Adaptation of a weighted regression approach to evaluate water quality trends in an estuary

Marcus W. Beck[1], **James D. Hagy III**[2]

[1]*ORISE Research Participation Program*
USEPA National Health and Environmental Effects Research Laboratory
Gulf Ecology Division, 1 Sabine Island Drive, Gulf Breeze, FL 32561
Phone: 850-934-2480, Fax: 850-934-2401, Email: beck.marcus@epa.gov

[2]*USEPA National Health and Environmental Effects Research Laboratory*
Gulf Ecology Division, 1 Sabine Island Drive, Gulf Breeze, FL 32561
Phone: 850-934-2455, Fax: 850-934-2401, Email: hagy.jim@epa.gov

Abstract

The increasing availability of long-term monitoring data can improve resolution of temporal and spatial changes in water quality. In many cases, the fact that changes have occurred is no longer a matter of debate. However, the relatively simple methods that have been used to evaluate trends in environmental monitoring data in estuaries are often not sufficient to disaggregate the complex effects of multiple environmental drivers, limiting the potential to relate changes to possible causes. To improve the description of long-term changes in water quality, we adapted a weighted regression approach developed to describe trends in pollutant transport in rivers to analyze a long-term water quality dataset from Tampa Bay, Florida. The weighted regression approach allows for changes in the relationships between water quality and explanatory variables by using dynamic model parameters and can more clearly resolve the effects of both natural and anthropogenic drivers of ecosystem response. The model resolved changes in chlorophyll-*a* (chl-*a*) from 1974 to 2012 at seasonal and multi-annual time scales while considering variation associated with changes in freshwater influence. Separate models were developed for each of 4 Bay segments to evaluate spatial differences in patterns of long-term change. Observed trends reflected the known long term decrease in nitrogen loading to Tampa Bay since the 1970s, such that chl-*a* concentration decreased by 32% after municipal improvements in wastewater treament. Although trends in mean chl-*a* have remained constant since mitigation of point-sources of pollution, model predictions indicated that variation has increased in recent years for upper Bay segments, as well as an overall decrease in low productivity events for the lower Bay. Changes in model parameters for the last decade indicated that the relationship of salinity with chl-*a* was unimodal rather than continuous for upper Bay segments, such that flushing effects were observed beyond specific thresholds. Results from our analyses have allowed additional insight into water quality changes in Tampa Bay that has not been possible with traditional modeling approaches and the approach could easily be applied to other systems with long-term datasets.

Key words: chlorophyll, estuary, salinity, Tampa Bay, trend evaluation, weighted regression

1 Introduction

Eutrophication has been documented in aquatic systems worldwide and is of particular concern for coastal waters that support numerous aquatic life and human uses. Eutrophication is defined as an increase in the rate of supply of organic matter to a system (Nixon 1995) and is typically caused by elevated nitrogen or phosphorus loads. Although nutrients are necessary for the growth of primary producers, excessive anthropogenic inputs can have serious consequences for the structure and function of aquatic systems. Eutrophication of coastal systems has been associated with depletion of dissolved oxygen from the decomposition of organic matter (Diaz 2008), increases in the frequency and severity of harmful algal blooms (Glibert *et al.* 2013), and reduction or extirpation of seagrass communities (Duarte 1995, Tomasko *et al.* 2005). System-wide changes can occur as the effects of eutrophication on primary production propogate to upper trophic levels (Powers *et al.* 2005).

The effects of eutrophication are generally well understood, particularly for freshwater systems. The consequences of nutrient pollution were increasingly obvious by the 1960s such that eutrophication became a central focus of limnological research (Cloern 2001). However, the importance of understanding the relative effects of eutrophication on coastal systems were not realized until several decades later. For example, Rosenberg (1985) described the future hazards of coastal eutrophication nearly twenty years after similar issues were the focus of intense study in freshwater systems. Approaches for describing nutrient dynamics in coastal systems have relied heavily on freshwater eutrophication models that may not adequately describe idiosyncratic behaviors of individual estuaries. For example, Cloern (2001) suggests that system-specific attributes modulate coastal response to nutrient inputs, such that more appropriate conceptual models that recognize linked changes in relevant state variables are needed. To date, empirical models that are flexible and appropriate for site-specific conditions have not been extensively applied to describe nutrient-response dynamics in estuaries.

The increasing availability of long-term, high resolution datasets has further underscored the need to develop quantitative nutrient-response models given the potential to extract detailed information on system dynamics. In many cases (e.g., Caffrey 2003, Greening and Janicki 2006), long-term datasets have sufficiently described general trends in response to changing nutrient regimes or seasonal dynamics, although unambiguous and quantitative descriptions of responses have been lacking. For example, temporal variations in phytoplankton growth dynamics are often apparent by season with typical late summer blooms in temperate or tropical systems (Cloern 1996), whereas climate variation may contribute to substantial deviation in growth patterns between years (Jassby *et al.* 2002). Additionally, spatial heterogeneity in algal response to nutrients is common across salinity gradients such that effects of flux variation are most apparent near freshwater inflows (Cloern 1996). Simple statistical models that are constrained by assumptions of linearity and stationarity of variables through time may not adequately characterize subtleties in the variation of nutrient-response measures at different scales. Novel techniques that leverage the descriptive capabilities of large datasets are needed to improve our understanding of temporal and spatial variation in chlorophyll dynamics as a measure of eutrophication.

Use of simple descriptive statistics to evaluate the effects of water quality management may be ill-advised given that general trends in monitoring data may reflect both management actions and natural variation in system characteristics. Hirsch *et al.* (2010) developed the Weighted Regressions on Time, Discharge, and Season (WRTDS) approach to model pollutant concentration in rivers to address these issues and shortcomings of previously-developed models. WRTDS enables a flexible interpretation of water quality changes by estimating multiple parameters that are specific to a given season, year, and discharge for individual observations across the time series. This allows for a more detailed description of water quality changes than standard regression models that characterize trends using a single set of parameters. Accordingly, the approach addresses the need to focus on descriptions of change in relation to water quality variables across time, rather than hypothesis testing. The approach has been applied to model pollutant delivery from tributary sources to Chesapeake Bay (Hirsch *et al.* 2010, Moyer *et al.* 2012, Zhang *et al.* 2013), Lake Champlain (Medalie *et al.* 2012), and the Mississippi River (Sprague *et al.* 2011). The successful applications to water quality trends in rivers suggest the approach could potentially be applied to estuaries to characterize and better understand long-term changes in water quality. Moreover, changes in pollutant sources and variation in freshwater inputs over time for many coastal systems warrant the use of novel methods for trend evaluation. Resolving these changes may improve our understanding of linkages between drivers and responses over time.

Water quality data have been collected in the Tampa Bay estuary (Florida, USA) for approximately forty years. The natural history of Tampa Bay and the corresponding data provide a useful opportunity for applying quantitative methods to model nutrient dynamics. Nitrogen loads in the mid 1970s were estimated at 8.2×10^6 kg yr^{-1}, with approximately 5.5×10^6 kg yr^{-1} entering the upper bay alone (Poe *et al.* 2005, Greening and Janicki 2006). Reduced water clarity associated with phytoplankton biomass contributed to dramatic reduction in the areal coverage of seagrass (Tomasko *et al.* 2005) and development of hypoxic events causing a decline in benthic faunal production (Santos and Simon 1980). Extensive efforts to reduce nutrient loads to the Bay occurred by the late 1970s, with the most notable being improvements in infrastructure for wastewater treatment in 1979. Improvements in water clarity and decreases in chlorophyll-*a* (chl-*a*) were observed bay-wide in the 1980s, with conditions generally remaining constant to present day. Although the nutrient management program has been successful in improving water quality, variation in water quality drivers over time suggests the WRTDS method could provide information on system dynamics that are not apparent from the observed data.

The goal of the analysis was to describe changes in algal biomass in an estuary in relation to time, season, and freshwater inputs. We adapted the WRTDS approach developed by Hirsch *et al.* (2010) to describe water quality trends using a multi-decadal dataset from Tampa Bay, Florida. The analysis addressed four main objectives. First, we described the weighted regression model and provided a rationale for its adaptation to estuaries. Second, we applied the model to the time series in different segments of Tampa Bay to characterize trends in both the mean response of chl-*a* and the frequency of occurrence of extreme events using quantile regression. Third, additional factors related to water quality were used to describe the unexplained variance in chl-*a* growth patterns not characterized by the model. Specifically, model residuals were compared with variation in seagrass growth, El Niño-Southern Oscillation (ENSO) effects, and nitrogen load and concentrations in the Bay. Finally, we developed informed hypotheses to explain temporal and spatial patterns in chl-*a* growth in response to large scale drivers that affect water quality. Results from the analysis provide a natural history of water quality changes in Tampa Bay that is temporally consistent with drivers of change. This analytical approach may improve our understanding of the nutrient-response paradigm in coastal systems.

2 Methods

2.1 Data

We compiled a time series of chl-a concentration (μg L^{-1}) in Tampa Bay using data from the Hillsborough County Environmental Protection Commission (EPC) (TBEP 2011). Data are monthly at mid-depth for each of 50 stations throughout the Bay (Fig. 1) from 1974 to 2012, producing approximately 456 observations per station ($n = 1820$, Fig. 2). Stations were visited on a rotating schedule such that one third of all stations were sampled each week. Bay segments represent management units of interest with distinct chemical and physical differences (Table 1, Lewis and Whitman 1985). Accordingly, station data were averaged by segment. In addition to chl-a, salinity data were obtained and used as a tracer of freshwater influence on water quality. We expected that salinity was an important factor influencing interpretation of chl-a trends relative to the effects of additional factors (e.g., date, nutrient load, seagrass, etc.). Salinity data were converted to dimensionless values that represent the fraction of freshwater (Dyer 1973), such that:

$$Sal_{ff} = 1 - \frac{Sal_{mea}}{Sal_{ref}} \tag{1}$$

where Sal_{mea} is the measured salinity for a given station and Sal_{ref} is salinity at the seaward reference station for each observation date. Station 94 in the Gulf (Fig. 1) was used for reference salinity. Chlorophyll concentrations below the detection limit (censored data) were set to one-half the value from the limit to zero (Gilbert 1987). Chlorophyll data were ln-transformed because observations were skewed right, similar to a log-normal distribution. Kolomogorov-Smirnov tests indicated that the raw data were not signicantly different from theoretical log-normal distributions.

2.2 *Weighted regression*

WRTDS was adapted to relate chlorophyll concentration to salinity and time:

$$\ln(Chl) = \beta_0 + \beta_1 t + \beta_2 Sal_{ff} + \beta_3 \sin(2\pi t) + \beta_4 \cos(2\pi t) + \epsilon \quad (2)$$

where the natural log of chl-*a* is related to decimal time t, salinity Sal_{ff}, and unexplained variation ϵ. Salinity and time are linearly related to chl-*a* on a sinuisoidal annual time scale (i.e., cylical variation by year). The parameters $\beta_0, \ldots, \beta_4$ are estimated for each observed salinity at time t such that multiple sets of parameters are used to characterize the period of observation. Decimal time was calculated as the the year and month of each observation as an equivalent decimal (e.g., July 1974 as 1974.5). Although data were typically not collected on the first of each month, we considered the decimal time coincident with the period of observation. Additionally, quantile regression models (Cade and Noon 2003) were used to characterize trends at extreme conditional distributions of the data. Specifically, we adapted the weighted regression approach to model the conditional response at the 10$^{\text{th}}$ and 90$^{\text{th}}$ quantiles ($\tau = 0.1$, and 0.9, respectively) of the chlorophyll distribution. Quantile regression is analogous to least-squares regression such that a set of β parameters that minimizes the error term is estimated, where the minimization function is the sum of the weighted absolute deviations of the fitted values from the observed quantile. A general interpretation of the fitted values is the distribution of chl-*a* conditional upon time and salinity for low ($\tau = 0.1$) or high ($\tau = 0.9$) biomass events, rather than a characterization of 'average' conditions using mean models.

