# Effect of seasonal wind-driven upwelling on phytoplankton biomass in Central San Francisco Bay, 1990-2010

Daphne J. Szutu

## ABSTRACT

Changing ocean conditions will impact the intensity and strength of upwelling and ultimately affect variability in phytoplankton biomass, but because estuaries are a relatively unstudied habitat, it is uncertain how coastal upwelling affects estuarine phytoplankton biomass. I examined the connection between coastal seasonal upwelling and phytoplankton biomass in San Francisco Bay using water quality data collected monthly from the Central Bay by the United States Geological Survey (1990 to 2010). I examined water temperature, salinity, the concentration of dissolved oxygen, and chlorophyll *a* (chl *a*, a proxy for phytoplankton biomass). I separated the data into upwelling on season (May through August) and upwelling off season (November through February) to explore the seasonality of upwelling and the subsequent movement of upwelled water into the Bay. Temperature, dissolved oxygen, and chl *a* were significantly different ($p<0.05$) between the on season and off season. Of three regression models (univariate, multiple, and principle components), multiple regression was the best model for both the on season ($R^2 = 30.2\%$) and off season ($R^2 = 34.8\%$) in explaining the variation in surface chl *a* based on the physical indicators. Longitudinally, the dataset was characterized by non-constant variance and weak correlations for all variables, suggesting naturally very variable data and the presence of other factors beyond imported upwelling-induced chl *a* that may have impacted the measured chl *a* in the Central Bay. A baseline understanding of how upwelling affects estuarine phytoplankton variability will provide a basis against which to evaluate the impacts of future climate change.

## KEYWORDS

chlorophyll *a*, ocean-estuary coupling, Pacific Decadal Oscillation (PDO), gravitational circulation, longitudinal study

## INTRODUCTION

Climate change is affecting our oceans and leading to shifts in physical oceanographic conditions such as surface temperature and wind variability, bringing into question the stability of marine trophic systems, a concern for both ecological integrity and future management planning (Rost, Zondervan, & Wolf-Gladrow, 2008). For example, the base of almost all marine trophic systems is phytoplankton, the photosynthetic organisms that act as primary producers. Phytplankton serve in an ecologically critical role of converting the sun's energy and inorganic nutrients into chemical energy available to marine consumers (Hays, 2005). Consequently, shifts in phytoplankton populations will affect the rest of the ecosystem; the shifts are especially relevant in productive fisheries, which rely on phytoplankton (Hays, 2005; Brown et al., 2010). Sustaining these higher levels of trophic relationships require increases in productivity provided by phytoplankton blooms, important ecological events that consist of a rapid increase in phytoplankton growth and reproduction (Cloern & Jassby, 2008). Although the exact effects of climate change on phytoplankton blooms are unclear, changing oceanic conditions will ultimately impact phytoplankton biomass.

Upwelling is an important factor influencing phytoplankton biomass, but it is uncertain how the intensity or timing of upwelling will be impacted by changing oceanic conditions. Upwelling is a wind-driven coastal process that brings water from the deep ocean up to the surface (Kudela et al., 2008). The deeper ocean water is colder, more saline, and has lower amounts of dissolved oxygen relative to the surface (Kudela et al., 2008) and also serves to replenish the nutrient supply of the surface waters where phytoplankton exist (Martin, Fram, & Stacey, 2007). These increases in nutrients are vital for phytoplankton growth, and an upwelling event usually precedes a phytoplankton bloom. Climate change may lead to increases in greenhouse gas forcing and to wind intensification, potentially impacting the strength or frequency of upwelling (Bakun, 1990; Snyder, Sloan, Diffenbaugh, & Bell, 2003).

Upwelling, though a coastal phenomenon, also impacts marine-influenced habitats such as estuaries (Cloern & Dufford, 2005). Estuarine habitats have a wide variability in physical conditions, and as a result estuarine upwelling and its effect on phytoplankton are not well studied (Cloern, Cole, Wong, & Alpine, 1985). For example, in San Francisco Bay (SFB), phytoplankton blooms have historically occurred annually during the spring, but since 1999 there

have been annual bloom events in both spring and autumn (Cloern, Jassby, Thompson, & Hieb, 2007). This change in bloom events was caused by a regime change in the Pacific Decadal Oscillation, a multi-decadal variation in sea surface temperature (MacDonald & Case, 2005). In 1999, the Pacific Decadal Oscillation shifted to a "cold phase" for the eastern Pacific, marked by intensified southerly flows, strengthened upwelling, and a trophic cascade that reduced the bivalve population and its top-down control on phytoplankton biomass (Cloern et. al, 2007). These irregular regime changes and the accompanying impacts on the marine ecosystem underlie the importance of a long-term study on phytoplankton variability (Cloern et. al, 2007).

Understanding how upwelling affects estuarine phytoplankton variability will give a baseline that can be used to evaluate the effects of climate change in the future. Tracking the historic seasonality of physical and biological indicators of upwelled water inside the Bay can help explore how upwelling affects phytoplankton biomass. A long-term dataset of water quality inside the Bay has been produced by the United States Geological Survey (USGS), which has been continuously sampling SFB every month over the last 20 years (USGS, 2010). Upwelled water enters the bay through gravitational circulation, where denser, high salinity water tends to flow into the bay at depth while fresh water tends to flow seaward at the surface (Monismith, Kimmerer, Burau, & Stacey, 2002). Because of this phenomenon, the physical signature of upwelled water would appear near the bottom of the water column, whereas the strongest signature of phytoplankton biomass would be near the surface of the water column where phytoplankton thrive (Cloern, 1996). The effect of seasonal upwelling on phytoplankton biomass has not been studied in SFB.

In this study, I examine the relationship between seasonal upwelling and phytoplankton biomass in the Central Bay of SFB. I use USGS data on physical oceanographic variables (temperature, salinity and dissolved oxygen concentration) and a biological oceanographic variable (phytoplankton biomass), collected over the past two decades. I seek to answer the questions: (1) Are there non-biological (physical) signals of upwelling in SFB? (2) Are there biological signals of upwelling in SFB? And (3) Is there a change in the reflected seasonality after 1999? I expect that there will be non-biological signs of upwelling (lower temperature, higher salinity, and lower DO) in the bay, but the biological indicators of upwelling (increased phytoplankton biomass) will not necessarily be transported into the bay. The biological indicator

of upwelling will be less distinct because the North and South Bays are also sources of phytoplankton.

## METHODS

### Study site

SFB is an estuary system on the west coast of California, along the eastern boundary of the Pacific Ocean, bordered by the Golden Gate Bridge. SFB is composed of three embayments, South, North, and Central. The estuary system is influenced by both freshwater inputs from land and marine input from the Pacific Ocean (Cloern, 1996). The main source of freshwater is through the North Bay, and includes water collected in the Sacramento and San Joaquin Rivers. Input of marine water, primarily influenced by the tide, enters through the channel; the majority of marine input influence on phytoplankton takes place in the Central Bay (Cloern, 1996).

