Application of Generalized Additive Models for long-term and seasonal trend analysis of water quality: South San Francisco Bay case study

Marcus W. Beck ([mbeck@tbep.org](mailto:mbeck@tbep.org)), Perry de Valpine ([pdevalpine@berkeley.edu](mailto:pdevalpine@berkeley.edu)), Rebecca Murphy ([rmurphy@chesapeakebay.net](mailto:rmurphy@chesapeakebay.net)), Ian Wren ([ianw@sfei.org](mailto:ianw@sfei.org)), Ariella Chelsky ([ariellac@sfei.org](mailto:ariellac@sfei.org)), Melissa Foley ([melissaf@sfei.org](mailto:melissaf@sfei.org)), David B. Senn ([davids@sfei.org](mailto:davids@sfei.org))

Compiled 2020-09-23 16:12:20

# Abstract

# Introduction

Many environmental monitoring programs generate fine-scale but potentially irregular time-series data for the purpose of assessing long-term trends for regulatory, management, or scientific purposes. The mismatch between the scales of monitoring versus analysis questions can lead to statistical challenges. At finer temporal scales, environmental systems may show short-term fluctuations from multiple factors (e.g., weather events or seasonal changes) that may not be of direct interest for longer-term trends. For basic questions, simple smoothing or averaging over time may be appropriate for analyzing longer-term trends. However, simple analyses that aggregate or generalize observations to assess trends can create mismatches between questions of interest and what is returned by the summarized results. Further, many analysis questions are not easily addressed by simple smoothing methods, such as an assessment of trends that occur at different temporal scales.

In this paper we propose analysis methods that separate the signal-extraction of fine-scale data from the trend analysis at longer time-scales, with full propagation of statistical uncertainties between the two analysis stages. For the signal-extraction stage, we use smoothing splines, or the larger framework of generalized additive models (GAMs) that encompasses smoothing splines. While GAMs are a common statistical method and have been used for trend-analysis of environmental monitoring data, alternative and more commonly used approaches can entangle the trend-analysis goal with the signal-extraction goal in the formulation of splines. Analysis questions or management goals that require information about about non-trivial temporal features, such as across-year trends in timing of within-year seasonal events or baseline levels, may not be amenable to a single GAM formulation. For the trend-analysis stage, error propagation in summary statistics can be addressed using mixed-effects meta-analysis, enabling the application of established tools that incorporate uncertainties in model estimates of summary statistics. This two-stage approach is useful because across-year time-scales of interest can be imposed or estimated explicitly, which contrasts with implicit absorption of time-scales into spline formulations in other methods. For example, across-year smoothing methods can be applied to seasonal averages from the signal-extraction stage while accurately propagating uncertainties. On a practical level, the signal-extraction stage can be viewed as creating a “data product” that can be used in second-stage analyses without the difficulty of embedding every long-term trend question in a GAM modeling framework to model the fine-scale fluctuations.

The smoothing methods under discussion here are distinct from older methods of trend detection, such as seasonal Kendall tests or other non-parametric approaches. The former has been used for decades in water quality trend assessments to identify monotonic changes that account for seasonal variation between years. The limitations are such that only the magnitude and direction of a change between selected a priori years can be identified, nor are any quantitative descriptions of changes occurring at different temporal scales explicitly provided. Alternative approaches that decompose a time series into additive or multiplicative components have also been used, such as the seasonal time series composition that isolates a trend, seasonal, and residual component from an observed time series. Although computationally efficient, such approaches may over-simplify trends into absolute components that do not smoothly vary over time, nor do they allow for explicit hypothesis testing. More recently, moving window regression approaches have been applied to dynamically estimate parameters that influence temporal changes in water quality time series. The Weighted Regression on Time, Discharge, and Season (WRTDS) method can describe multi-decadal changes by estimating effects of the predictors that are specific to a given season, year, and flow regime. The approach provides results that are comparable to non-linear, spline-based methods, although at the expense of computation time and uncertainty estimates that are not easily obtainable.