The WRTDS approach obtains fitted values of the response variable by estimating regression parameters for each unique observation. Specifically, a regression model was estimated for each of 1,. . ., 456 data points for each Bay segment. Each regression model was weighted by month, year, and salinity such that a unique set of regression parameters for each observation in the time series was obtained. For example, a weighted regression for October 2003 weights other observations in the same year, month, and similar salinity with higher values, whereas observations for different months, years, or salinities receive lower weights (Fig. 3). This weighting approach allows estimation of regression parameters that vary in relation to observed conditions. Hirsch *et al.* (2010) used a tri-cube weighting function:

$$w = \begin{cases} \left(1 - (d/h)^3\right)^3 & \text{if } |d| \leq h \\ 0 & \text{if } |d| > h \end{cases} \tag{3}$$

where the weight w for each observation is defined by the distance d from the current observation within a window h. The weights are diminishing in relation to the current observation until the maximum window width is exceeded and a weight of zero is used. The weight for each observation is the product of all three weights assigned to month, year, and salinity. Window widths of six months, 10 years, and half the range of Sal_{ff} for each Bay segment were used (Fig. 3). Window widths were increased by 10% increments during model estimation until a minimum of 100 observations with non-zero weights were obtained (Hirsch *et al.* 2010).

The adapted WRTDS approach was used to model and interpret chl-*a* trends from 1974–2012 for each of the 4 bay segments. In contrast with (Hirsch *et al.* 2010), estimates were made using monthly observations rather than daily predictions given the available data for Tampa Bay. Particular attention was given to trends that have not been previously described. Following Hirsch *et al.* (2010), predicted values were based on interpolation matrices for each model type (mean, 90th percentile , and 10th percentile) to reduce computation time. Specifically, a sequence of 20 salinity values based on the minimum and maximum values for each segment were used to predict chl-*a* using the observed month and year. Model predictions were then linearly interpolated from the grid using closest salinity value to the actual for each date. Hirsch *et al.* (2010) notes that the introduction of bias associated with using imprecise values in place of actual observations to estimate predictions was minimal. Model performance was based on coefficients of determination (R^2) for the mean regression models and pseudo-R^2 values that are specific to given quantiles (Koenker and Machado 1999). Additionally, root mean square error ($RMSE$) was calculated as an alternative measure of performance such that:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{n}} \tag{4}$$

where n is the number of observations from $1, \ldots, n$ for a given segment, y_i is the observed value of $\ln$-chl-*a* for observation i, and $\hat{y}_i$ is the predicted value for of $\ln$-chl-*a* for observation i. $RMSE$ values closer to zero represent model predictions closer to observed. The performance of weighted models were compared to conventional (i.e., non-weighted) additive linear models to show potential improvements using the WRTDS approach.

A potential issue for predictions with regression models in $\ln$-transformed space is bias associated with back-transformation (Duan 1983). Specifically, predicted values that are back-transformed by exponention may be biased due to variation in the concentration-salinity (or concentration-discharge) relationship through changes in residual variation across the data domain. We followed the approach in Moyer *et al.* (2012) that corrected for back-transformation bias using a scale parameter that is independently estimated for all regression models in the time series. The scale parameter describes the variance of the residuals such that:

$$\hat{\sigma}_\epsilon^2 = \frac{\sum\limits_{i=1}^{n} w_i \left(y_i - \hat{y}_i\right)^2}{\sum\limits_{i=1}^{n} w_i} \tag{5}$$

where residual variance $\hat{\sigma}_\epsilon^2$ (scale parameter) is the weighted sum of squared errors for chlorophyll for observations $1, \ldots, n$. Scale parameters were obtained for each unique regression across the time series and used to determine the correction bias in the back-transformation such that:

$$\alpha = \exp\left(\frac{\hat{\sigma}_\epsilon^2}{2}\right) \tag{6}$$

$$\hat{Chl} = \alpha \exp\left(\beta_0 + \beta_1 t + \beta_2 Sal_{ff} + \beta_3 \sin\left(2\pi t\right) + \beta_4 \cos\left(2\pi t\right)\right) \tag{7}$$

where the back-transformed chlorophyll concentration $\hat{Chl}$ is the exponentiated model prediction multiplied by the correction factor α (Moyer *et al.* 2012). Unique scale and correction bias parameters were obtained for each observation that was back-transformed. Although differences between results from bias-corrected predictions and simple exponentiation were minimal, eqs. (5) to (7) were used to create more accurate representations of chlorophyll trends in accordance with developed methods (i.e., Hirsch *et al.* 2010, Moyer *et al.* 2012).

In addition to trend description, the WRTDS approach can be used to normalize predicted values for a given explanatory variable to allow interpretation of trends in the absence of random variation. For example, water quality trends related to management actions cannot be precisely evaluated if pollutant concentrations vary with discharge. Hirsch *et al.* (2010) used the approach to normalize trends by flow, whereas our adapted approach was used to normalize by salinity which accounts for both freshwater inputs and tidal exchange. Normalized predictions were obtained for each observation date by assuming that salinity values for the same month in each year were equally likely to occur across the time series. That is, salinity is assumed to be uniformly distributed within the range of observed values for the same month between years. For example, normalization for January 1st 1974 considers all salinity values occuring on January 1st for each year in the time series as equally likely to occur on the observed data. A normalized value for January 1st 1974 is the average of the predicted values using each of the salinity values as input, while holding month and year constant. Normalization across the time series is repeated for each observation to obtain salinity-normalized predictions.

2.3 Evaluation of model residuals

An advantage of the WRTDS approach is the ability to describe water quality trends by considering changes in the relationships among variables for different observation periods. Additional factors not related to time or salinity could be used to describe the unexplained variation in the models (ϵ, eq. (2)). Residuals for the mean and quantile models in each Bay segment were related to additional variables with considerable management importance: seagrass growth, ENSO climate effects by season and year, and nitrogen load and concentrations. Conventional statistics were used to obtain a general description of the relationships, such as correlation coefficients and linear regression.

Seagrass coverage in Tampa Bay has been estimated bi-annually since 1988 (Tomasko *et al.* 2005). Coverage data are based on interpretation of aerial photos to produce raster surfaces with pixels coded as continuous (>75%) or patchy (25%–75%) coverage. Areal coverage of seagrass for years with available data ($n = 12$) were estimated by considering seagrass as present (continous or patchy) or absent within each Bay segment. ENSO data obtained from the Climate Prediction Center (CPC 2013) were based on a three year running-average of Sea Surface Temperature (SST) anamolies in the Niño 3.4 region of the Pacific Ocean (5°N–5°S, 120°–170°W). SST index values greater (less) than 0.4 (-0.4) were considered El Niño (La Niña) conditions, neutral otherwise. SST index values were categorically and quantitatively summarized by year and season using designations in Lipp *et al.* (2001): winter - January, February, March; spring - April, May, June; summer - July, August, September; fall - October, November, December. Finally, monthly loads for total nitrogen (TN, kg/mo) from 1985–2007 were obtained (Zarbock *et al.* 1994, Pribble *et al.* 2001, Poe *et al.* 2005), in addition to TN concentration from monitoring data (TBEP 2011). Nitrogen loads are based on estimated and measured contributions from nonpoint sources, point sources, atmospheric deposition, groundwater, and losses of phosphate rock and fertilizer from industrial processes.

3 Results

3.1 Observed trends in chlorophyll

Observed chl-*a* for all dates indicated mean values decreasing from Hillsborough (13 μg L^{-1}), to Old (8.8), to Middle (7.3), and to Lower Tampa Bay (3.8). Observed trends from 1974 to 2012 indicated a consistent decrease from 1974 to present as previously documented, with the most dramatic declines observed in the 1980s (Fig. 2a). Annual peaks in chl-*a* have also been observed associated with El Niño effects (Greening and Janicki 2006) in the mid-1990s. For example, 29.6 μg L^{-1}of chl-*a* was observed for Old Tampa Bay in October 1995. More extreme observations have been observed for individual stations during El Niño events.

Seasonal trends in observed chl-*a* were also consistent with documented changes, primarily in response to precipitation patterns. Maximum concentrations were generally observed in late summer whereas, minimum concentrations were observed in mid winter (Fig. 2b). Mean concentrations for the entire Bay were 12.8 μg L^{-1}for September and 4.5 for February. Trends by Bay segment were similar except that the amplitude of seasonal peaks diminished with proximity to the Gulf. For example, mean September and February concentrations for Lower Tampa Bay were 6.5 and 2.3, whereas concentrations in the same months for Hillsborough Bay were 21.9 and 8.4. Relationships of observed chl-*a* with salinity (as Sal_{ff}) indicated higher proportion freshwater was associated with higher chl-*a* (Pearson $\rho = 0.6$, $p < 0.005$, all observations). Correlations between chl-*a* and salinity by bay segment were similar (Pearson $\rho \approx 0.4$, $p < 0.005$ for all), with a slightly lower correlation in Old Tampa Bay ($\rho = 0.32$).

3.2 Predicted trends in chlorophyll

Predicted values obtained from the adapted WRTDS approach accounted for the effects of time and salinity on chl-*a* and generally followed observed trends as expected (Figs. S1 and 4). Predicted changes in chl-*a* by segment from 1974 to 2012 showed consistent declines throughout the time series (Table 2), as suggested by observed chl-*a* in Fig. 2a. Weighted regressions were also more precise than non-weighted additive linear regressions for all model types (Table 3). Mean explained variance using R^2 for all Bay segments was 0.45, 0.44, and 0.64 for the 90th percentile, 10th percentile, and mean models, respectively, compared to 0.3, 0.3, and 0.52 for the null models. Mean error using $RMSE$ for all Bay segments was 0.61, 0.62, and 0.37 for the 90th percentile, 10th percentile, and mean models, respectively, compared to 0.69, 0.7, and 0.43 for the null models. Additionally, increases in predictive performance from a non-weighted to weighted approach were slightly higher for the quantile models as compared to the mean models. Increases in predictive performance based on R^2 for both quantile models was 0.15 averaged for all Bay segments, whereas mean models increased an average of 0.12 over standard regression models. Similar trends were observed for $RMSE$ values (Table 3).

Substantial variation in chl-*a* response from the mean predicted values was observed despite high explained variance (Fig. 4). Observed values close to the mean response were fit well by the mean model, whereas extreme observations at low or high ends of the distribution were better predicted by the quantile models. For example, Fig. 5 shows the predicted and observed values for a two year period in Hillsborough Bay such that model fit varies depending on the month of observation. Model fit for peak observed chl-*a* in September and October of 1994 is best fit by the 90$^{\text{th}}$ percentile models, whereas a low seasonal peak observed in the winter of 1994 was best fit by the 10$^{\text{th}}$ percentile model. Larger differences between the predicted values for the 90$^{\text{th}}$ and 10$^{\text{th}}$ models was also observed in earlier years of the time series, such that the period from 1974-1980 had larger variation in predicted chl-*a*, in addition to higher overall mean values (Fig. 4).

Aggregation of model results by year allowed an evaluation of annual trends for predicted and salinity-normalized concentrations (Fig. 6). Predicted values illustrated response of chlorophyll by model type, whereas salinity-normalized estimates indicated annual trends independent of variation in tidal effects or tributary inputs by model segment. In general, trends were similar by model type such that increases or decreases in chl-*a* were similar regardless of the distribution that was characterized (i.e., mean, 0.9 τ, or 0.1 τ). Exceptions are noted for the 90$^{\text{th}}$ percentile models in the early years of the time series such that the frequency of high chl-*a* events were more common. Decreases in the variability of chl-*a* for Lower Tampa Bay in recent years are also apparent such that the 90$^{\text{th}}$ percentile model is decreasing and 10$^{\text{th}}$ percentile model is increasing, wheras predictions from the mean model are relatively constant. Additionally, recent trends for Middle and Lower Tampa Bay for the 10th percentile models suggest an increase in chl-*a* such that the occurrence of low concentration events are decreasing. An annual peak in predicted chl-*a* for 1998 was observed for all Bay segments which is not apparent for the salinity-normalized data. Further aggregation of the salinity-normalized results in Table 4 illustrated trends on decadal and seasonal time scales. In particular, trends prior to treatment of point sources of pollution from 1974–1980 generally indicated high and increasing chl-*a* for all segments and model types, excluding the 10th percentile model for Hillsborough Bay which showed consistent declines for the period. In contrast, the most dramatic declines in chl-*a* were estimated from 1980 to 1990 for all Bay segments. Accordingly, mean chl-*a* concentrations from 1980–1990 were less than the previous time period. A slight positive increase for the 10th percentile model for Lower Tampa Bay in recent years is also evident on a decadal time scale. Seasonal trends in salinity-normalized estimates indicated higher chl-*a* concentrations in warmer months and generally decreasing concentrations throughout the time series.

