankton that were used were those that were selected via CCM (see Fig. 6) at the lowest significant lag.

Table 2: Kelp Densities for La Jolla and Southern California (adapted from MBC, 2017)

Year	Kelp Density (km^2) in La Jolla, CA	Kelp Density (km^2) in Southern CA
2011	2.565	10.379
2012	1.569	11.882
2013	4.006	17.064
2014	2.790	14.053
2015	2.968	12.667
2016	0.927	5.134