GAMs have only recently been applied to trend analysis of water quality time-series, particularly in the evaluation of multi-decadal time series from long-term monitoring programs. The Chesapeake Bay Program has developed and integrated GAM methods into routine reporting assessments to track and understand long-term changes in nutrients and other eutrophication indicators. Central to their method is flexibility in describing the individual components of the time series with varying levels of complexity, including modeling additive annual and seasonal changes and the interaction between the two. Post-hoc methods were also integrated into the GAM framework to address regionally-specific programmatic questions, such as an assessment of GAM fit before/after a management or policy intervention of interest (e.g., impoundment removal or point-source control). Other studies have addressed specific technical issues of applying GAMs to environmental time series, such as appropriate transformations for response variables, accounting for back-transformation bias, addressing serial correlation in high resolution data, or addressing different time lags in describing relationships between response and predictor variables.

Although our approach is not the first to apply GAMs for trend analysis environmental monitoring data, our approach is distinct from previous applications in several ways. First, some approaches use the GAM smoothing functions to approximate components of a time series to estimate average conditions for a finite period of time. This can represent a mismatch between processes occurring at different time scales and does not make sense as an estimation method if, for example, a seasonal treatment factor is not an explicit component of the time series. Second, some approaches formulate GAMs with multiple spline terms to decompose time-scales of fluctuations within the boundaries of the GAM estimation methods. There may be some analysis questions for which this is appropriate, but it is less general than our approach and it goes against caution of Wood (2nd edition book) that the degrees of smoothness attributed to multiple splines can be confounded. Perhaps more important, this approach risks obscuring the different time-scales implicit in choices such as an appropriate number of spline knots, leaving the rendering results more difficult to interpret. Third, over-confidence in conclusions derived from insufficient treatment of model uncertainty, including propogation to secondary analyses, can lead to misinformed decisions. GAM applications to environmental monitoring data have not fully treated error propagation as a fundamental component of model application to decision-making, particularly when wrong decisions can lead to expensive consequences.

We illustrate the proposed methods with the motivating example of water quality monitoring in the southern portion of San Francisco Bay, California, USA. For several decades, approximately twice-monthly monitoring has been conducted at fixed locations (stations) of the longitudinal axis of the Bay. Analysis of these data is complicated by irregularities in timing and consistency of data collection, such that simple seasonal averages of raw data may not adequately describe trends. Examples of long-term trend questions include: Are there significant trends in spring mean chlorophyll at four-year (or other) time-scales? At what across-year scale do within-year summer-fall mean chloropyll levels change? Are there significant across-year trends in within-year timing of the spring bloom in chlorophyll or in baseline levels during periods of low productivity? We also contrast different GAM formulations that can yield similar signal-extraction results but carry somewhat different interpretations. The approach we give for using meta-analysis methods as a second stage of analysis following signal extraction with GAMs appears to be new to environmental trend-detection problems. For this step, we give methods for isolating seasonal trends secondarily from GAM results with reasonable certainty and evaluating these trends between year.

# Methods

## Study area and data sources

The San Francisco Estuary (SFE) is the largest estuary on the Pacific Coast of North America and drains an area of approximately 200 thousand km in the US state of California. Major freshwater inputs enter the system through the Sacramento-San Joaquin Delta complex upstream of Suisun Bay, where the combined inflow from both rivers is approximately 28 km per year. The northern subembayments are river-dominated (salinity ranging from 0 to 15 ppt), whereas the southern subembayments are marine-dominated with salinity ranging from 5 to 35 ppt depending on the tidal cycle, effluent discharge from wastewater treatment plants, and stormwater runoff. The South Bay embayment is heavily urbanized and includes thirty-seven wastewater treatments plants that serve 7.2 million people. Secondary treatment occurs at a majority of the treatment plants and the remaining effluent is discharged into the SFE. Agricultural runoff from the upper watershed also contributes to nutrient loading in the SFE with the annual nutrient export estimated as approximately 30 thousand kg dy of nitrogen from the Delta.