### Data sources

I downloaded water quality and phytoplankton datasets from a government agency internet data source, the USGS. The SFB Water Quality dataset has monthly data available from January 1990 through December 2010 for both physical water quality variables and biological phytoplankton biomass data (http://sfbay.wr.usgs.gov/access/wqdata). The water quality variables I examined were (a) water temperature, (b) salinity, and (c) dissolved oxygen (DO) (Table 1). The biological variable I studied was chl *a*, a proxy for phytoplankton biomass. All water quality and phytoplankton biomass data was collected by the USGS along a single transect, from Calaveras Point in the South Bay to the mouth of Sacramento River in the North Bay. For this study focusing on the Central Bay, I downloaded data from "Station 18" (37° 50.8'N, 122°2536'W), which is located east of Golden Gate Bridge and in the vicinity of Point Blunt. The data was collected from the surface to approximately 45 meter depth at 1-meter depth intervals.

**Table 1. Summary of oceanographic variables used in the study.** Data was downloaded from the United States Geological Survey San Francisco Water Quality database. Chlorophyll *a* is a proxy for phytoplankton biomass.

| Category | Variable (units) | Depth (m) used to calculate median |
|---|---|---|
| Physical water quality data | Temperature (°C) | 30-34 |
| | Salinity (psu) | |
| | Dissolved oxygen concentration (mg/L) | |
| Biological data | Chlorophyll *a* ($mg/m^3$) | 1-5 |

I also used the NOAA Upwelling Index to define the upwelling "on season" and "off season." I downloaded a graph of the smoothed daily NOAA upwelling index from the past 18 months (October 2009 to March 2011) at 36N latitude (see Appendix A for upwelling index, http://www.pfel.noaa.gov/products/PFEL/modeled/indices/upwelling/NA/daily_upwell_graphs.html#p10daily.gif ). The upwelling index is calculated based on Ekman's theory of mass transport by wind stress: a combination of wind parallel to the shore and the Coriolis effect from the Earth's rotation cause a net movement of water perpendicular to the shore (Pacific Fisheries Environmental Laboratory, n.d.; Mann & Lazier, 2006). The volume of upwelled water is based on six-hourly surface pressure analysis (Pacific Fisheries Environmental Laboratory, n.d.). The pressure gradient is used to approximate upwelling by calculating wind speed because wind flows down the pressure gradient, and a larger gradient indicates a higher wind speed, creating a larger wind stress. A large positive upwelling index over several days indicates a prolonged period of high wind stress and therefore the upwelling "on season", whereas a negative or zero upwelling index indicates the upwelling "off season." I defined the on season as May, June, July, and August of all years, and the off season as January, February, November, and December of all seasons.

## Data processing

I separated the 4 variables into the upwelling on and off seasons. Chl *a*, temperature, salinity, and DO measurements from May, June, July, and August from every year 1990 through 2010 was considered to be part of the upwelling on season dataset. Chl *a*, temperature, salinity, and DO measurements from January, February, November, and December from every year 1990 through 2010 was considered to be part of the upwelling off season dataset.

This study used the measurements of temperature, salinity, and DO near the bottom of the water column and the measurements of chl *a* near the surface of the water column to capture the transport of coastally upwelled water into the Central Bay. Low water temperature, high salinity, and low levels of dissolved oxygen are characteristic of deep upwelled waters (Hickey & Banas, 2003). Because of the lower temperature and higher salinity, upwelled water is denser than surface water; consequently, upwelled water would first enter the Central Bay near the bay floor before being mixed with the rest of the water column inside the Bay (Monismith et al., 2002). I used a bin of 30-34 meters to calculate the median of temperature, salinity, and dissolved oxygen (Table 1). At this depth, the measurements are still representative of bottom water while still taking into account most of the sampling dates over the 20 year sampling period (J. Cloern, personal communication, March 21, 2011). Station 18 is 45 meters deep, but the slightly shallower bin was used to calculate the medians of temperature, salinity, and DO because not every single sampling date had taken measurements to 45 meters. Out of 226 samplin dates, 21 sampling dates were not coded for temperature, salinity, and DO because the maximum depth of sampling on those dates was less than 34 meters. I calculated the median of chl *a* from a depth of 1-5 meters of the water column on most of the sampling dates (Table 1). On 14 sampling dates when sample measurements did not begin until a depth of 2 meters, I calculated the median of chl *a* using a bin of 2-5 meters (see Appendix B for all sampling dates that were not used or were used with unusual bins).

## Analysis

### *Assumption checking and transformations of chl* a

To check for functional form and constant variance, I created plots of standardized residual and fitted values for all regression models, using both the year-long, on season, and off season datasets for each variable. Because of non-constant variance in all of the physical and biological oceanographic variables in the year-long and on season datasets (as shown by the megaphone shape in the standardized residuals vs. fitted values plots), I performed a natural log transformation on the independent variable, chl *a*, for use in the year-long and on season regression models (see Appendix C for standardized residuals vs. fitted value plots for both non-

transformed and transformed chl *a*). I retained the non-transformed chl *a* in off season regression models.

*Year-long data*

**Regressions and best-fit model.** To investigate how each physical water quality variable separately affects phytoplankton biomass, I performed linear regressions in Stata 11 (StataCorp, 2009) to examine the relationship between the transformed ln(chl *a*) (the dependent variable) and each physical water quality variables (the independent variable). I performed both univariate regression and multiple regression to determine if a full model using all the physical water quality variables together could better explain the variation in chl *a* than the univariate models. I produced three individual-variable univariate regression models: (1) ln(chl *a*) with temperature, (2) ln(chl *a*) with salinity, and (3) ln(chl *a*) with DO. I created one multiple-regression model: the independent variables were all of the physical oceanographic variables (temperature, salinity, and DO) with the single dependent variable of ln(chl *a*).

Examining the four physical variables separately to pinpoint periods of upwelling can be cumbersome. To simplify the independent factors in the study system, I used Principal Component Analysis (PCA) with Stata 11(StataCorp, 2009) to create a single indicator that is a linear combination of the physical water quality variables (temperature, salinity, and dissolved oxygen) to represent upwelled water. I then performed a linear regression between the essential principal components and ln(chl *a*) to determine the proportion of the variation in ln(chl *a*) that the principal components could explain. I compared measurement of goodness of fit ($R^2$) values of the six models generated to determine the best-fit model out of four individual-variable models, one multiple regression model, and one PCA model.

*Seasonal data*

**Differences between on and off season.** To explore the movement of upwelled water into the bay, I used a 2-sample t-test to determine if the four variables (temperature, salinity, DO, and chl *a*) were significantly different between the upwelling on and off seasons.

**On and off season regression models and best-fit model.** To test if the variation in surface chl *a* could be explained by the physical indicators measured near the bottom of the water column, I also performed univariate and multiple regressions, as well as regression with principal components in Stata 11 (StataCorp, 2009). The univariate regressions produced six individual-variable models: (1) ln(chl *a*) with temperature during the on season, (2) chl *a* with temperature during the off season, (3) ln(chl *a*) with salinity during the on season, (4) chl *a* with salinity during the off season, (5) ln(chl *a*) with dissolved oxygen during the on season, and (6) chl *a* with dissolved oxygen during the off season. I created two multiple-regression models, one during the on season and one during the off seasons: (1) on season ln(chl *a*) with the physical variables during the on season, and (2) off season chl *a* with the physical variables during the off seasons.