Of potential interest is an evaluation of between-year variability for all salinity-normalized estimates. Specifically, between-year comparisons of chl-*a* estimates for each model indicated that the range has not been constant throughout the time series (Table 5 and Fig. 7). Maximum within-year variability (as annual standard deviation divided by the mean) for all models was generally observed in recent years, with exceptions for models in Lower Tampa Bay where maximum variability was observed in 1993 (40.9%) for the mean model, 1975 (40.8%) for the 90th percentile model, and 1988 (38.7%) for the 10th percentile model. Increasing variability throughout the time series was particularly pronounced for the 90th percentile model for Old Tampa Bay with annual variation ranging from 25.4% in 1977 to 63.4% in 2012. Variability in salinity-normalized chl-*a* estimates across seasons were comparable, although variability was reduced in summer months (Table 5). Additionally, high variability was observed for Hillsborough Bay in winter and for Lower Tampa Bay in fall.

Evaluation of model predictions given changes in freshwater inputs and different periods of observation provides insight into the dynamic relationships between the response and predictor variables. Specificallly, an interpolation grid is produced for each model that is used to obtain both predictions and salinity-normalized results. The grid provides estimates of chl-*a* across the range of salinity values for each segment that are specific to each observation. Changes in the response of $\ln$-chl-*a* across salinity gradients for each Bay segment can be interpreted by plotting chl-*a* against Sal_{ff} for different dates (Fig. 8). For example, the response of chl-*a* in Hillsborough Bay with increasing freshwater input for early years was minimal, whereas a strong positive relationship is observed in later years. Higher freshwater inputs in recent years may also be associated with a threshold effect such that chl-*a* concentrations do not increase beyond a given value (e.g., 0.45 Sal_{ff}). Other Bay segments also show changes in the relationship between chl-*a* and freshwater inputs. For example, Lower Tampa Bay shows a stronger relationship between chl-*a* and Sal_{ff} for recent years.

3.3 Evaluation of model residuals

Mean residual values by segment indicated that the 90th and 10th percentile models over- and under-fit the respective quantile distributions, whereas residual values for the mean models were centered at approximately zero. In other words, the 90th percentile and 10th percentile models produced residuals that were negative and positive in sign, respectively, which is expected given the definition of quantile distributions. Correlations of residuals to additional explanatory variables indicated that chloropyhll response could be attributed to factors other than time and salinity (Table 6). Not surpisingly, significant correlations were observed with TN for all segments and models, although correlations were observed for concentration rather than load. All models and segments had positive correlations with concentration except the 90th percentile model for Hillsborough, Old, and Lower Tampa Bay and the mean model for Middle Tampa Bay. Only the 90th percentile model for Old Tampa Bay was positively correlated with TN load. Correlations with seagrass coverage and ENSO index values binned by year and season were not significant (Table 6). Regression models relating residuals to ENSO categories by year and season (e.g., El Niño fall) were not signficant. Regression models using continuous seasonal index values were also unable to resolve variance in the residuals, with the exception of the 10th percentile model for Lower Tampa bay such that a significant and positive relationship was observed between residuals and ENSO index values for spring dates ($F = 5.2$, $R^2 = 0.12$, $p = 0.028$).

4 Discussion

Application of the Weighted Regressions on Time, Discharge, and Season (WRTDS) model to a analyze a long-term record of chl-*a* in 4 segments of Tampa Bay provided an improved quantitative description of long-term changes relative to commonly applied methods. Because the descriptions are conceptually related to expected causes, the results enabled generation of informed hypotheses regarding ecosystem behavior and change and could suggest a potential approach for developing quantitative thresholds for water quality management. These conclusions are supported by several key aspects of the results. First, the WRTDS model provided improved predictions of chl-*a* relative to non-weighted regression, measured as both higher R^2 and lower $RMSE$ (Table 3). Second, WRTDS results for segments of the Bay that were historically most impacted by nutrient loading pointed to shifts in the response of chl-*a* to changes in freshwater inflows. These changes are temporally coherent with known changes in nutrient sources, suggesting that the WRTDS results quantify a response to changes in nutrient forcing. Finally, adaptation of WRTDS to predict quantiles in addition to the mean response provided information about long-term shifts in phytoplankton dynamics that are ecologically informative. In total, the results obtained by applying WRTDS to the Tampa Bay chl-*a* time series suggest that this model could be broadly useful for analyzing and interpreting the growing number of long-term data sets for water quality in estuaries.

4.1 Improved description of chl-a using WRTDS

The primary advantage of applying the WRTDS approach to the Tampa Bay dataset was an empirical description of water quality trends that accounted for the effects of freshwater variation over time. The approach allows for reconstruction of observed trends with more accuracy (Figs. S1 and 4), as well as the ability to predict chl-*a* response to changes in freshwater inputs that are temporally consistent for different periods of observation (Table 2). The increased predictive abilities of the WRTDS approach was apparent by comparison with unweighted linear model (Table 3). Hirsch *et al.* (2010) indicated similar improvements with application to Chesapeake Bay river inputs such that an increase in R^2 from 35% to 56% was observed using the weighted approach. Relative increases in predictive performance were not as dramatic for the Tampa Bay dataset, although R^2 values were higher than those in Hirsch *et al.* (2010). Improved model fit results in part from more flexible parameterization. This increases the ability of the model to describe historical patterns, but reduces application to predicting future chl-*a*. If drivers of chl-*a* are changing over time, predicting future chl-*a* while assuming that drivers are not changing could be of limited value. For example, WRTDS showed that the relationship between chl-*a* and freshwater forcing changed over time, such that predictions of chl-*a* in the near future would by necessity be based on the most recent estimates of the ecosystem response to freshwater forcing rather than the long-term average response. As such, the primary use of the WRTDS is a description of historical change that can lead to *post hoc* formulation of hypotheses. Hirsch *et al.* (2010) also used WRTDS to quantify changes ecological drivers, pointing to long-term changes in the strength and direction of discharge effects on nutrient concentrations in rivers. Watershed drivers of changes described by Hirsch *et al.* (2010) suggests similar conclusions can be made regardings drivers of observed changes in chl-*a* in Tampa Bay.

Pollutant sources for Tampa Bay have changed over time with an increasing dominance of non-point sources in recent years. Changes in pollutant sources may affect the relationship between chl-*a* and freshwater inputs. Nutrient concentrations and discharge are correlated regardless of pollutant sources, whereas the relationship between nutrient loading and discharge may vary. Increasing discharge with non-point sources of pollution is related to both increasing load and decreasing concentration of nutrients. Conversely, increasing discharge with point-sources of pollution may only be related to decreasing concentration since total load remains constant. Reduction of point sources of pollution in Tampa Bay and increasing dominance of non-point sources suggests that chl-*a* relationships with discharge may be dynamic over time. Application of the WRTDS model to the Tampa Bay dataset provided evidence of these shifts in the salinity-chlorophyll relationship over time. The shifts were most apparent for Bay segments that received large tributary inputs (Fig. 8). For example, the relationship of salinity with chlorophyll for Hillsborough Bay during earlier periods indicated no trend as expected, whereas the opposite was true for later periods. However, our measure of fraction of freshwater differs from discharge in that the effects of tidal exchange are also implicitly included. Accordingly, fraction of freshwater only partially explains the effects of tributary inputs. Hirsch *et al.* (2010) developed the WRTDS approach for rivers and streams where discharge effects are considered the primary variable affecting interpretation of water quality trends. Therefore, salinity effects were included in eq. (2) as being more appropriate for estuaries that are influenced by natural variation in both tidal flow and freshwater inputs (Cloern 1996).

The final objective of the analysis was to develop informed hypotheses of temporal and spatial patterns of chl-*a* growth in response to drivers of eutrophication in Tampa Bay. The most informative indication of changes for hypothesis development is illustrated by changes in chl-*a* response to freshwater inputs over time and by Bay segments (Fig. 8), particularly for Hillsborough Bay. Earlier periods (1974–1980) showed little response of chl-*a* to freshwater inputs, which is likely related to the dominance of point sources. An alternative explanation is provided by Wofsy (1983), such that phytoplankton growth dynamics in nutrient-saturated systems may be invariant to freshwater inputs. Biological processes, such as phytoplankton self-shading, may be more limiting for algal growth. Later periods showed significantly stronger responses of chl-*a* to freshwater inputs, likely related ot the relative influence of non-point sources in recent years, followed by a specific threshold response. Temporal dynamics for other Bay segments are also illustrative of changes in causal mechanisms. For example, Lower Tampa Bay shows increased sensitivity to freshwater inputs in recent years, despite relatively consistent mean concentrations in Fig. 6. Overall, differences between concurrent periods of observation and Bay segments remains a question of interest and results from the weighted regressions provide descriptions that facilitate interpretations.

4.2 Changes in chl-a variability

Most analyses of changes in water quality focus on changes in mean water quality over time. Linear models generally fit a constant seasonal cycle, a constant response to freshwater inflow, and a linear trend to describe the long-term change. The flexible parameterization of the WRTDS approach can substantially improve descriptions of water quality trends by addressing limitations of simple models. As a result, predicted values from WRTDS results are appropriate for evaluating change in direction of the response, whereas salinity-normalized values are useful for evaluating more subtle changes in variation. Direction and magnitude of change were primarily in agreement with expectations, whereas changes in variation over time have not been previosly described. Salinity-normalized predictions suggested that the variability of chlorophyll response between-years has generally been increasing, i.e., variability for most Bay segments has been larger than the most heavily polluted periods in the 1970s (Table 5 and Fig. 7). Differences were also observed by mean or quantile response, particularly for the 90^{th} percentile model in Old Tampa Bay. Mechanisms describing heterogeneity of chlorophyll between years is uncertain, although increasing variation in water quality parameters is a potential indicator of ecological transition in lakes (Carpenter and Brock 2006). Variation in chlorophyll could be an indication of impending changes despite constant mean values for several decades. We further emphasize that characterization of between-year variation is only possible with methods such as WRTDS. Less complex approaches that are not data-driven may be unable to resolve this variation (e.g., additive seasonal models, Cloern and Jassby 2010).

The inclusion of quantile models represents an important extension of the WRTDS approach by allowing insight into conditional response of chl-*a* not described by mean models.

Quantile models are particularly useful for characterizing response variables that exhibit considerable heterogeneity about the mean (Terrell *et al.* 1996, Cade and Noon 2003). Practical interpretation of the quantile models are such that the 90th percentile models show variation in the occurrence of extreme events whereas the 10th percentile models show variation in low productivity events. Quantification of extreme events may provide a more informative measure of progress towards ecosystem change in response to management. For example, a previous description for developing numeric criteria for Florida waters used the 90th percentile value from cumulative distribution models of chlorophyll for multiple coastal segments (Schaeffer *et al.* 2012). Although the exact upper percentile for criteria definition is arbitrary, consistency among methods could facilitate adoption in water quality standards. Similarly, variation in low productivity events could provide information of system departure from baseline or reference conditions (e.g., Stoddard *et al.* 2006). For example, variation in the 10th percentile model for Lower Tampa Bay in recent years suggests a consistent decrease in events with low chlorophyll concentrations (Fig. 6).