Nitrogen and phosphorus levels in SFE generally exceed concentrations that have been observed to promote excess primary production in other large estuarine systems. However, eutrophic conditions have not been regularly observed since routine monitoring began in the 1970s. Historical resistance of SFE to eutrophication has been attributed to several factors, including elevated suspended sediments that reduce light penetration in the water column, regular exchange and mixing with low-nutrient marine waters and export of estuarine nutrients to the Pacific Ocean, and benthic grazing by filter-feeding bivalves that reduce algal concentrations. Renewed interest in the potential for nutrient loading to negatively affect water quality has occurred recently, particularly in South Bay, where harmful algal blooms (HABs), increases in summer-fall chlorophyll concentrations, and low dissolved oxygen concentrations beginning in 1999 (Figure 1) (Cloern et al. [2020](#ref-Cloern20)). Although visual changes in observed data are apparent, statistical analyses to quantify current status and to provide estimates of annual and seasonal trends with appropriate bounds on uncertainty have not been sufficiently developed, particularly on a seasonal basis.

The analysis evaluated near-surface chlorophyll-a data collected biweekly to monthly along the South Bay axis extending from Central Bay (stations 18-23), South Bay (stations 24-32), and Lower South Bay (stations 34-36) (Table 1, Figure 2). Monitoring data were obtained from the SFE Research Program of the US Geological Survey (Cloern and Schraga [2016](#ref-Cloern16), Schraga et al. [2020](#ref-Schraga20)). Discrete chlorophyll concentrations at each station were determined by fluorimetric analysis with 90% acetone pigment extraction on GFF filters. Data collected between 1990-2017 were selected for analysis because it represented a suitable balance among three factors relevant to testing the statistical approaches, including sufficient length of record, consistent biweekly-monthly sampling, and a diverse set of stations covering the salinity gradient across multiple subembayments. While sampling frequency varied somewhat over time or by station, all data were treated as unique time series within the statistical models (i.e., no spatial or temporal binning or averaging was done).

## GAM application

The basic structural form of a Generalized Additive Model is (Hastie and Tibshirani [1990](#ref-Hastie90)):

where the expected value of a dependent variable conditional on predictors X through X is the sum of smoothing functions for each predictor plus an intercept term . The smoothing functions are standardized and have expected values equal to zero so that . Because GAMs are an extension of Generalized Linear Models, they can also include linear predictors in addition to the smoothed functions. All methods herein use functions provided by the *mgcv* R package to fit GAMs using multiple parameter smoothing estimation methods (Wood [2017](#ref-Wood17)).

To evaluate GAMs with SFE chlorophyll data, we followed an approach that built on past work by the Chesapeake Bay Program (Murphy et al. [2019](#ref-Murphy19)). We specifically build our models using the gam1 and gam2 structures in Murphy et al. ([2019](#ref-Murphy19)), but present a different naming convention to better distinguish the functional components of the models. The gam1 and gam2 models are shown in equations (2) and (3), denoted as “Constant season” and “Interactive season”, respectively. We also describe the "Constant season\*" model below as a slight modification of the former. Each model has a similar functional form for time in relation to chlorophyll with slight differences in the smoothers that describe time.

For each GAM, chlorophyll was modelled using a link function that relates the response variable to the independent variables and day of year () (Table 2, Figure 3). The variable is a continuous measure for date in decimal time, where date is expressed as a continuous numeric value with year on the left of the decimal and date within the year to the right of the decimal (e.g., July 1st 2019 would be 2019.5). The variable is an integer value for the Julian day in each year (from 1 to 366) that represents the seasonal component of the time series. The estimate is an intercept value and is a slope estimate for a linear year effect included in all models. The and functions are different smoothers for and where the former uses a thin-plate spline and the latter uses a cyclic cubic spline to capture the periodic seasonal component (Wood [2003](#ref-Wood03)). The function is a tensor product interaction between and that allows the seasonal component to change between years. Within the function are the two separate smoothers for year as a thin-plate spline and as a cyclic cubic spline. As with conventional linear models, the remaining variance in the response not captured by the functional form of the model is expressed as an error term that is normally-distributed having mean equal to zero and variance as . The residual variance is used to back-transform model estimates for trend testing, described below.