I used PCA with Stata 11 (StataCorp, 2009) to create two indicators that represent a linear combination of the physical water quality variables to indicate upwelled water, one during the on season and one during the off season. I then performed a linear regression between the on season essential principal components with ln(chl *a*) during the on season and the off season essential principal components with chl *a* during the off season to determine proportion of the variation in chl *a* that the principal components could explain. I compared $R^2$ values of the ten models generated to determine the best-fit model: six individual-variable models, two multiple regression models, and two PCA models.

**The effect of the PDO shift.** To compare the physical and biological data before and after the shift in PDO, I divided the seasonal datasets into two periods, 1990-1998 and 1999-2010. This division was to explore if the 1999 change in annual bloom pattern Cloern et. al (2007) recorded had affected chl *a* at the surface or temperature, salinity, or DO at depth. Each variable then has four subsets: (1) on season before 1999, (2) on season after 1999, (3) off season before 1999, and (4) off season after 1999. I plotted box and whisker plots and used 2-tailed t-tests to examine if the medians between the four subsets were significantly different for each variable.

# RESULTS

## Study site

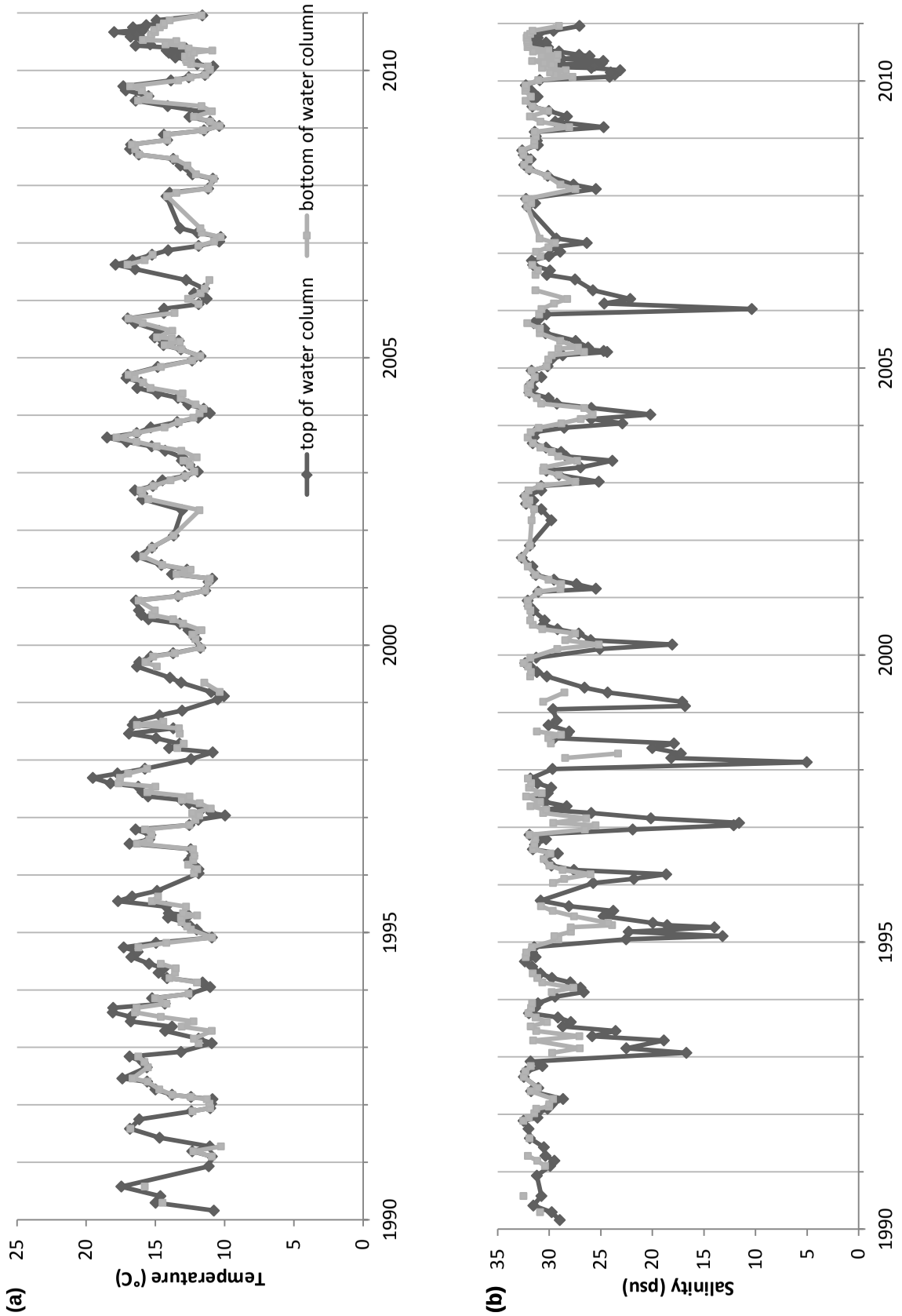
The sampling method had varied minimum and maximum depths of measurement. Over the 20-year period, the minimum depth of measurement ranged from 1 to 3 meters and the maximum depth of measurement ranged from 22 to 55 meters. Out of the total 226 sampling dates, 217 sampling dates had measurements for the 30-34 meter bin and were used to calculate the median of temperature, salinity, and DO. 76 sampling dates comprised the upwelling off season and 75 sampling dates comprised the upwelling on season.

## Longitudinal trends

I found a high level of variability for the long-term time series of each variable, although all seemed to vary annually (Table 2). Taking into account the whole water column, a water sample had median values of 13.57 °C, 31 psu, 7.6 mg DO/L and a chl *a* measurement of 3.1 $mg/m^3$ (Table 2). Temperature varied predictably on an annual scale, with dips during the winter months and peaks in the summer months of each year (Fig. 1a). Salinity was relatively constant from 1990 to 1993, but there were large dips in the median surface salinity in 1993, and annually 1995-2000, and in 2006, although the salinity of the bottom of the water column remained relatively constant (Fig. 1b). Dissolved oxygen was sampled beginning in 1993. It displayed the most variability but also seemed to follow an annual cycle, with higher values in the beginning months of a year. (Fig. 1c). Chl *a* also displayed a high level of variability, with higher peaks in the periods 1999-2003 and 2006-2011 relative to the rest of the sampling period (Fig. 1d). Compared to the top of the water column, the bottom of the water column tended to have higher chl *a*, lower water temperature, higher salinity, and lower dissolved oxygen (Fig. 1).