4.3 Limitations and future applications

The adaptation of the WRTDS approach to quantify chl-*a* trends in estuaries shows promise, although our analysis differs in several key aspects from the original model. First, issues of spatial scale will continue to have relevance given specific research objectives. The application of the WRTDS approach to Tampa Bay considered individual segments as being most relevant given our goal to provide a quantitative history of eutrophication that has importance for regional planning and decision-making processes. Different research objectives may warrant the use of Bay segments as inappropriate since phytoplankton growth patterns can be characterized at multiple scales. Cloern (1996) reviews spatial patterns of phytoplankton growth in estuaries such that longitudinal, lateral, and vertical dynamics are commonly observed. Growth dynamics may also be evident at scales ranging from meters to several kilometers. More subtle differences in spatial patterns are likely observed at individual stations in the Bay, which could serve as a focus for additional evaluation. Similarly, phytoplankton dynamics may be evident at different temporal scales. Hirsch *et al.* (2010) developed the WRTDS approach for daily water quality observations, although the Tampa Bay dataset prohibits analysis at time scales shorter than a month. A second consideration in our adaptation of the WRTDS model is the treatment of censored data. All censored data were set to one half the detection limit, as compared to a more quantitative approach by the WRTDS method using survival regression (Moyer *et al.* 2012). A post-hoc analysis of the Tampa Bay data suggested that the treatment of censored data affected the results, particularly for Lower Tampa Bay where chl-*a* concentrations are generally lower. However, effects were minimal and the overall conclusions were unchanged. Regardless, future modifications of the approach should include more robust treatment of censored data.

Additional considerations not unique to our adaption of the WRTDS approach deserve further investigation. The WRTDS method currently does not provide measures of uncertainty associated with model predictions, although development is in progress (R. Hirsch, personal communication May 2014). Lack of confidence in model predictions is a primary disadvantage of the approach that distinguishes it from alternative methods. For example, Moyer *et al.* (2012) compares the WRTDS methods with ESTIMATOR, an alternative regression-based approach (Cohn *et al.* 1992). Although WRTDS provided more accurate and precise descriptions, indications of uncertainty provided by ESTIMATOR suggested variation may be considerable in some cases. Moreover, the determination of appropriate window widths for defining model weights has been an issue of concern since initial development of the approach. A systematic evaluation of different combinations of window widths for reducing prediction error could be conducted to identify optimal widths. However, results may be specific to individual datasets and computional time may be excessive such that the increase in predictive performance may be trivial relative to time spent defining optimum widths. Regardless, the window widths used for our analysis produced useful results and could be used for additional applications.

The lack of correlation between model residuals and additional variables was unexpected, particularly for the seagrass and ENSO data. Previous analyses have illustrated the effects of precipitation events associated with ENSO on Tampa Bay. For example, Schmidt and Luther (2002) described ENSO effects on salinity profiles for Tampa Bay such that high precipitation events (i.e., El Niño spring or winters) were correlated with depressed salinity profiles. Our analyses indicated that residuals were not related to ENSO variation. The WRTDS model included salinity effects such that residual variation accounts for changes in freshwater inputs, which potentially explains lack of correlation with ENSO. Lack of correlation with seagrass data may have been related to sample size given that only annual estimates of seagrass coverage were available. Additionally, correlations between seagrass growth and chlorophyll may have been present for lags in the time series that were not evaluated. For example, high chl-*a* concentrations may have an effect on seagrass growth the following year, rather than within the same year.

4.4 Conclusions

Management over several decades has been successful in improving water quality in Tampa Bay from heavily degraded to more culturally desirable conditions (Greening and Janicki 2006). These changes have been most dramatic for Bay segments that receive a majority of nutrient pollution from tributary or point-sources, particularly Hillsborough and Old Tampa Bay. The general effects of management actions are therefore obvious, although quantitative descriptions of these changes that consider the effects of confounding variables on water quality dynamics have been lacking. Establishing direct links between management actions and changes in water quality are critical to inform the prioritization of limited resources for future decisions. Application of the WRTDS approach to Tampa Bay has provided a novel description of eutrophication dynamics that can be evaluated in the context of observed changes over time. Conclusions from the analysis showed that 1) improved statistical performance can be obtained using WRTDS as compared to traditional regression models, 2) the results reflected dynamic relationships between chl-*a* and salinity over time that suggested temporal shifts in nutrient forcing, and 3) considerable variation in chl-*a* response can be described by quantile distributions. Overall, the ability to describe the data and aspects of long-term changes has been improved by adaptation of the WRTDS approach to Tampa Bay. Such techniques are critical for informing the nutrient-response paradigm in coastal systems, providing an incentive for validation with additional long-term datasets.

Acknowledgments

We acknowledge the significant efforts of the Hillsborough County Environmental Protection Commission and the Tampa Bay Estuary Program in developing and providing access to high quality data sets. We also acknowledge the work of R. Hirsch and colleagues on the WRTDS model. This study was funded by the US Environmental Protection Agency (EPA), but the contents are solely the views of the authors. Use of trade names does not constitute endorsement by the US EPA.

References

Boler, R., 2001. Surface water quality 1999-2000 Hillsborough County, Florida, Technical report, Environmental Protection Commission of Hillsborough County, Tampa, Florida, USA.

Cade, B. S. and B. R. Noon, 2003. A gentle introduction to quantile regression for ecologists. Frontiers in Ecology and the Environment 1(8):412–420.

Caffrey, J. M., 2003. Production, respiration and net ecosystem metabolism in U.S. estuaries. Environmental Monitoring and Assessment 81(1-3):207–219.

Carpenter, S. R. and W. A. Brock, 2006. Rising variance: A leading indicator of ecological transition. Ecology Letters 9(3):308–315.

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

—, 2001. Our evolving conceptual model of the coastal eutrophication problem. Marine Ecology Progress Series 210:223–253.

Cloern, J. E. and A. D. Jassby, 2010. Patterns and scales of phytoplankton variability in estuarine-coastal ecosystems. Estuaries and Coasts 33(2):230–241.

Cohn, T. A., D. L. Caulder, E. J. Gilroy, L. D. Zynjuk, and R. M. Summers, 1992. The validity of a simple statistical model for estimating fluvial constituent loads: An empirical study involving nutrient loads entering Chesapeake Bay. Water Resources Research 28(9):2353–2363.

CPC (Climate Prediction Center), 2013. National Weather Service Climate Prediction Center: Cold and warm episodes by season. http://www.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ensoyears.shtml. (Accessed September, 2013).

Diaz, R. J., 2008. Spreading dead zones and consequences for marine ecosystems. Science 321:926–929.

Duan, N., 1983. Smearing estimate: A nonparametric retransformation method. Journal of the American Statistical Association 78(383):605–610.

Duarte, C. M., 1995. Submerged aquatic vegetation in relation to different nutrient regimes. Ophelia 41:87–112.

Dyer, K. R., 1973. Estuaries: A Physical Introduction, Wiley, London, United Kingdom.

Gilbert, R. O., 1987. Statistical Methods for Environmental Pollution Monitoring, Van Nostrand Reinhold, New York, USA.

Glibert, P. M., D. C. Hinkle, B. Sturgis, and R. V. Jesien, 2013. Eutrophication of a Maryland/Virginia coastal lagoon: A tipping point, ecosytem changes, and potential causes. Estuaries and Coasts .

Greening, H. and A. Janicki, 2006. Toward reversal of eutrophic conditions in a subtrophical estuary: Water quality and seagrass response to nitrogen loading reductions in Tampa Bay, Florida, USA. Environmental Management 38(2):163–178.

Hirsch, R. M., D. L. Moyer, and S. A. Archfield, 2010. Weighted regressions on time, discharge, and season (WRTDS), with an application to Chesapeake Bay river inputs. Journal of the American Water Resources Association 46(5):857–880.

Jassby, A. D., J. E. Cloern, and B. E. Cole, 2002. Annual primary production: Patterns and mechanisms of change in a nutrient-rich tidal ecosystem. Limnology and Oceanography 47(3):698–712.

Koenker, R. and J. A. F. Machado, 1999. Goodness of fit and related inference processes for quantile regression. Journal of the American Statistical Association 94(448):1296–1310.

Lewis, R. R. and E. D. Estevez, 1988. The ecology of Tampa Bay, Florida: An estuarine profile, Technical Report Biological Report 85(7.18), US Fish and Wildlife Service, Washington DC, USA.

Lewis, R. R. and R. L. Whitman, 1985. A new geographic description of the boundaries and subdivisions of Tampa Bay, in: Proceedings of the Tampa Bay Area Scientific Information Symposium, S. F. Treat, J. L. Simon, R. R. Lewis, and R. L. Whitman, editors, Florida Sea Grant Report No. 65, Bellwether Press, Tampa, Florida, USA, pp. 10–18.

Lipp, E. K., N. Schmidt, M. E. Luther, and J. B. Rose, 2001. Determining the effects of El Niño-Southern Oscillation events on coastal water quality. Estuaries and Coasts 24(4):491–497.

Medalie, L., R. M. Hirsch, and S. A. Archfield, 2012. Use of flow-normalization to evaluate nutrient concentration and flux changes in Lake Champlain tributaries, 1990-2009. Journal of Great Lakes Research 38(SI):58–67.

Moyer, D. L., R. M. Hirsch, and K. E. Hyer, 2012. Comparison of two regression-based approaches for determining nutrient and sediment fluxes and trends in the Chesapeake Bay Watershed, Technical Report Scientific Investigations Report 2012-544, US Geological Survey, US Department of the Interior, Reston, Virginia.

Nixon, S. W., 1995. Coastal marine eutrophication: A definition, social causes, and future concerns. Ophelia 41:199–219.

Poe, A., K. Hackett, S. Janicki, R. Pribble, and A. Janicki, 2005. Estimates of total nitrogen, total phosphorus, total suspended solids, and biochemical oxygen demand loadings to Tampa Bay, Florida: 1999-2003, Technical Report #02-05, Tampa Bay Estuary Program, St. Petersburg, Florida, USA.

Powers, S. P., C. H. Peterson, R. R. Christian, E. Sullivan, M. J. Powers, M. J. Bishop, and C. P. Buzzelli, 2005. Effects of eutrophication on bottom habitat and prey resources of demersal fishes. Marine Ecology Progress Series 302:233–243.

Pribble, R., A. Janicki, H. Zarbock, S. Janicki, and M. Winowitch, 2001. Estimates of total nitrogen, total phosphorus, total suspended solids, and biochemical oxygen demand loadings to Tampa Bay, Florida: 1995-1998, Technical Report #05-01, Tampa Bay Estuary Program, St. Petersburg, Florida, USA.

Rosenberg, R., 1985. Eutrophication – the future marine coastal nuisance? Marine Pollution Bulletin 16(6):227–231.

Santos, S. L. and J. L. Simon, 1980. Marine soft-bottom community establishment following annual defaunation: Larval or adult recruitment. Marine Ecology - Progress Series 2(3):235–241.

Schaeffer, B. A., J. D. Hagy, R. N. Conmy, J. C. Lehrter, and R. P. Stumpf, 2012. An approach to developing numeric water quality criteria for coastal waters using the SeaWiFS satellite data record. Environmental Science and Technology 46(2):916–922.

Schmidt, N. and M. E. Luther, 2002. ENSO impacts on salinity in Tampa Bay, Florida. Estuaries and Coasts 25(5):976–984.

Sprague, L. A., R. M. Hirsch, and B. T. Aulenbach, 2011. Nitrate in the Mississippi River and its tributaries, 1980 to 2008: Are we making progress? Environmental Science and Technology 45(17):7209–7216.

Stoddard, J. L., D. P. Larsen, C. P. Hawkins, R. K. Johnson, and R. H. Norris, 2006. Setting expectations for the ecological condition of streams: The concept of reference condition. Ecological Applications 16(4):1267–1276.

TBEP (Tampa Bay Estuary Program), 2011. Tampa Bay Water Atlas. http://www.tampabay.wateratlas.usf.edu/. (Accessed October, 2013).

Terrell, J. W., B. S. Cade, J. Carpenter, and J. M. Thompson, 1996. Modeling stream fish habitat limitations from wedge-shaped patterns of variation in standing stock. Transactions of the American Fisheries Society 125(1):104–117.

Tomasko, D. A., C. A. Corbett, H. S. Greening, and G. E. Raulerson, 2005. Spatial and temporal variation in seagrass coverage in Southwest Florida: Assessing the relative effects of anthropogenic nutrient load reductions and rainfall in four contiguous estuaries. Marine Pollution Bulletin 50(8):797–805.