A challenge for optimizing GAMs to time series data with significant intra- and interannual variation is choosing an appropriate functional form (i.e., equation (2) or (3)) and determining how much variation could be explained by the smoothers in each function. The two models above provide tradeoffs in the functional forms that balance descriptive and computational efficiency (Table 2, Figure 3). However, the individual smoothing functions (i.e, , , and ) require user input on the potential upper limit for how much variation could be explained by the predictors during model fitting. Each smoother is approximated by multiple, localized spline functions that are connected by knots (*k*-values) to create a continuous function. Increasing the number of knots creates a smoother fit of the response variable against the predictor, whereas choosing a small number of knots creates a more rigid function that may underfit the data. Initial testing by Murphy et al. ([2019](#ref-Murphy19)) showed that the default *k* values in the mgcv package were insufficient for describing the interannual variation in chlorophyll data from Chesapeake Bay. The upper limit for the number of knots for the smoother was modified from the default value to the maximum between 10 or the number of years in the time series. We follow the same approach herein both models. Further, the default *k*-values for and were considered appropriate and not tested further. For the application to SFE chlorophyll data, an extension of the constant season model was also added as "Constant season\*", where the total number of knots for the smoother, , was increased to 12 times the total number of years in the time series. This approach potentially allows the smoother for to be unconstrained in how much variance is fit within and between years.

The link functions for each model were chosen as either logarithmic transformations (base 10) or best estimates of Box-Cox power transformations for chlorophyll. A comparison of GAM performance between transformation methods has not yet been explored in the application of GAMs to long-term water quality monitoring data. The logarithmic transformation is commonly used for chlorophyll to approximate a log-normal distribution of the response variable. However, alternative transformations could provide improved model fits and reduced uncertainty estimates in the fitted parameters by better satisfying assumptions for GAMs. The Box-Cox method was also used to transform chlorophyll as a comparison to the logarithmic transformation (Box and Cox [1964](#ref-Box64)). The Box-Cox method requires an estimate of the parameter , whereby the the optimal value is based on a minimization function of a log-likelihood profile vector. The profile vector for each chlorophyll time series was obtained using the *boxcox* function from the MASS package (Venables and Ripley [2002](#ref-Venables02)) to evaluate log-likelihood estimates of across the range -4 to 4. Once the optimal value was identified, chlorophyll was transformed using the following power transformation (Box and Cox [1964](#ref-Box64)):

where is the transformation function for the appropriate GAM at the estimated value. Note that in some cases the Box-Cox transformation is a logarithmic transformation if is estimated as zero. The *BoxCox* function from the forecast package (Hyndman et al. [2020](#ref-Hyndman20)) in R was used to transform the chlorophyll time series with the optimal value. Equation (4) is only appropriate for positive values of a response variable, as for chlorophyll.

All models were compared using standard summary statistics describing overall model fit to the observed data. In addition to the Generalized Cross-Validation score provided by the mgcv package, models were compared using the Akaike Information Criterion (AIC) and summmary statistics. Comparisons were made between the different GAMs and for the transformations used to define the function for chlorophyll at each station. Model predictions were based on the standard model output from the *predict()* function from the mgcv package. A long-term trend independent of seasonal variation was also estimated for each model by subtracting the seasonal terms (i.e., and ) from the predictions. Back-transformation of results were obtained by exponentiation if the log-transformed variable was used or using the *InvBoxCox()* function from the forecast package if the Box-Cox transformation was used. For hypothesis tests that required estimates of central tendency and confidence intervals for mean values, back-transformations were based on an exponentiation or inverse function of the transformed response variable as before, with the addition of a correction factor. This bias-correction factor was simply the variance of the model residuals divided by two for variables in log-space and correction with a forecast variance estimate as specified in the *InvBoxCox()* function from the forecast package.