**Table 2. Summary of variables measured from USGS San Francisco Water Quality database.**

| Variable | Range | Median | Standard Deviation |
|---|---|---|---|
| Temperature (°C) | 9.96 - 19.53 | 13.56 | 1.94 |
| Salinity (psu) | 1.76 – 32.67 | 30.67 | 3.04 |
| DO (mg/L) | 4.2 - 10.4 | 7.60 | 0.80 |
| Chl *a* ($mg/m^3$) | 0.1 – 20.5 | 3.1 | 2.63 |


(a)
Temperature (°C)
0
5
10
15
20
25
1990
1995
2000
2005
2010
top of water column
bottom of water column
(b)
Salinity (psu)
0
5
10
15
20
25
30
35
1990
1995
2000
2005
2010

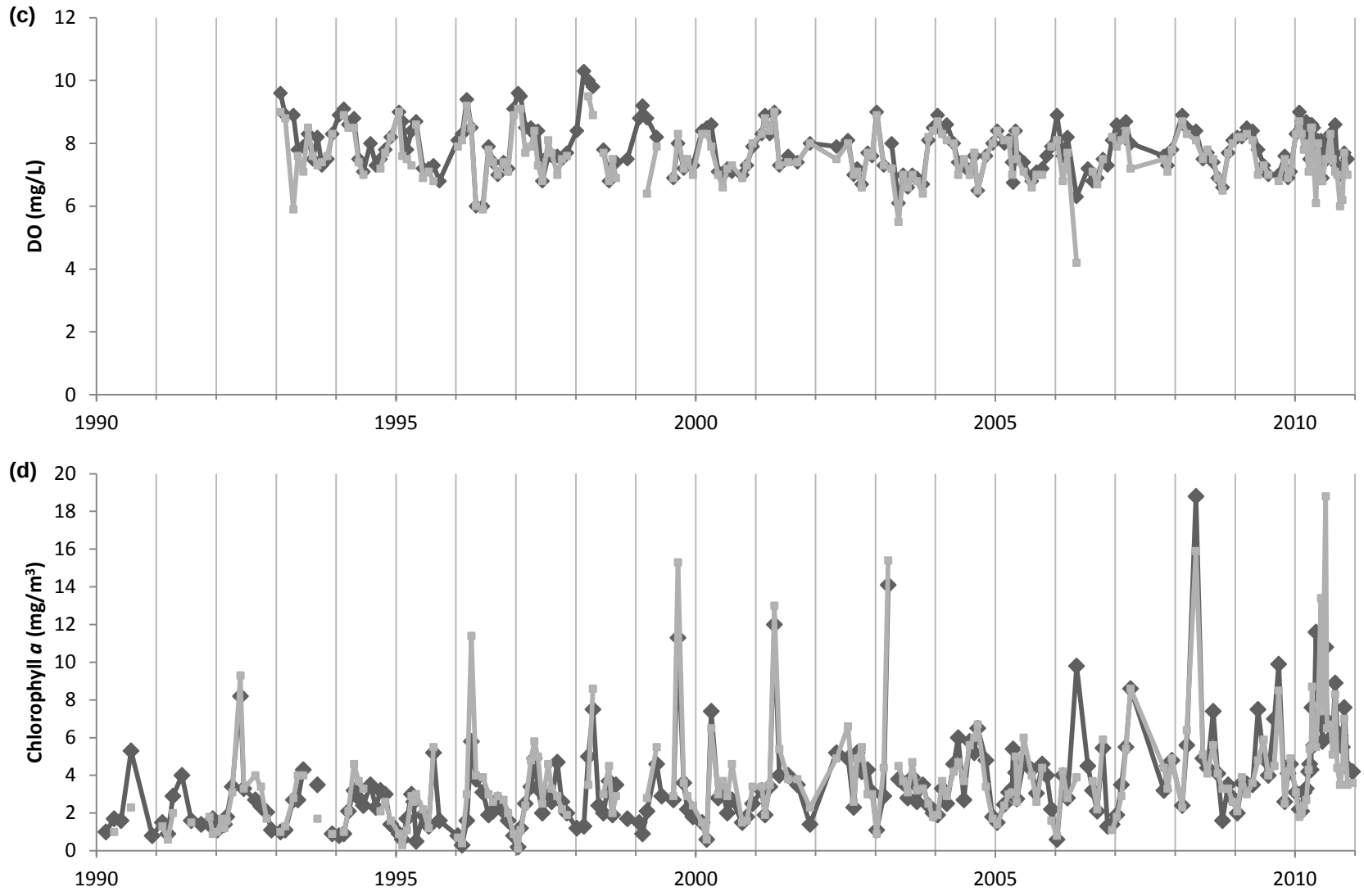



**Figure 1. Time series of (a) temperature, (b) salinity, (c) dissolved oxygen concentration, and (d) chlorophyll *a*, from 1990 to 2010.** The lighter gray line represents the top of the water column (calculated as median of meters 1-5), and the darker gray line represents the bottom of the water column (calculated as median of meters 30-34).

## Year-long data

*Regressions models*

**Univariate regression model.** After log transformation of the outcome variable, I found that salinity and DO were significant in the univariate regression model (Table 3), rejecting the null hypothesis that changing the value of salinity and DO have no impact on ln(chl *a*). $R^2$ values ranged from approximately 0.014 to 0.042 for the top of the water column, indicating that the individual variables predicted 1.4% -4.2% of the variability in ln(chl *a*) when looked at individually (Table 3).

**Table 3. Individual regression models for ln(chl *a*).** * denotes a significant p-value (p<0.05), ** denotes a very significant p-value (p<0.01)

| Models | Coefficient | P-value | $R^2$ |
|---|---|---|---|
| Temperature | 0.1646 | 0.096 | 0.014 |
| Salinity | 0.2353 | 0.016* | 0.029 |
| DO | -0.6478 | 0.006** | 0.042 |

**Multiple regression model.** None of the explanatory variables, temperature, salinity, or DO, were significant in the multiple regression model, with all variables having p-values > 0.5 (Table 4). I could not reject the null hypothesis and concluded that there was no relationship between ln(chl *a*) at the surface and the physical variables near the bottom of the water column. When examining year-long data, the $R^2$ value was 0.057, explaining 5.7% of the variation in ln(chl *a*).

**Table 4. Multiple regression model for ln(chl *a*).** $R^2 = 0.057$

| Explanatory Variable | Coefficient | P-Value |
|---|---|---|
| Temperature | 0.0109 | 0.922 |
| Salinity | 0.1834 | 0.109 |
| DO | -0.4582 | 0.084 |

**PCA and regression with principal components.** In examining year-long temperature, salinity, and DO, I found two essential (useful) principal components. The two essential principal components together explained 80.7% of the variation in the model (see Appendix D,

Table D1 and Equations D1, D2 for PCA details). Both components, pc1 and pc2, were highly statistically significant, with p-values <0.0005 (Table 5). Taking pc1 and pc1 together, the $R^2$ was 0.150, explaining 15.0%of the variance in the ln(chl *a*) (Table 5).

**Table 5. Regression with principal components for ln(chl *a*).** $R^2$ = 0.150, *** denotes a very significant p-value (p<0.01)

| Explanatory Variable | Coefficient | P-Value |
|---|---|---|
| pc1 | 0.1904 | <0.0005*** |
| pc2 | -0.0847 | 0.139 |

*The best-fit model*

The best-fit models for describing the relationship between the physical and phytoplankton datasets using the criteria of $R^2$ was the regression model with principle components using the log-transformed chl *a* variable. The PCA model had the highest $R^2$ value of 0.150 (Table 5) calculated relative to the other models examining year-long data: $R^2$ values of 0.014 to 0.042 for the univariate regression models and $R^2$ values of 0.057 for the multiple regression model (Table 3, 4).