Wofsy, S. C., 1983. A simple model to predict extinction coefficients and phytoplankton biomass in eutrophic waters. Limnology and Oceanography 28(6):1144–1155.

Zarbock, H., A. Janicki, D. Wade, D. Heimbuch, and H. Wilson, 1994. Estimates of total nitrogen, total phosphorus, and total suspended solids loadings ot Tampa Bay, Florida, Technical Report No. 04-94, Prepared for the Tampa Bay National Estuary Program, St. Petersburg, Florida, USA.

Zhang, Q., D. C. Brady, and W. P. Ball, 2013. Long-term seasonal trends of nitrogen, phosphorus, and suspended sediment load from the non-tidal Susquehanna River Basin to Chesapeake Bay. Science of the Total Environment 452-453:208–221.

Table 1: Summary of characteristics for Tampa Bay segments. Mean chlorophyll and salinity data for 2012 are shown. Sources: Lewis and Whitman (1985), Lewis and Estevez (1988).

Segment	Area (km^2)	Shoreline length (km)	Mean depth (m)	Watershed area (km^2)	Chlorophyll-*a* (μg L^{-1})	Salinity
HB	105.3	128.6	3.2	3319.9	9.9	24.4
OTB	200.7	339.8	2.8	874.4	7.6	23.5
MTB	309.9	262.8	4.1	1062.7	6.1	27.1
LTB	246.6	121.6	3.8	330.5	4.1	32.2

Note: HB: Hillsborough Bay, OTB: Old Tampa Bay, MTB: Middle Tampa Bay, LTB: Lower Tampa Bay

Table 2: Expected changes in chlorophyll (μg L^{-1}, percent in parentheses) from 1974 to 2012 and 2005 to 2012 from each model and Bay segment at low, moderate, and high values for fraction of freshwater (Sal_{ff}). Values are the estimated differences in predicted chl-*a* concentrations for July in each year. Low, moderate, and high are relative values for each segment for all years during July.

Models	**1974-2012**			**2005-2012**		
	Low	Moderate	High	Low	Moderate	High
HB						
mean	-16.7 (-60.6)	-10.8 (-39.2)	-2.8 (-9.8)	-1.6 (-12.9)	-0.5 (-2.8)	6.0 (30.7)
0.9 τ	-29.0 (-74.4)	-24.7 (-57.2)	-13.4 (-32.8)	-10.2 (-50.5)	-9.5 (-33.8)	-7.6 (-21.8)
0.1 τ	-9.4 (-46.9)	-6.4 (-32.2)	0.4 (2.3)	1.7 (19.7)	0.9 (6.9)	9.8 (94.7)
OTB						
mean	0.2 (1.1)	0.4 (2.4)	2.0 (13.3)	4.6 (40.3)	2.7 (21.2)	1.9 (12.8)
0.9 τ	11.5 (50.2)	6.2 (26.4)	9.6 (42.1)	12.4 (56.4)	3.8 (14.8)	-0.2 (-0.7)
0.1 τ	-0.5 (-6.0)	1.5 (18.7)	6.8 (100.8)	0.9 (11.9)	0.7 (7.4)	3.5 (34.7)
MTB						
mean	-3.6 (-36.2)	-1.5 (-14.4)	-0.4 (-2.9)	-0.6 (-8.8)	0.6 (6.8)	2.2 (22.2)
0.9 τ	-6.3 (-43.1)	-3.5 (-23.4)	0.2 (1.1)	-1.5 (-14.9)	-0.4 (-3.3)	1.0 (6.7)
0.1 τ	-2.0 (-31.7)	0.9 (14.2)	0.2 (2.1)	-0.1 (-2.7)	1.5 (27.0)	3.1 (43.8)
LTB						
mean	0.6 (15.4)	1.7 (38.7)	3.3 (67.8)	-0.03 (-0.7)	0.5 (9.2)	1.1 (15.3)
0.9 τ	0.05 (0.9)	0.7 (9.1)	2.1 (24.0)	-1.0 (-14.4)	-1.9 (-18.6)	-3.2 (-22.7)
0.1 τ	1.8 (78.6)	2.0 (72.8)	3.0 (99.7)	0.8 (22.8)	1.2 (34.1)	2.1 (54.4)

Table 3: Model performance by bay segment comparing non-weighted and weighted regression. Performance is evaluated using R^2 for mean models, pseudo-R^2 for 90$^{\text{th}}$ and 10$^{\text{th}}$ percentile (τ) models, and $RMSE$ for all models (statistics by Bay segment).

Statistic	**mean**		**0.9 τ**		**0.1 τ**	
	Non-wtd	Wtd	Non-wtd	Wtd	Non-wtd	Wtd
HB						
R^2	0.54	0.66	0.32	0.47	0.31	0.45
$RMSE$	0.48	0.41	0.78	0.66	0.74	0.67
OTB						
R^2	0.54	0.65	0.29	0.45	0.34	0.47
$RMSE$	0.41	0.36	0.65	0.61	0.67	0.59
MTB						
R^2	0.60	0.71	0.34	0.51	0.38	0.51
$RMSE$	0.37	0.31	0.60	0.52	0.61	0.52
LTB						
R^2	0.40	0.51	0.26	0.37	0.18	0.34
$RMSE$	0.45	0.40	0.72	0.65	0.77	0.68

Note: HB: Hillsborough Bay, OTB: Old Tampa Bay, MTB: Middle Tampa Bay, LTB: Lower Tampa Bay

Table 4: Decadal and seasonal summaries of salinity-normalized chl-*a* (μg L^{-1}) trends by Bay segment. Trends are evaluated for models fit through the mean response and the 10th and 90th percentile (τ) distributions. Mean and slope (Δ chl-*a*) estimates are aggregated by year or season categories using monthly results. Slopes indicate the change in chl-*a* with increasing time for each year or season category.

Models	**1974-1980**		**1980-1990**		**1990-2000**		**2000-2012**	
	Mean	Δ chl-*a*	Mean	Δ chl-*a*	Mean	Δ chl-*a*	Mean	Δ chl-*a*
HB								
mean	24.91	0.06	18.30	-0.86***	13.00	-0.11	11.30	0.02
0.9 τ	43.51	1.56*	33.33	-2.00***	22.54	-0.10	19.23	-0.15
0.1 τ	16.41	-0.55**	10.88	-0.45***	8.10	-0.02	7.37	0.08
OTB								
mean	12.45	0.55	10.94	-0.31*	8.84	0.06	9.10	0.12
0.9 τ	19.26	0.72	17.26	-0.36*	14.91	0.12	16.30	0.16
0.1 τ	8.45	0.57**	7.32	-0.23**	5.59	0.02	6.08	0.07
MTB								
mean	10.33	0.77***	10.05	-0.34***	7.51	-0.04	6.39	0.01
0.9 τ	15.45	1.27***	16.40	-0.65***	11.61	-0.11	9.46	-0.05
0.1 τ	6.88	0.46**	6.81	-0.14	5.29	-0.05	4.57	0.05
LTB								
mean	4.68	0.33**	4.39	-0.11*	3.88	0.07	4.06	0.02
0.9 τ	8.32	0.42	8.05	-0.13	6.82	0.12	6.56	-0.12*
0.1 τ	2.84	0.22**	2.62	-0.10**	2.29	0.07**	2.75	0.07***
	winter		**spring**		**summer**		**fall**	
	Mean	Δ chl-*a*	Mean	Δ chl-*a*	Mean	Δ chl-*a*	Mean	Δ chl-*a*
HB								
mean	10.65	-0.45***	13.75	-0.44***	22.37	-0.33***	14.88	-0.39***
0.9 τ	18.20	-0.72***	23.83	-0.98***	39.67	-0.77***	26.48	-0.56***
0.1 τ	6.49	-0.32***	8.58	-0.21***	14.35	-0.19***	9.43	-0.26***
OTB								
mean	5.49	-0.10***	8.88	-0.12***	15.09	-0.06***	10.46	-0.12***
0.9 τ	9.81	-0.19***	13.72	-0.11***	24.84	0.17***	18.19	-0.22***
0.1 τ	3.53	-0.07***	6.28	-0.06***	9.92	-0.07***	6.67	-0.09***
MTB								
mean	5.23	-0.13***	7.50	-0.18***	11.95	-0.15***	8.14	-0.12***
0.9 τ	8.86	-0.25***	11.33	-0.28***	17.71	-0.23***	12.85	-0.22***
0.1 τ	3.55	-0.08***	5.31	-0.11***	8.47	-0.10***	5.37	-0.05***
LTB								
mean	2.70	-0.01***	3.41	-0.01*	5.97	0.00	4.67	-0.05***
0.9 τ	4.91	-0.05***	5.85	-0.03***	9.96	-0.05***	8.34	-0.15***
0.1 τ	1.71	0.01***	2.22	0.00	3.73	0.00	2.78	0.00

Note: *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; HB: Hillsborough Bay, OTB: Old Tampa Bay, MTB: Middle Tampa Bay, LTB: Lower Tampa Bay

Table 5: Variability in chl-a (μg L^{-1}) for Bay segments by year and seasons using salinity-normalized predictions. Variability (%) was quantified as the standard deviation of predictions by year (or season) category divided by the mean of predictions by year (or season) category. Trends are evaluated for models fit through the mean response and the 10$^{\text{th}}$ and 90$^{\text{th}}$ percentile (τ) distributions.

Models	**1974-1980**	**1980-1990**	**1990-2000**	**2000-2012**
HB				
mean	13.3	32.2	41.5	45.2
0.9 τ	24.2	37.2	41.4	49.1
0.1 τ	16.7	35.2	45.8	48.6
OTB				
mean	33.6	34.9	41.7	47.4
0.9 τ	26.4	30.4	43.0	54.3
0.1 τ	35.3	35.5	46.1	43.7
MTB				
mean	27.4	30.4	42.0	39.5
0.9 τ	25.0	26.4	37.5	39.1
0.1 τ	31.1	33.8	42.9	39.1
LTB				
mean	34.4	34.1	36.6	34.8
0.9 τ	37.2	31.6	34.7	33.5
0.1 τ	36.8	37.4	33.2	32.7
	winter	**spring**	**summer**	**fall**
HB				
mean	53.1	41.8	20.5	39.9
0.9 τ	47.9	55.6	29.4	35.5
0.1 τ	65.3	35.9	21.8	44.2
OTB				
mean	23.0	29.3	13.9	30.5
0.9 τ	23.9	25.8	20.1	31.3
0.1 τ	29.0	31.1	13.6	33.2
MTB				
mean	33.0	33.0	18.5	31.9
0.9 τ	38.0	36.4	20.6	31.4
0.1 τ	30.6	29.4	19.2	31.2
LTB				
mean	12.5	18.7	13.0	25.9
0.9 τ	17.8	16.9	17.8	28.6
0.1 τ	13.6	21.8	15.7	27.1

Note: winter: JFM, spring: AMJ, summer: JAS, fall: OND; HB: Hillsborough Bay, MTB: Middle Tampa Bay, OTB: Old Tampa Bay, LTB: Lower Tampa Bay

Table 6: Correlations between model residuals for each bay segment and potential drivers of chl-a (μg L^{-1}) independent of season, year, or salinity effects. Residuals were compared with seagrass area (hectares), mean ENSO index values by season and year, and total nitrogen load (kg$\cdot$mo^{-1}) and concentration (μg L^{-1}).