## Using GAMs to evaluate trends

Each of the GAM structures provided different approaches to quantify long-term variation in annual and seasonal changes. Although useful as a general approach to view functional forms for different time series, additional information on whether an observed change is statistically significant can be useful to inform environmental management decisions. We present two methods for performing trend tests on model results that address different changes that may be of management interest. First, a method for evaluating the percent change in annual averages between two time periods is briefly summarized (Murphy et al. [2019](#ref-Murphy19)). Second, a method for evaluating means and confidence intervals from GAM results for seasonal periods is presented, including an estimate of direction and rate of seasonal change over time.

First, for two time periods of interest, trends were estimated as the difference in means from the GAM results between the periods, an estimate of percent change, and an estimate of statistical certainty of the comparison. Murphy et al. ([2019](#ref-Murphy19)) defined the two time periods as “base” and “test” periods, where the former represents a baseline condition earlier in the time series as a basis of comparison to the more recent “test” period. The differences are estimated as:

where is a point estimate of the differences of the predictions between the base and test years and is the estimated parameter vector from the fitted GAM. The standard error of this difference is then:

defined as the square root of the product of , the estimated variance-covariance matrix of parameter vector for the fitted GAM , and the transposition of . The trend is considered significant if the appropriate confidence interval at a specified does not include zero.

The second trend test evaluated changes in seasonal estimates between years using results from a fitted GAM. This test provided both an estimate of a seasonal mean, its associated confidence interval, and if the rate of change in the seasonal means has varied over a specific period of interest. For example, changes that occur during late fall or early spring can be evaluated between years to determine if the seasonal averages differ and if the direction and magnitude of change is significant. This has implications for understanding potential changes in seasonal bloom phenology resulting from shifting system responses to nutrient inputs or changes in the latter.

Seasonal averages and confidence intervals were estimated from fitted GAMs by first creating a matrix of the predictors (i.e., decYear, doy) covering defined seasons in the time series. For example, seasonal changes from late summer could be obtained from a prediction matrix covering all dates from Aug 1st to Sep. 30th for the period from 2000 to 2010. The prediction matrix is then used to obtain fitted values from a GAM that describe the estimated values of chlorophyll that occurred during the periods of interest. Rather than the estimated values for the predictions, a linear predictor matrix was obtained that provided a parameter vector for the supplied covariate values. This predictor matrix was then used to estimate the mean seasonal value within a year and associated variance using the sum of predictions from the matrix. As a result, each year had a mean value of chlorophyll with an associated confidence interval that describes the modelled expectation for the seasonal period.

Although the seasonal averages obtained from GAMs described the expected values for each year, describing the direction and magnitude of the seasonal change across years was also of interest. Conventionally, trend analysis of changes in a response variable could be achieved with ordinary least squares regression, where the response variable is the expected seasonal value of chlorophyll and the predictor is year as a numeric value. Although this approach could be used to estimate a seasonal rate of change, the linear fit through the seasonal means does not account for the variance in the seasonal averages. To address this issue, we use mixed effects regression meta-analysis (mixed-meta hereafter) to evaluate seasonal changes and their uncertainty over a finite set of years. This analysis is an extension of conventional regression by accounting for uncertainty attributed to both variation of the true seasonal mean around the trend and variation of the estimated seasonal mean around the true seasonal mean:

where the estimated value of chlorophyll at year for the predetermined season is a function of the predicted seasonal mean plus deviation of the seasonal mean from the trend () and deviation of the seasonal mean from the true mean (). Both the error terms are assumed to be normally distributed with mean zero and variance for equal to an unknown term and variance for equal to the squared standard error from the seasonal chlorophyll averages. The R package *mixmeta* was used to estimate trends in the seasonal averages over time (Sera et al. [2019](#ref-Sera19)). Results have a similar interpretation as those from regression analysis (i.e., slope and intercept estimates), although parameter uncertainty is more accurately estimated with mixed meta-analysis.