## Seasonal data

*Differences between on and off seasons*

Temperature, DO, and chl *a* values were significantly different between the on season and the off season, with p-values 8 to 69 magnitudes of order smaller than the significance level of p=0.05 (Table 6). Salinity was not significantly different between the on and off season, with a p-value of 0.215 (Table 6). During the on season, temperature was 2.81°C higher, salinity was 0.15 psu higher, DO was 0.8 mg/L lower, and chl *a* concentration was 1.9 mg/$m^3$ higher relative to the off season (Table 6).

**Table 6. Differences in temperature, salinity, DO, and chl *a* between upwelling on and off seasons**. P-values calculated from 2-tailed t-tests to test for significant differences between the on and off season, *** denotes highly significant p-values (p<0.001)

| Variable | Median | | P-value |
|---|---|---|---|
| | On season | Off season | |
| Temperature (°C) | 14.93 | 12.11 | 3.05E-71*** |
| Salinity (psu) | 31.15 | 31 | 0.215 |
| DO (mg/L) | 7.3 | 8.1 | 2.56E-11*** |
| Chl *a* ($mg/m^3$) | 3.7 | 1.8 | 1.37E-37*** |

*Regressions models*

**On and off season univariate regression models.** I found that the only significant relationship for ln(chl a) during the on season for the individual regression model was with salinity (p-value=0.001), but models were not significant for temperature or DO (Table 7). Salinity (p-value=0.012) and DO (p-value<0.0005) were significant in the individual regression model during the off season with chl *a*, but temperature was not significant (p-value = 0.0410) (Table 7). $R^2$ values ranged from approximately 0.028 to 0.150 during the on season, indicating that the individual variables predicted 2.8%-15.0% of the variability in chl *a* when examined individually (Table 8). During the off season, I found that the $R^2$ values ranged from 0.011 to 0.328, predicting 1.1% to 32.8% of chl *a* variability (Table 8).

**Table 7. Individual regression models for on season ln(chl *a*).** ** denotes a very significant p-value (p<0.01)

| Models | P-value | $R^2$ |
|---|---|---|
| Temperature | 0.061 | 0.055 |
| Salinity | 0.001** | 0.150 |
| DO | 0.199 | 0.028 |

**Table 8. Individual regression models for off season chl *a*.** * denotes a significant p-value (p<0.05), *** denotes a highly significant p-value (p<0.001)

| Models | P-value | $R^2$ |
|---|---|---|
| Temperature | 0.410 | 0.011 |
| Salinity | 0.012* | 0.099 |
| DO | <0.0005*** | 0.328 |

**On and off season multiple regression models.** In comparing on and off season multiple regression models for chl *a*, the explanatory variables showed differences in significance.

During the on season, temperature and salinity were highly significant with p-values less than 0.01 (Table 9). While DO was highly significant in the off season multiple regression model, the other variables were not significant (Table 10). The $R^2$ value indicated that 30.2% of variation of ln(chl *a*) during the on season was accounted for by temperature, salinity and DO in the multiple regression model (Table 9). During the off season, the $R^2$ value was higher, indicating that the three physical variables together explained about 34.8% of the variance in chl *a* (Table 10).

**Table 9. Multiple regression model for on season ln(chl *a*).** $R^2$ = 0.302, ** denotes a very significant p-value (p<0.01), *** denotes a highly significant p-value (p<0.001)

| Explanatory Variable | Coefficient | P-Value |
|---|---|---|
| Temperature | -0.1251 | 0.004** |
| Salinity | 0.1780 | <0.0005*** |
| DO | -0.0256 | 0.793 |

**Table 10. Multiple regression model for off season chl *a*.** $R^2$ = 0.348, *** denotes a highly significant p-value (p<0.001)

| Explanatory Variable | Coefficient | P-Value |
|---|---|---|
| Temperature | -.1286 | 0.360 |
| Salinity | .0939 | 0.340 |
| DO | -1.082 | <0.0005*** |

**PCA and regression with principal components.** For both on and off season datasets, I found two essential principal components. Regressions with ln(chl *a*) during the on season and with chl *a* during the off season were not statistically significant (p-value>0.05) (Table 11, 12). During the on season, the two essential principal components together explained 17.2% of the variation in the model (Table 11; see Appendix D, Table D2 and Equations D3, D4 for PCA details). During the off season, the two essential principal components together explained 31.2% of the variation (Table 12; see Appendix D, Table D3 and Equations D5, D6 for PCA details).

**Table 11. Regression with principal components for on season ln(chl *a*).** $R^2$ = 0.172, ** denotes a very significant p-value (p<0.01)

| Explanatory Variable | Coefficient | P-Value |
|---|---|---|
| $pc1_{on}$ | -0.0417 | 0.496 |
| $pc2_{on}$ | 0.2287 | 0.001*** |

**Table 12. Regression with principal components for off season chl *a*.** $R^2 = 0.312$, * denotes a significant p-value ($p<0.05$), *** denotes a highly significant p-value ($p<0.001$)

| Explanatory Variable | Coefficient | P-Value |
|---|---|---|
| $pc1_{off}$ | 0.4365 | <0.0005*** |
| $pc2_{off}$ | -0.3986 | 0.023* |

*Effect of 1999 PDO shift*

Salinity, DO, and chl *a* all showed significant difference before and after the 1999 shift in PDO, with chl *a* concentration showing the largest change before and after 1999. Year-long temperature was lower, salinity was higher, DO was higher, and chl *a* was higher in the post-shift period (Table 13). On season salinity, temperature, and chl *a* were slightly higher than off season salinity and temperature both before and after 1999 (Fig. 2a, b, d). Chl *a* concentration showed much more variability above the median during the on season after 1999 (Fig. 2d). The off season DO was higher than on season DO both before and after 1999 (Fig. 2c). Between the two time periods (1990-1998 and 1999-2010), salinity, DO, and chl *a* were significantly different taking into account year-long and off season data, but only chl *a* was significantly different during the on season when comparing the two time periods (Table 14).

**Table 13. Medians of temperature, salinity, DO, and chl *a*.** Medians are calculated separately using year-long , on season, and off season datasets during the two time periods, 1990-1998 and 1999-2010.

| Variable | 1990-1998 median values | | | 1999-2010 median values | | |
|---|---|---|---|---|---|---|
| | Year-long | On season | Off season | Year-long | On season | Off season |
| Temperature (°C) | 12.79 | 14.92 | 12.21 | 13.1 | 14.93 | 11.9 |
| Salinity (psu) | 30.72 | 30.82 | 30.02 | 31.22 | 31.33 | 31.03 |
| DO (mg/L) | 7.7 | 7.4 | 8.25 | 7.55 | 7.25 | 8 |
| Chl *a* ($mg/m^3$) | 1.7 | 2.6 | 1.1 | 3.4 | 4.45 | 2.4 |