Models		**ENSO**		**TN**	
	seagrass	annual	season	load	conc.
HB					
mean	0.23	0.25	0.03	0.03	0.11*
0.9 τ	0.26	0.16	0.01	0.00	0.07
0.1 τ	-0.07	0.30	0.04	0.02	0.12*
MTB					
mean	-0.40	0.10	-0.04	0.00	0.11
0.9 τ	-0.24	0.06	-0.01	0.06	0.12*
0.1 τ	-0.43	0.06	-0.07	-0.06	0.12*
OTB					
mean	0.04	0.08	-0.04	0.06	0.19***
0.9 τ	0.23	0.02	-0.06	0.16**	0.11
0.1 τ	-0.03	0.07	-0.04	0.03	0.23***
LTB					
mean	0.09	0.08	0.00	0.01	0.20***
0.9 τ	0.28	0.08	0.03	0.06	0.10
0.1 τ	-0.19	0.07	-0.02	0.01	0.33***

Note: *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; for Pearson correlations, sample size varies from 11 to 308, HB: Hillsborough Bay, MTB: Middle Tampa Bay, OTB: Old Tampa Bay, LTB: Lower Tampa Bay

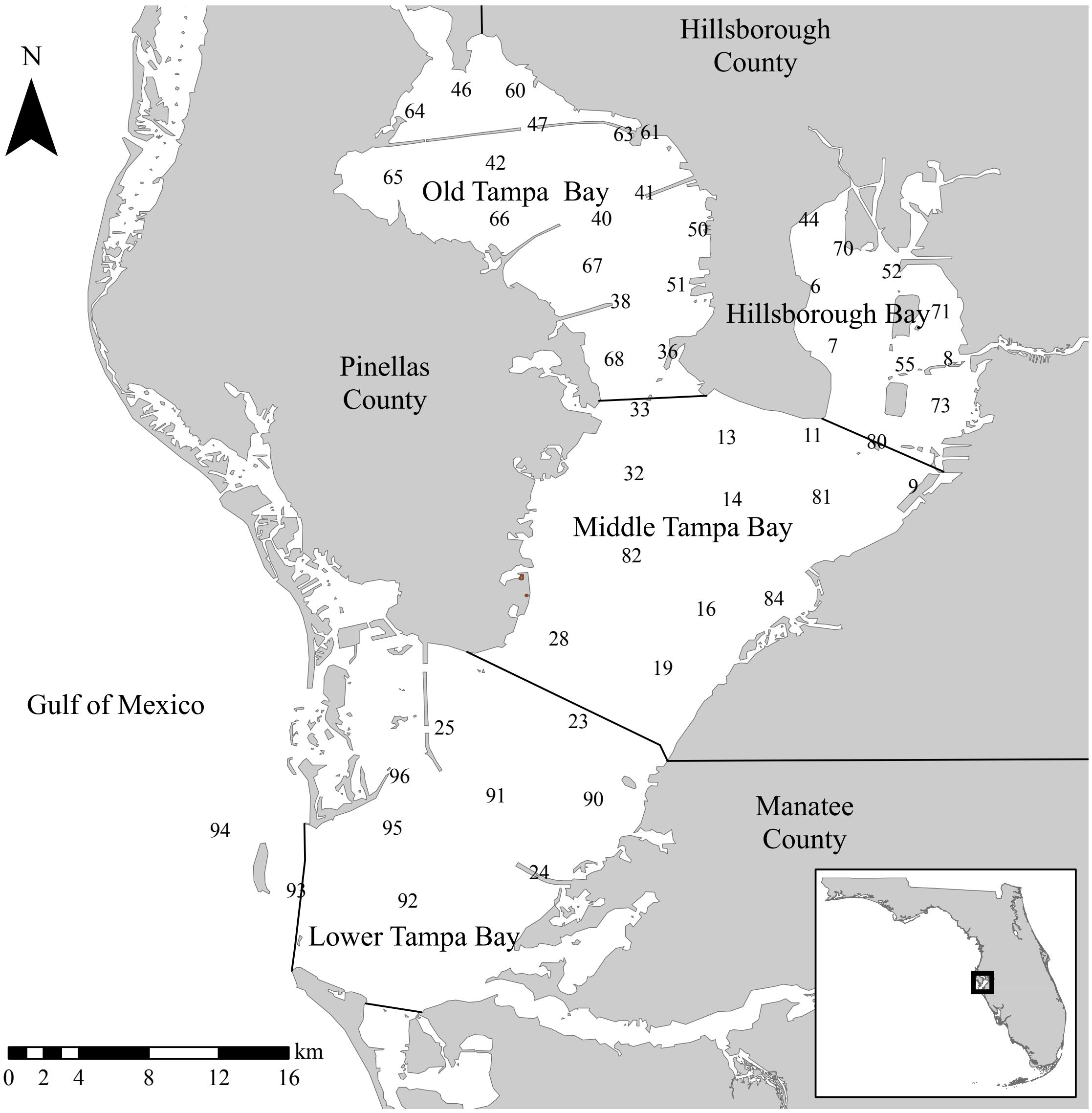
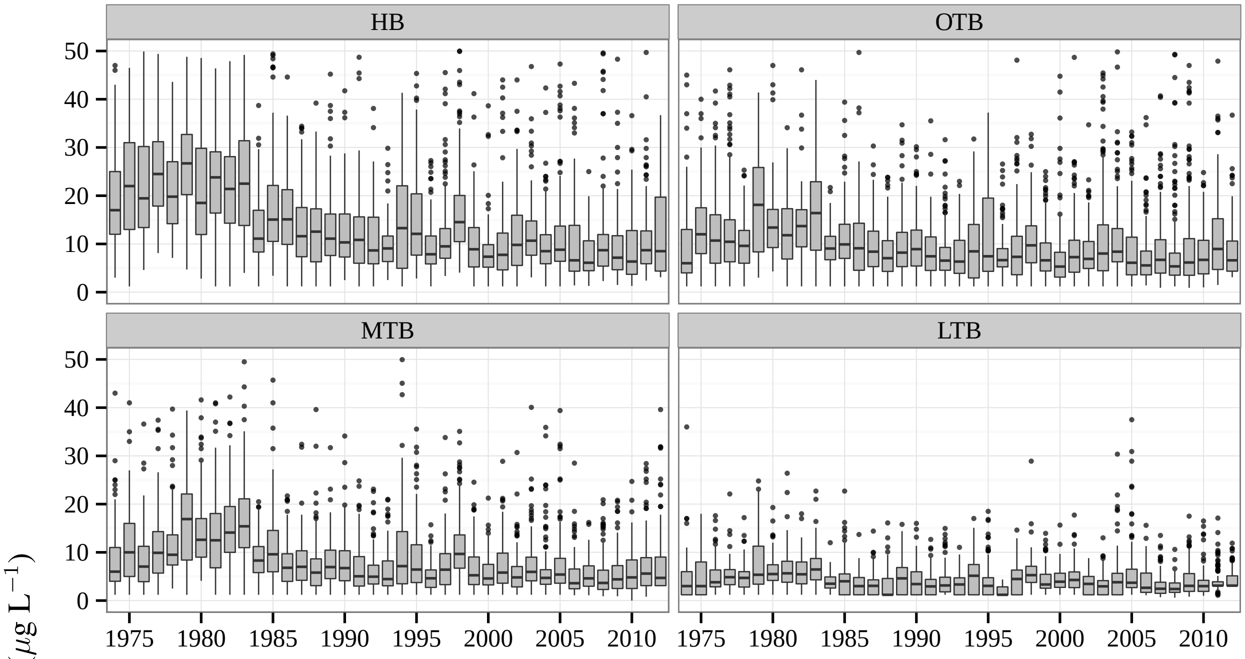


Fig. 1: The Tampa Bay estuary located on the west coast of central Florida. The Bay is separated into four segments defined by chemical, physical, and geopolitical boundaries (Lewis and Whitman 1985). Monthly water quality monitoring stations are also indicated by their identification number (Boler 2001).

(a) By year



(b) By month

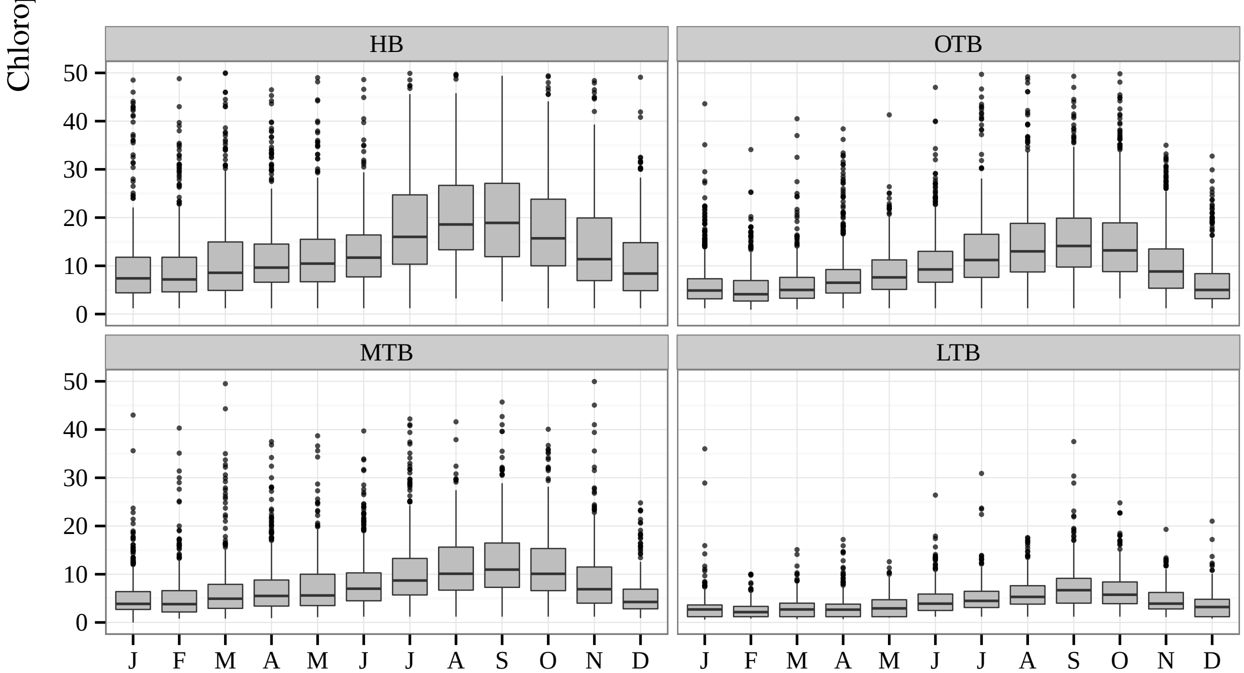


Fig. 2: Observed chl-*a* data for Tampa Bay segments by (a) year and (b) month aggregations. Each box is bisected by the median and represents the IQR (25$^{\text{th}}$ to 75$^{\text{th}}$ percentile). Outliers are present beyond whiskers (1.5$\cdot$IQR) and were observed beyond 50 μg L^{-1}. HB: Hillsborough Bay, MTB: Middle Tampa Bay, OTB: Old Tampa Bay, LTB: Lower Tampa Bay.