## Model application and interpretation

Each of the three GAM structures (constant season, interactive season, and constant season\*) were used to fit models to the chlorophyll time series at each of the nine stations from central to lower South Bay. For each model and station, the fit was assessed using the standard summary statistics described above (i.e., GCV, AIC, and ) and relative significance of the individual smoother terms. Models were also evaluated using different transformations to determine which had improved fit for specific time series. Trend analyses with fitted GAM predictions were also compared between stations and model types, particularly to determine if and when significant results varied with alternative methods. Temporal and spatial patterns within and between stations were also identified to better quantity changes in chlorophyll production throughout the period of record using GAMs to demonstrate improved descriptions relative to interpretations from the observed data.

# Results

## Model performance and observed results

A total of six models (three GAM structures, two link functions) were fit to each of nine stations, producing summary statistics for 18 different models (Table 3). Chlorophyll trends were generally well-explained across all models and stations, with an average R-squared value equal to 53% and ranging from 37% (Const seas log-model at station 34) to 77% (Const seas\* log-model at station 32). The constant season model had the worst performance for all stations (average R 44%), with either the interactive season or constant season\* model having improved measures of model fit. For all but stations 18 and 22, the constant season\* model provided significantly lower AIC and GCV scores and higher R-squared values compared to the interactive season models. For station 18, the interactive season and constant season\* models provided only a marginal improvement over the constant season model. For station 22, the interactive season and constant season\* models provided similar but much improved fit to the data as compared to the constant season model. Moeel performance comparing either Box-Cox or logarithmic link functions were similar (average R values of 54% and 53% respectively), although the logarithmic models had slightly improved model performance. Station 18 was the only location where the three GAM models were not significantly different from each other for the Box-Cox transformation.

GAM predictions from north to south on the longitudinal axis provided different descriptions of annual and seasonal changes in chlorophyll. Four representative stations are shown in Figure 5 that capture spatial variation in chlorophyll and different temporal trends described by the models. All models suggested an increase in chlorophyll in the period of record from 1990 to approximately 2005 to 2010, when predictions suggested a decreasing trend until the end of the record. Seasonally, all models indicated a larger spring peak in chlorophyll with more southern stations, whereas a subsequent peak in fall concentrations did not have any noticeable difference in magnitude by location. However, the magnitude of the fall peak at more southern stations shifted over time depending on the model. In particular, the interactive season model indicates a shift in the magnitude of the fall peak at stations 24, 27, and 32 with maximum concentrations observed around the mid-early 2000s. The annual trends for all models suggested similar trends over time, i.e., an increase followed by a recent decrease, whereas the constant season\* models with an unconstrained smoother on the decYear predictor varied considerably at station 24, 27, and 32. Annual trends at these stations were difficult to generalize and had substantial overlap with the seasonal trends. Conversely, station 18 had minimal visual differences in annual and seasonal trends across all models, which is reflected by similarities in the model performance statistics in Table 3.

## Trend estimates

Examples of results provided by the trend tests are shown for station 34, with estimates of percent change between year pairs and seasonal trend results for January to July and August to December (Figure 5). This figure illustrates the narrative descriptions that can be obtained for each test and the differences that can be observed depending on season. Figures 5a-c demonstrate that average chlorophyll at station 34 from 1990 to 2000 had a 57.7% increase, although the change is insignificant. However, the change from 2000 to 2010 was statistically significant with an increase of 78%. The period from 2010 to 2017 showed a significant decrease of 39.8%. An assessment of the annual differences indicated if changes overall were observed, but provides no information on seasonal shifts that may have contributed to these differences. The results in figure 5d-i show results from mixed-meta regression analysis to indicate which seasonal changes were contributing to the annual differences. Rows d-f show the seasonal averages and trend estimates for the same year periods in rows a-c, but only for months from January to July. Similarly, rows g-i show the seasonal averages and trend estimates from August to December. The seasonal trend results suggested that the increase in annual averages from 2000 to 2010 was likely caused by an increase in chlorophyll values in the fall (Figure 5h). Seasonal changes from 2010 to 2017 showed that neither the January to July nor August to December changes were significant, suggesting an overall decrease in chlorophyll across all seasons may have recently occurred.