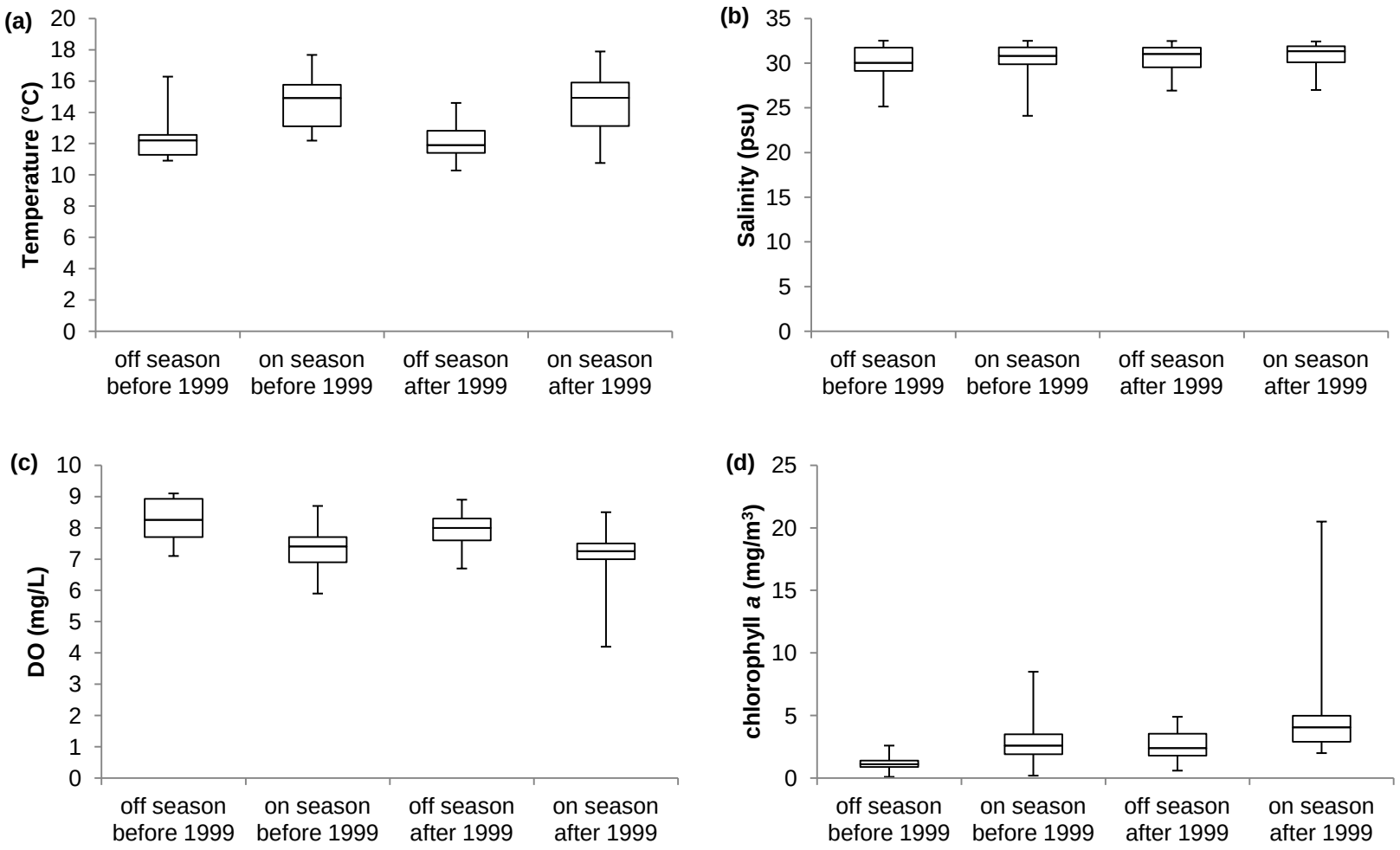



**Figure 2. Difference between variables for on and off season before and after 1999.** Box-and-whisker plot comparisons of the four variables between the off and on season in two time periods, 1990-1998 and 1999-2010: (a) temperature, (b) salinity, (c) dissolved oxygen concentration, and (d) chlorophyll *a*. The box indicates the middle 50% (between the 1st and 3rd quartile) of the data, the top whisker indicates the upper 25% of the data, and the bottom whisker indicates the lowest 25% of the data. Outliers (values great than 1.5 times the interquartile range above the median or less than 1.5 times the interquartile range below the median) are included in the whiskers.

**Table 14. Two-tailed t-tests comparing the periods 1990-1998 and 1999-2010.** P-values calculated for the variables temperature, salinity, DO, and chl *a*. ** denotes very significant p-value (p<0.01), *** denotes highly significant p-values (p<0.001)

| Variable | Year-long data | On season | Off season |
|---|---|---|---|
| Temperature (°C) | 0.128 | 0.545 | 0.16 |
| Salinity (psu) | 0.001*** | 0.08 | 0.005** |
| DO (mg/L) | 0.009** | 0.113 | 9.11E-6*** |
| Chl *a* ($mg/m^3$) | 6.12E-33*** | 3.52E-20*** | 3.48E-46*** |

## DISCUSSION

The objective of my study was to determine if there was a link between coastally upwelled water and the estuary of SFB by examining the seasonality of both physical and biological indicators of upwelled water inside the Bay. Because upwelling supplies essential nutrients to phytoplankton, understanding the effect of upwelling on San Francisco Bay phytoplankton biomass is helpful in modeling and planning for potential changes in the phytoplankton population and the rest of the food web. Chl *a*, measured at the surface, and water temperature and dissolved oxygen, measured near the bottom of the water column, were significantly different between the on and off upwelling seasons, but salinity, measured near the bottom of the water column, was not significantly different between seasons. During the upwelling on season, 30.2% of the variation in chl *a* could be explained by the variation in the physical factors, and during the upwelling off season, 34.8% of the variation in chl *a* could be explained.

### Seasonal trends

During the upwelling on season, there were non-biological indicators of upwelled water of lower water temperatures and lower dissolved oxygen in the bay bottom water, suggesting that over the 20 years of study in this dataset the bay is influenced by seasonal upwelling. Low temperature, high salinity, and low dissolved oxygen are indicative of deep ocean water (Kudela et al., 2008) and their presence at Station 18 appeared during the upwelling on season of May, June, July, and August of 1990 to 2010. The coherence of physical signatures of upwelled

waters appearing inside the bay during the coastal upwelling season confirms that the bay is connected with the marine system outside of the bay.

Although the upwelled waters did account for some of the variability in phytoplankton, there are several potential reasons why the biological and physical indicators of upwelled water did not have a higher association. For example, a high phytoplankton biomass measurement in the Central Bay could have originated from the a bloom event in the North or South Bays that was then transported to the Central Bay (Cloern et al., 1985). In addition, a complicated sequence of events was necessary for an upwelling-produced phytoplankton bloom to travel from the coastal waters into the Central Bay and may not always occur. The wind needed to blow strongly from the north for five to six days to induce an upwelling event; then, a reversal of wind direction was necessary to promote water moving toward the coast and into the Bay (Roegner, Hickey, Newton, Shanks, &Armstrong, 2002; J. Cloern, personal communication, March 21, 2011). The direct biological indicators of upwelling would be detected in our dataset only after the specific order of events and the appropriate phytoplankton bloom timing, which is four to ten days (J. Cloern, personal communication, March 21, 2011). This timing allows for phytoplankton to bloom after an upwelling event and for the elevated chl *a* signal to be transported into the bay and distinguished in our dataset (J. Cloern, personal communication, March 21, 2011).