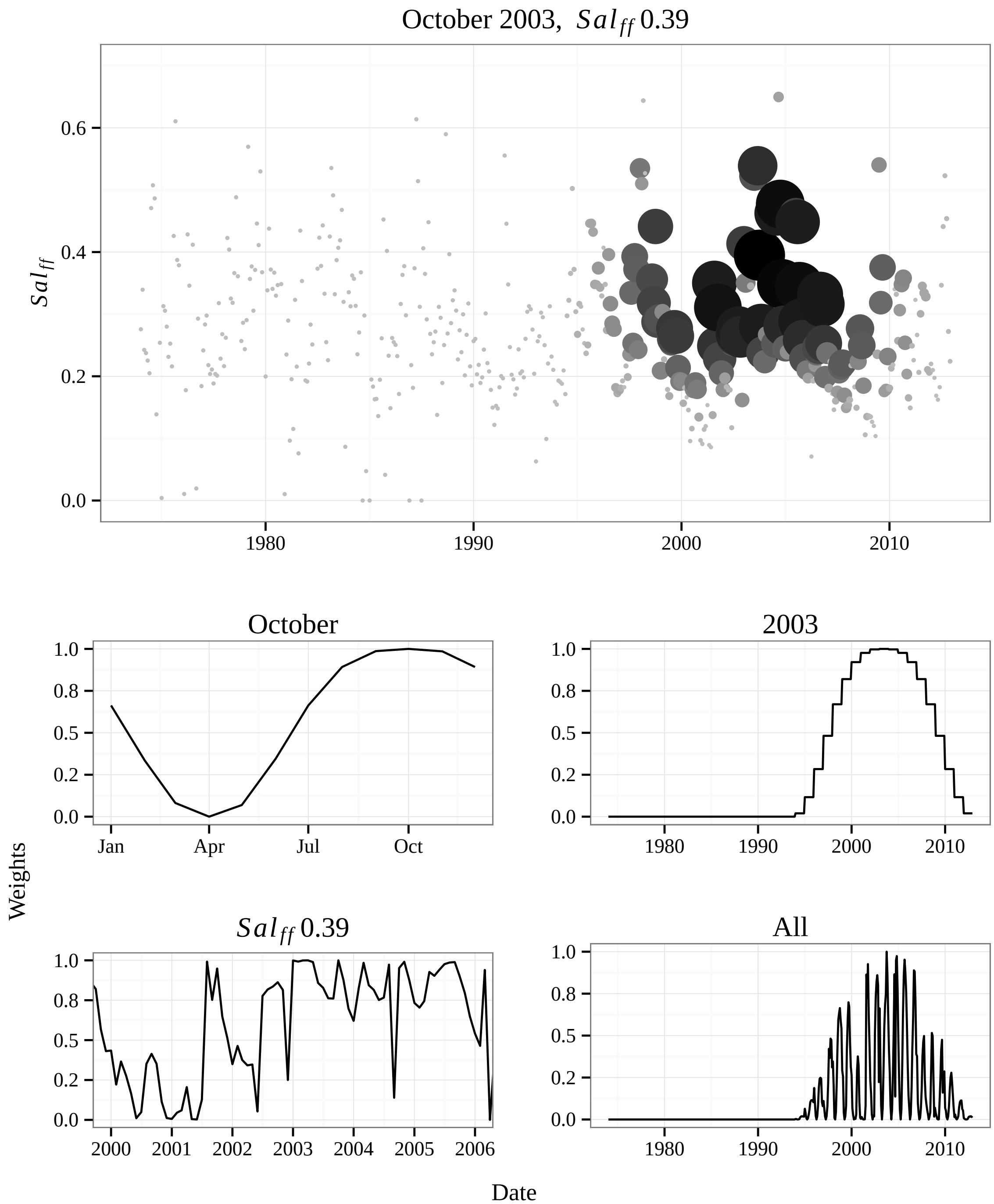


Fig. 3: Example of weighting for one observation in Hillsborough Bay. The top plot shows all data weighted for October 2003 when the proportion freshwater was 0.39. Point size and color are in proportion to weights (small grey points = 0, large black points = 1). The bottom plots show the individual weights for month, year, proportion freshwater, and all weights combined.

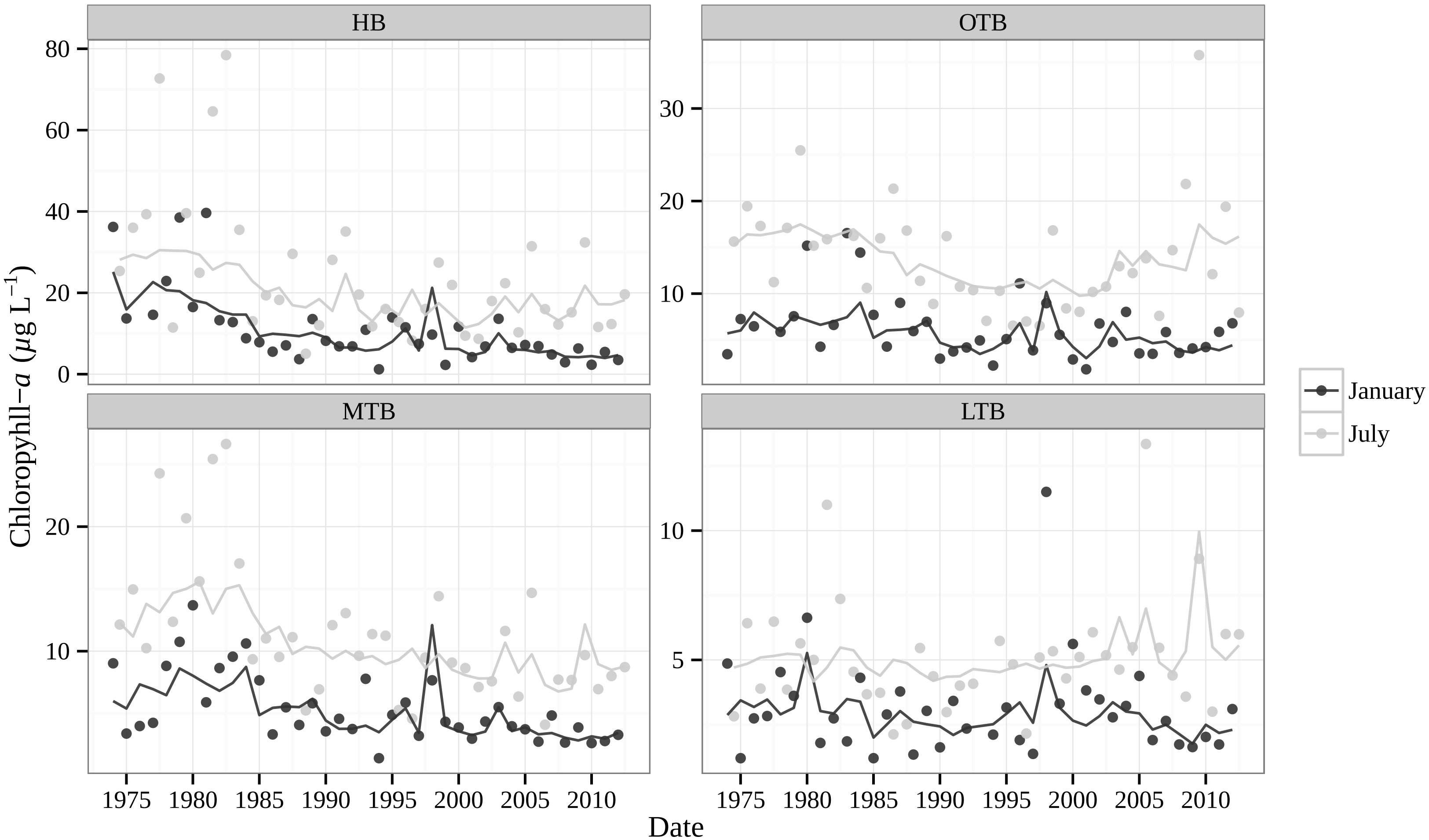


Fig. 4: Predicted (lines) and observed (circles) chl-*a* concentrations for Tampa Bay segments. Only January and July are shown to remove seasonal variation (see Fig. S1 for all months). Predicted values are for the weighted regression models fit through the mean response. HB: Hillsborough Bay, OTB: Old Tampa Bay, MTB: Middle Tampa Bay, LTB: Lower Tampa Bay.

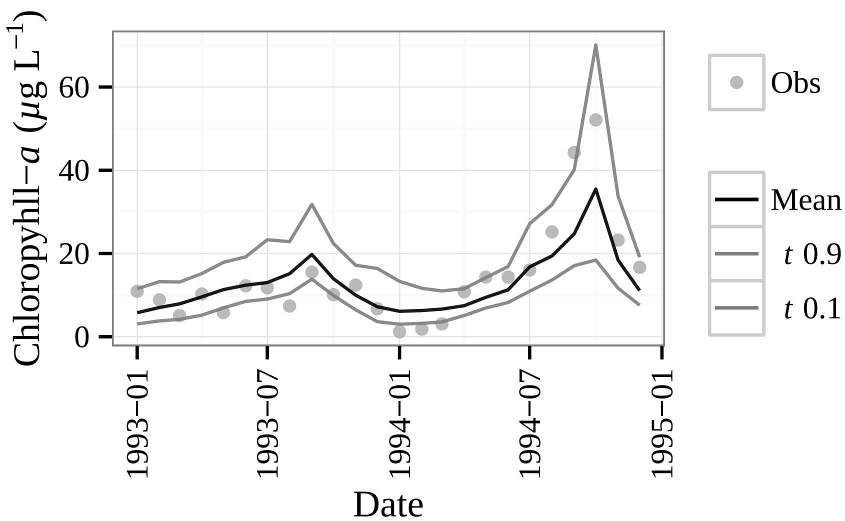


Fig. 5: Predicted and observed chl-*a* concentrations for Hillsborough Bay for 1993 to 1995 illustrating variation in model fit based on observation date. Predicted values are for the weighted regression models fit through the mean response and the 10th and 90th percentile (τ) distributions.

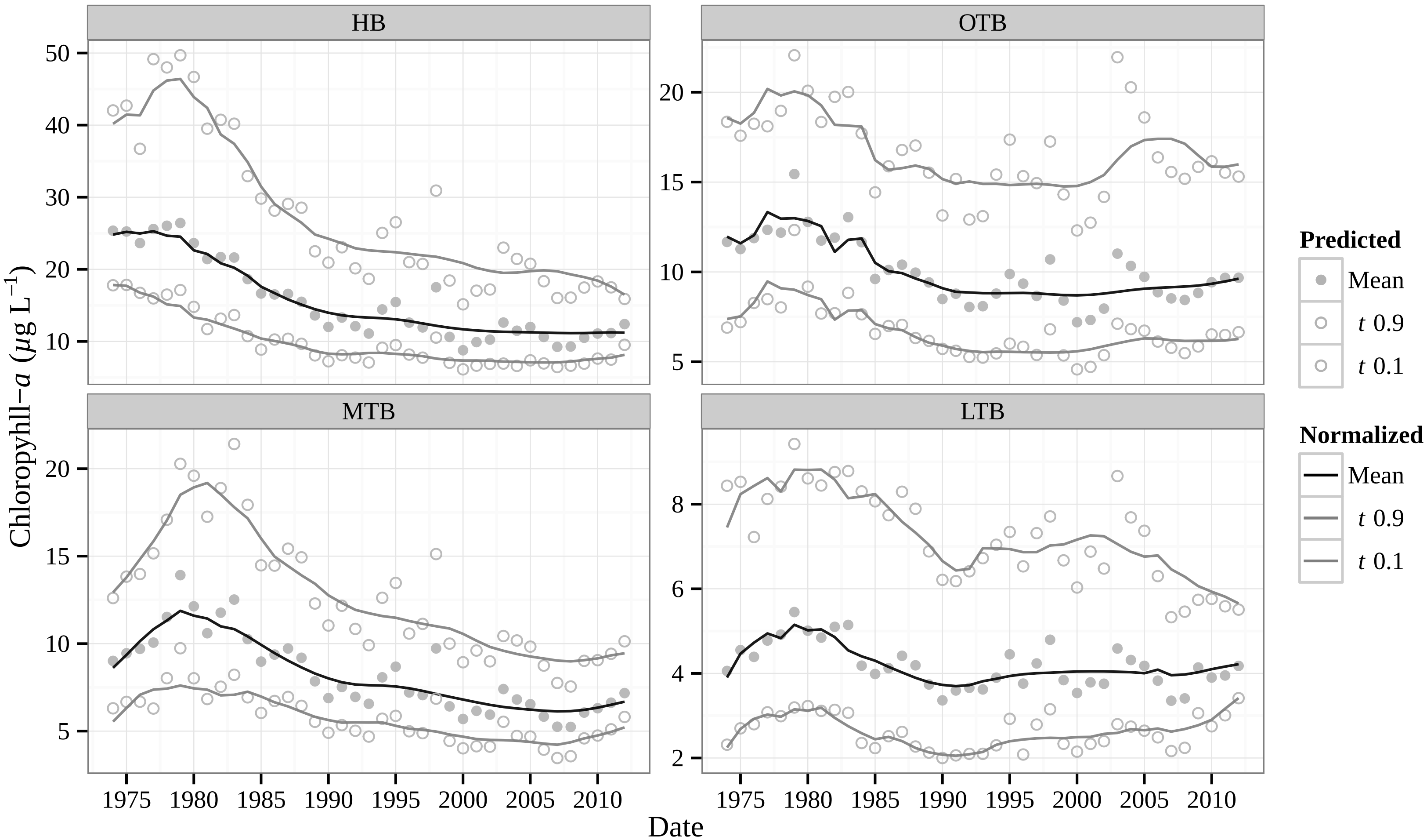


Fig. 6: Weighted regression predictions and salinity-normalized results aggregated by year for the mean response and the 10$^{\text{th}}$ and 90$^{\text{th}}$ quantile (τ) distributions. HB: Hillsborough Bay, OTB: Old Tampa Bay, MTB: Middle Tampa Bay, LTB: Lower Tampa Bay.