**Longitudinal trends**

The dataset displayed non-constant variance and weak correlations for all variables, suggesting that the environment and phytoplankton population biomass are naturally very variable and that other factors other than the import of coastally upwelled-induced phytoplankton biomass could have impacted the measured chl *a* at Station 18. One of the main factors is seasonality, which, though predictable, added variability to the physical and biological variables. The effect of seasonality can be seen in the increase of $R^2$ values. After separating the dataset into the upwelling on and off season, the variation in surface chl *a* explained by the variation in physical variables increased from 5.7% (considering data from the entire year) to 30.2% (on season data) and 34.8% (off season data). Besides seasonality, other factors contributing to how little of the variability in chl *a* was explained by the physical factors were (1) independent

variables not included in the analysis and (2) patchy distribution of phytoplankton. An independent variable that could be taken into account is suspended particulate matter, a measure of turbidity. Phytoplankton in SFB are generally light limited (Cloern et al., 1985; Dugdale, Wilkerson, Hogue, & Marchi, 2007), so adding turbidity to the regression models may yield higher $R^2$ values. Phytoplankton biomass, as a biological phenomenon, is spatially patchy with mesoscale variability, especially during upwelling events (Abbott & Zion, 1985). This day-to-day variability partly results from weather events such as rain or wind events and fluctuations in tidal mixing and partly from biological processes such as grazing of phytoplankton by zooplankton (Cloern, 1996; Lehman, 2000). Addressing some of this spatial variability by collecting data from more sites would be helpful in future studies.

There was a high level of variability for the long-term time series of each variable, implying that the system is naturally very patchy temporally and spatially. Some of the temporal patterns can be explained: the predictability of the annual variability in temperature in the water column is explained by the annual patterns in solar irradiance (Thompson, Baird, Ingleton, & Doblin, 2009). Chl *a* also varied annually, although less obviously so, suggesting that chl *a* was driven by more than physical processes – namely, the biological phenomena of phytoplankton blooms (Letelier et al., 1993; Cloern, 2006). Dissolved oxygen was the most variable over an annual scale relative to salinity and temperature, implying that dissolved oxygen concentration was driven mainly by respiration of marine organisms, a biological phenomenon that is naturally more spatially variable compared to the physical phenomenon of solar irradiation and wind stress that drives water temperature and salinity (Serret, Robinson, Fernández, Teira, & Tilstone, 2001).

Comparing the physical and biological variables during the upwelling on season and off season can indicate when oceanic water is entering the bay. Water measurements at the bottom of the water column during the upwelling season reflected oceanic water entering the bay, with statistically significant lower temperatures and lower dissolved oxygen relative to the top of the water column. Interestingly, the bottom of the water column also tended to have higher chl *a* level during both the upwelling on and off seasons, which could indicate the end of a bloom period (J. Cloern, personal communication, March 21, 2011). At the end of a bloom period, the phytoplankton die and sink to the bottom of the water column, thus producing a higher chl *a* measurement at the bottom of the water column compared to the top of the water column (J.

Cloern, personal communication, March 21, 2011). The reflected seasonality found at Station 18 of SFB suggests there is a link between coastal upwelling and estuarine phytoplankton biomass.

**Limitations**

To best understand upwelling, a sampling regime would need to sample on the scale of every three to four days. The monthly USGS sampling regime is designed for long-term characterization of mesoscale spatial variability along the entire estuary of SFB, on a time scale of weeks to years (Cloern, 1996). Twenty years of data is very useful in looking at long-term trends, but having only monthly sampling frequency at one site fails to capture the spatial and temporal patchiness of phytoplankton, temperature, salinity and DO. The sampling frequency is especially important when detecting the import of coastal phytoplankton blooms, which occur approximately four to ten days after an upwelling event (J. Cloern, personal communication, March 21, 2011). Other factors that impacted the variables may be more difficult to quantify, such as bathymetry, the surface features of the ocean floor, which influences the flow of water along the bottom of the bay floor, and the diurnal tides of SFB that change stratification and manipulate phytoplankton community dynamics (Cloern et al., 1985). Finally, I only focused on one sampling point for this study, Station 18. Although this station is closest to Golden Gate Bridge and therefore experiences the most influence from marine waters, using data from only one station is unlikely to be representative of the spatial variability of phytoplankton.

**Future Directions**

To address some of these limitations, future studies may include more frequent sampling regimes, as well as examining data from more than one sampling station. A more frequent sampling regime will help account for the time lag between an upwelling event and the phytoplankton bloom. In conjunction with data from additional stations, a more frequent sampling regime will help distinguish an influx of chl *a* from the coast, as compared to an influx of chl *a* from the North or South Bays. From two previous studies on the effect of upwelling on phytoplankton, phytoplankton biomass tended to increase four to six days after upwelling subsides; a more frequent sampling regime would follow a similar schedule to clearly show a

connection between an upwelling event along the coast and the movement of chl *a* moving from the coast into the bay (Palma, Mouriño, Silva, Barão, & Moita, 2005; Vahtera, Laanemets, Pavelson, Huttunen, & Kononen, 2005). Incorporating into the analysis more sites with more frequent sampling, but an overall shorter time-scale, will increase spatial resolution. Using data from at least 2 other sites, one site representative of the North Bay and one site representative of the South Bay, would help differentiate blooms that occur in the North, South, and Central Bays of SFB. Finally, using additional variables such as turbidity could potentially make the model more biologically meaningful (Cloern et al., 1985).

**Broader Implications**

Understanding how upwelling affects phytoplankton variability will give a baseline understanding to evaluate the impacts of climate change in the future. Climate change can affect the strength and timing of upwelling, and the changes in intensity and timing of upwelling-impacted nutrient and carbon fluxes in marine environments (Bakun, 1990). During the upwelling season, the pressure gradient between a warmer land mass and a cooler body of water maintains a coastal wind stress necessary to induce upwelling (Bakun, 1990). An increase in atmospheric carbon dioxide could lead to increased temperatures over land, thus increasing the pressure gradient between land and water (Bakun, 1990). The resulting intensification in wind stress will accelerate upwelling, and as a positive feedback could reduce the surface temperature of the ocean, further increasing the pressure gradient (Bakun, 1990). There are three way that changes in upwelling will impact marine ecosystems, through (1) providing a food and nutrient supply, (2) supporting a minimum concentration of food to sustain a population, and (3) retention of food supply and organisms in the same area (Snyder et al., 2003). Intensified upwelling would increase nutrient resupply from the deep ocean, but the increased wind stress would lead to more mixing, decreasing the concentration of food and scattering organisms spatially (Snyder et al., 2003). A delay in upwelling can lead to temporal mismatches among trophic levels, impacting fish populations and fisheries operations (Barth et al., 2007). Additional research exploring connection between phytoplankton biomass and seasonal upwelling will help develop estuarine and marine management to maintain ecological and economic integrity along the coasts.

## ACKNOWLEDGEMENTS

Patina Mendez, Kurt Spreyer, Lara Roman, and Seth Shonkoff comprised Team ES196, and their continual dedication and energy was essential in my completing this project. I would especially like to thank Patina Mendez for her enthusiasm and thoughtful feedback during office hours and through e-mail correspondence over the past year and a half. My subject matter support came from the Menlo Park USGS office. Dr. James Cloern, Valerie Greene, and Tara Schraga greatly helped me mold my project, answered my numerous questions, and gave me the opportunity to participate on the 11 March 2011 South Bay sampling cruise. Professor Zack Powell of UC Berkeley Integrative Biology cultivated my enthusiasm in oceanography and connected me with Dr. Cloern. Dr. Maureen Lahiff of UC Berkeley School of Public Health was instrumental in helping me with my statistical analysis and graciously mentoring me even after I was no longer a student in her class. Finally, I greatly benefitted from the energy and peer edit support from members of my Environmental Sciences cohort: Sarah Jarjour, Sophie You, Jim Gao, and Michael Young.

## REFERENCES

Abbott, M. R., & Zion, P. M. (1985). Satellite observations of phytoplankton variability during an upwelling event. *Continental Shelf Research, 4*(6), 661-680.

Bakun, A. (1990). Global climate change and intensification of coastal ocean upwelling. *Science, 247*, 198 -201.

Barth, J. A., Menge, B. A., Lubchenco, J., Chan, F., Bane, J. M., Kirincich, A. R., McManus, M. A., Nielsen, K.J., Pierce, S.D. & Washburn, L. (2007). Delayed upwelling alters nearshore coastal ocean ecosystems in the northern California current. *Proceedings of the National Academy of Sciences, 104*(10), 3719 -3724.

Brown, C. J., Fulton, E. A., Hobday, A. J., Matear, R. J., Possingham, H. P., Bulman, C., Christensen, V., Forrest, R.E., Gehrke, P.C., Gribble, N.A., Griffiths, S.P., Lozano-Montes, H., Martin, J.M., Metcalf, S., Okey, T.A., Watson, & Richardson, A.J. (2010). Effects of climate-driven primary production change on marine food webs: implications for fisheries and conservation. *Global Change Biology, 16*(4), 1194 -1212.

Cloern, J. E. (1996). Phytoplankton bloom dynamics in coastal ecosystems: A review with some general lessons from sustained investigation of San Francisco Bay, California. *Reviews of Geophysics, 34*(2), PP. 127 -168.

Cloern, J. E., Cole, B. E., Wong, R. L. J., & Alpine, A. E. (1985). Temporal dynamics of estuarine phytoplankton: A case study of San Francisco Bay. *Hydrobiologia*, *129*(1), 153 -176.

Cloern, J. E., & Dufford, R. (2005). Phytoplankton community ecology: principles applied in San Francisco Bay. *Marine Ecology Progress Series*, *285*, 11-28.

Cloern, J. E., Jassby, A. D., Thompson, J. K., & Hieb, K. A. (2007). A cold phase of the East Pacific triggers new phytoplankton blooms in San Francisco Bay. *Proceedings of the National Academy of Sciences*, *104*(47), 18561 -18565.

Hays, G. C., Richardson, A. J., & Robinson, C. (2005). Climate change and marine plankton. *Trends in Ecology & Evolution*, *20*(6), 337-344.

Kudela, R. M., Banas, N. S., Barth, J. A., Frame, E. R., Jay, D., Largier, J.L., Lessard, E.J., Peterson, T.D., & Vander Woude, A.J. (2008). New Insights into the Controls and Mechanisms of Plankton Productivity in Coastal Upwelling Waters of the Northern California Current System. *Oceanography*, *21*(4), 46-52.

Lehman, P. W. (2000). The Influence of Climate on Phytoplankton Community Biomass in San Francisco Bay Estuary. *Limnology and Oceanography*, *45*(3), 580-590.

Letelier, R. M., Bidigare, R. R., Hebel, D. V., Ondrusek, M., Winn, C. D., & Karl, D. M. (1993). Temporal Variability of Phytoplankton Community Structure Based on Pigment Analysis. *Limnology and Oceanography*, *38*(7), 1420-1437.

MacDonald, G. M., & Case, R. A. (2005). Variations in the Pacific Decadal Oscillation over the past millennium. *Geophysical Research Letters*, *32*, 4 PP.

Mann, K. H., & Lazier, J. R. N. (2006). *Dynamics of marine ecosystems: biological-physical interactions in the oceans*. Wiley-Blackwell.

Martin, M. A., Fram, J. P., & Stacey, M. T. (2007). Seasonal chlorophyll a fluxes between the coastal Pacific Ocean and San Francisco Bay. *Marine Ecology Progress Series*, *337*, 51-61.

Monismith, S. G., Kimmerer, W., Burau, J. R., & Stacey, M. T. (2002). Structure and Flow-Induced Variability of the Subtidal Salinity Field in Northern San Francisco Bay. *Journal of Physical Oceanography*, *32*(11), 3003-3019.

Pacific Fisheries Environmental Laboratory. (n.d.).*Upwelling Indices*. Retrieved from http:// http://www.pfeg.noaa.gov/products/PFEL/modeled/indices/upwelling/upwelling.html.

Palma, S., Mouriño, H., Silva, A., Barão, M. I., & Moita, M. T. (2010). Can Pseudo-nitzschia blooms be modeled by coastal upwelling in Lisbon Bay? *Harmful Algae*, *9*(3), 294-303.

Roegner, G. C., Hickey, B. M., Newton, J. A., Shanks, A. L., & Armstrong, D. A. (2002). Wind-Induced Plume and Bloom Intrusions into Willapa Bay, Washington. *Limnology and Oceanography*, *47*(4), 1033-1042.

Rost, B., Zondervan, I., & Wolf-Gladrow, D. (2008). Sensitivity of phytoplankton to future changes in ocean carbonate chemistry: current knowledge, contradictions and research directions. *Marine Ecology Progress Series*, *373*, 227-237.

StataCorp. (2009), *Stata Statistical Software: Release 11*. College Station, TX: StataCorp LP.

Serret, P., Robinson, C., Fernández, E., Teira, E., & Tilstone, G. (2001). Latitudinal Variation of the Balance between Plankton Photosynthesis and Respiration in the Eastern Atlantic Ocean. *Limnology and Oceanography*, *46*(7), 1642-1652.

Snyder, M. A., Sloan, L. C., Diffenbaugh, N. S., & Bell, J. L. (2003). Future climate change and upwelling in the California Current. *Geophysical Research Letters*, *30*, 1.

Thompson, P. A., Baird, M. E., Ingleton, T., & Doblin, M. A. (2009). Long-term changes in temperate Australian coastal waters: implications for phytoplankton. *Marine Ecology Progress Series*, *394*, 1-19.

USGS. (2010). Water Quality of San Francisco Bay Database. Retrieved from http://sfbay.wr.usgs.gov/access/wqdata.

Vahtera, E., Laanemets, J., Pavelson, J., Huttunen, M., & Kononen, K. (2005). Effect of upwelling on the pelagic environment and bloom-forming cyanobacteria in the western Gulf of Finland, Baltic Sea. *Journal of Marine Systems*, *58*(1-2), 67-82.