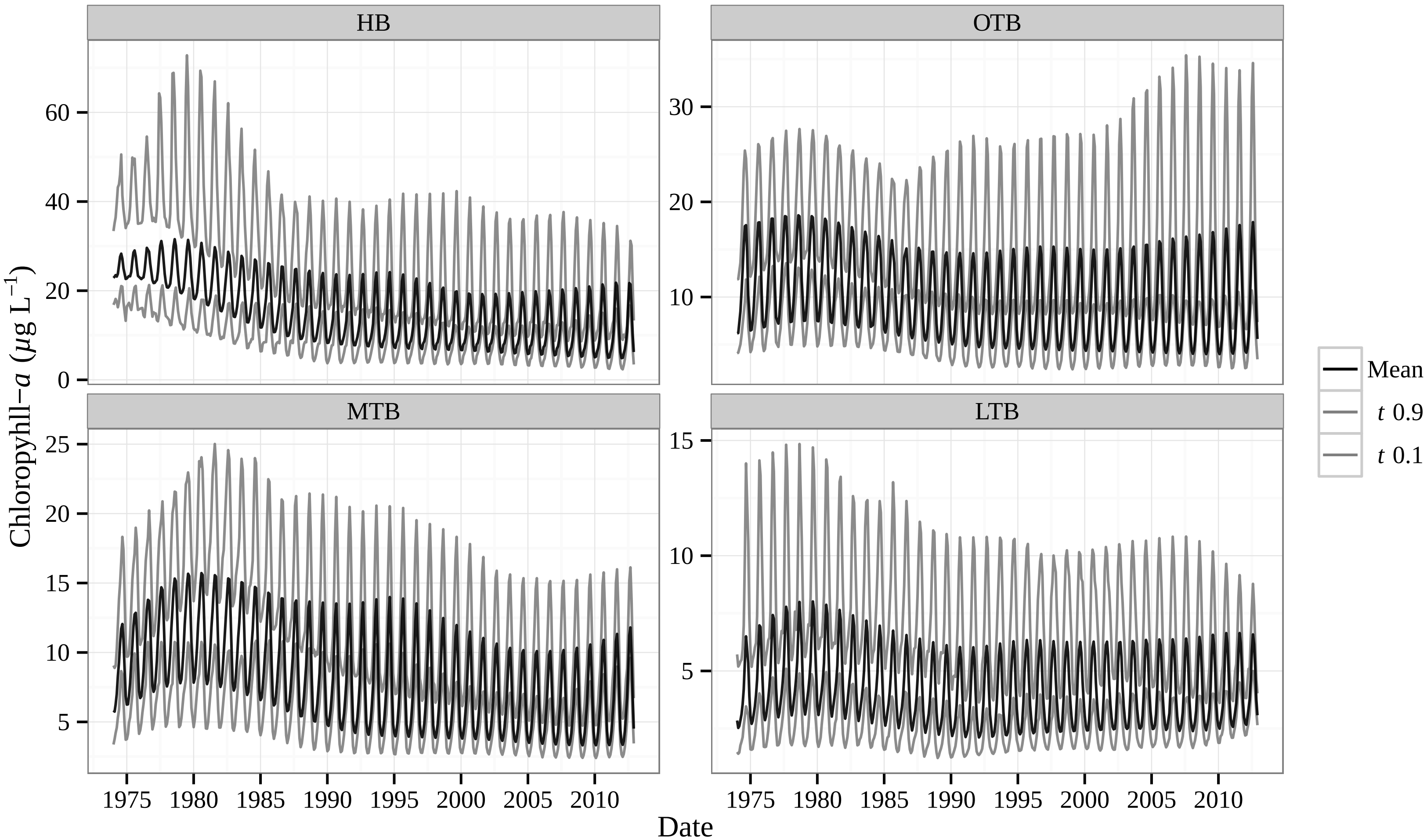


Fig. 7: Salinity-normalized results for the mean response and the 10th and 90th quantile (τ) distributions. Note changes in inter-annual variability by bay segment. HB: Hillsborough Bay, OTB: Old Tampa Bay, MTB: Middle Tampa Bay, LTB: Lower Tampa Bay.

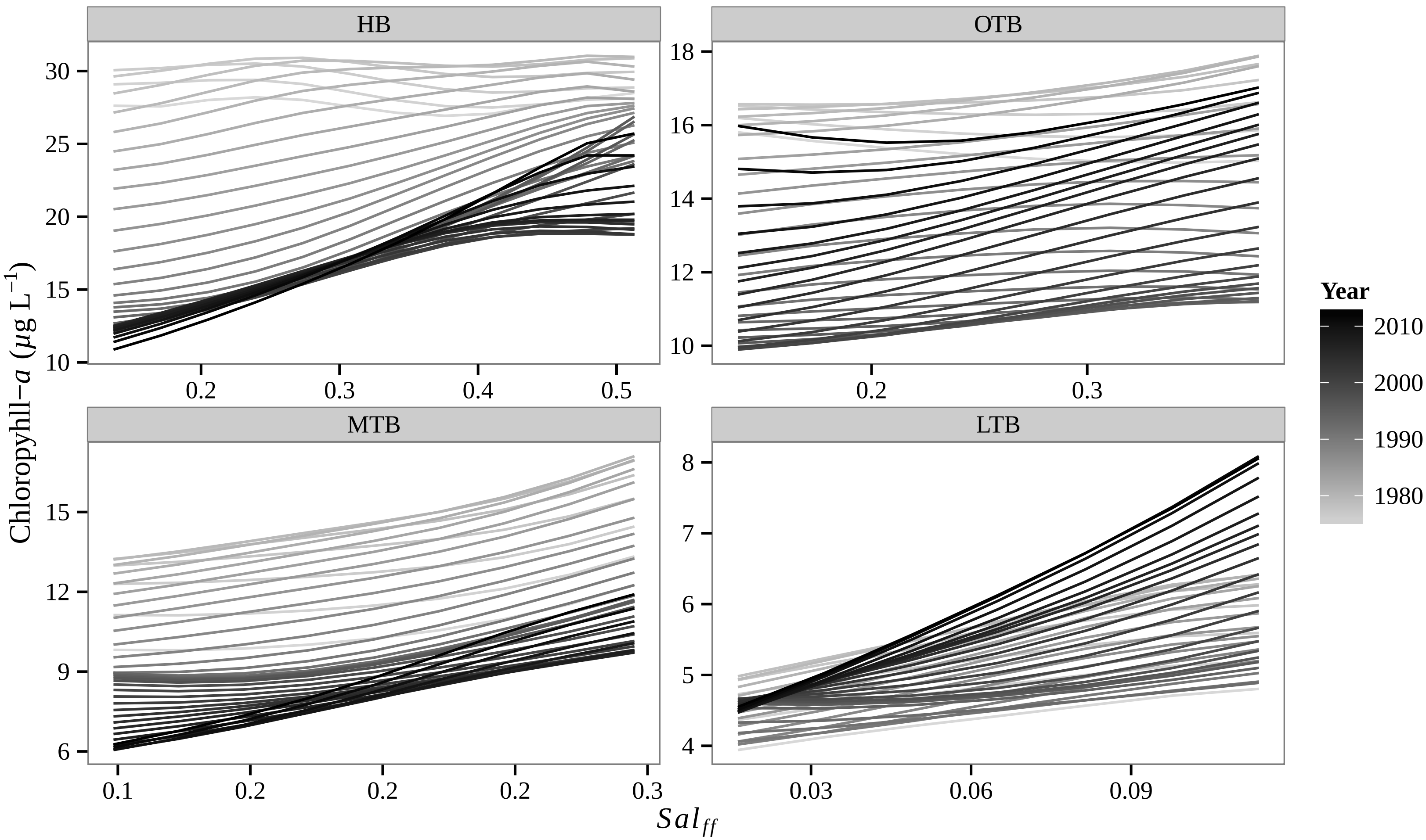


Fig. 8: Variation in the relationship between chl-a and salinity as fraction of freshwater (Sal_{ff}) across time series for Tampa Bay. Data are for July months to reduce seasonal variation. Only the mean response models are shown. HB: Hillsborough Bay, OTB: Old Tampa Bay, MTB: Middle Tampa Bay, LTB: Lower Tampa Bay.

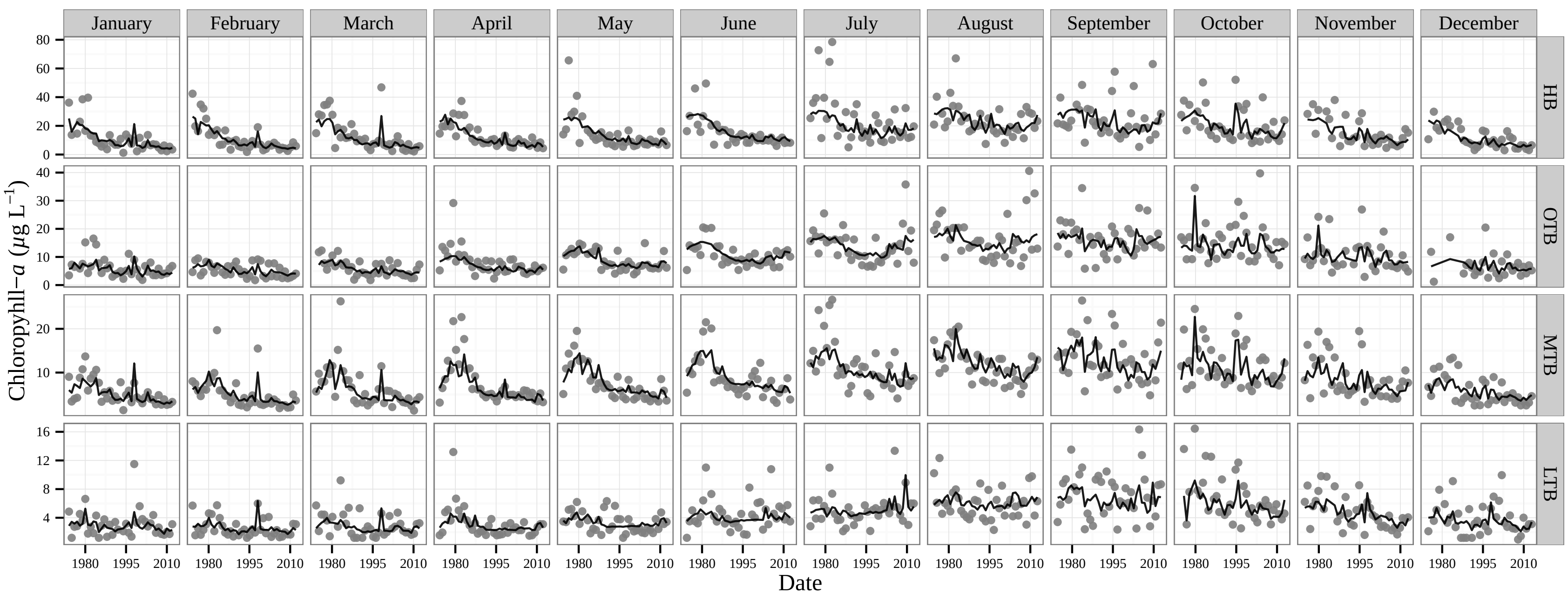


Fig. S1: Predicted (lines) and observed (circles) chl-*a* concentrations for Tampa Bay segments by month. Predicted values are for the weighted regression models fit through the mean response. HB: Hillsborough Bay, OTB: Old Tampa Bay, MTB: Middle Tampa Bay, LTB: Lower Tampa Bay.