Spatial variation in trends suggested that most stations in South Bay had increasing chlorophyll from 1990 to 2010 following a decrease in recent years, although many trends were not significant and varied by location (Figure 6a). From 1990 to 2000, only station 18 and 22 in the northern part of South Bay had significant increasing trends. The magnitude of the chlorophyll increase in the southernmost stations (34, 36) was higher during this same period, although the trends were not significant. From 2000 to 2010, significant increases were observed at the three northern-most stations (18, 21, and 22) and two southern stations (32, 34). From 2010 to 2016, all four southern stations had significant decreases in annual chlorophyll. The spatial trends were further evaluated with mixed-meta regression analysis to provide additional context to the annual changes by evaluating trends from January to July and August to December (Figure 6b). For the January to July period, significant changes were only observed at the southern stations, with increases at station 34 and 36 from 1990-2000, an increase at station 34 from 2000-2010, and a decrease at station 32 form 2010 to 2016. No significant changes were observed for the August to December period.

A final analysis of the seasonal trends evaluated more discrete monthly periods in three month blocks to provide finer resolution to explain trends from the GAM results (Table 4). Many stations and time periods did not show any seasonal chlorophyll trends or only had one significant trend (e.g., stations 21 and 24, JFM period from 2000-2010). However, several trends were observed in the southern stations, with two significant trends at station 32, six at station 34, and three at station 36. Trends at station 32 indicated a positive increase in the JFM (January, February, March) period from 2000-2010, followed by a significant decrease for the same seasonal period from 2010-2016. Station 34 had significant increases in the JFM period from 1990-2010 and all other periods the OND (October, November, December) period for 2000 to 2010. Significant decreases were observed at station 34 in the JFM and AMJ (April, March, July) period from 2010 to 2016. Station 36 had significant chlorophyll increases in all seasons, except OND for the 1990 to 2000 period. Accordingly, table 4 demonstrates the value of obtaining greater insight into seasonal changes by evaluating different blocks as compared to the larger seasonal blocks shown in Figure 6b.

## Trend comparisons

Trends results from the mixed-meta regression method for each season and different time periods were compared to alternative trend analyses to demonstrate different conclusions that can be derived depending on method. Figure 7 shows estimates of chlorophyll change per year for the same seasonal periods and annual groupings as in figure 6b. In addition to mixed-meta analysis, two alternative trend methods were used where the first shows slope estimates from a linear regression on annual averages of observed chlorophyll for the specified seasonal period and the second shows slope estimates from a linear regression model on annual averages derived from the GAM seasonal estimates. The latter uses the same averages from the mixed-meta analysis, but is a simpler approach that does not account for the uncertainty in the average estimates.

The different trend analysis methods provided conflicting information on the magnitude and direction of the seasonal chlorophyll changes in each decade (Figure 7). The slope estimates from the linear model applied to the observed data were understandably more variable than the slope estimate from the average and mixed-meta methods, with much larger slopes observed especially at the more southern stations. Slope estimates from the linear model applied to the averages from the GAMs as compared ot the mixed-meta results were identical at almost all stations, excluding station 32 in the January-July period from 1990-2000 when the former showed a negative slope and the latter showed a positive slope. The mixed-meta analyses had less significant trends compared to the average results for the same GAM estimates, which reflects the ability of the former to account for uncertainty in the average estimates when considering trends. Importantly, completely opposite and/or varying significance of trends were observed between the methods, suggesting choice of analysis method can influence the certainty of conclusions.

# Discussion

# Conclusions

# Acknowledgments

# Figures

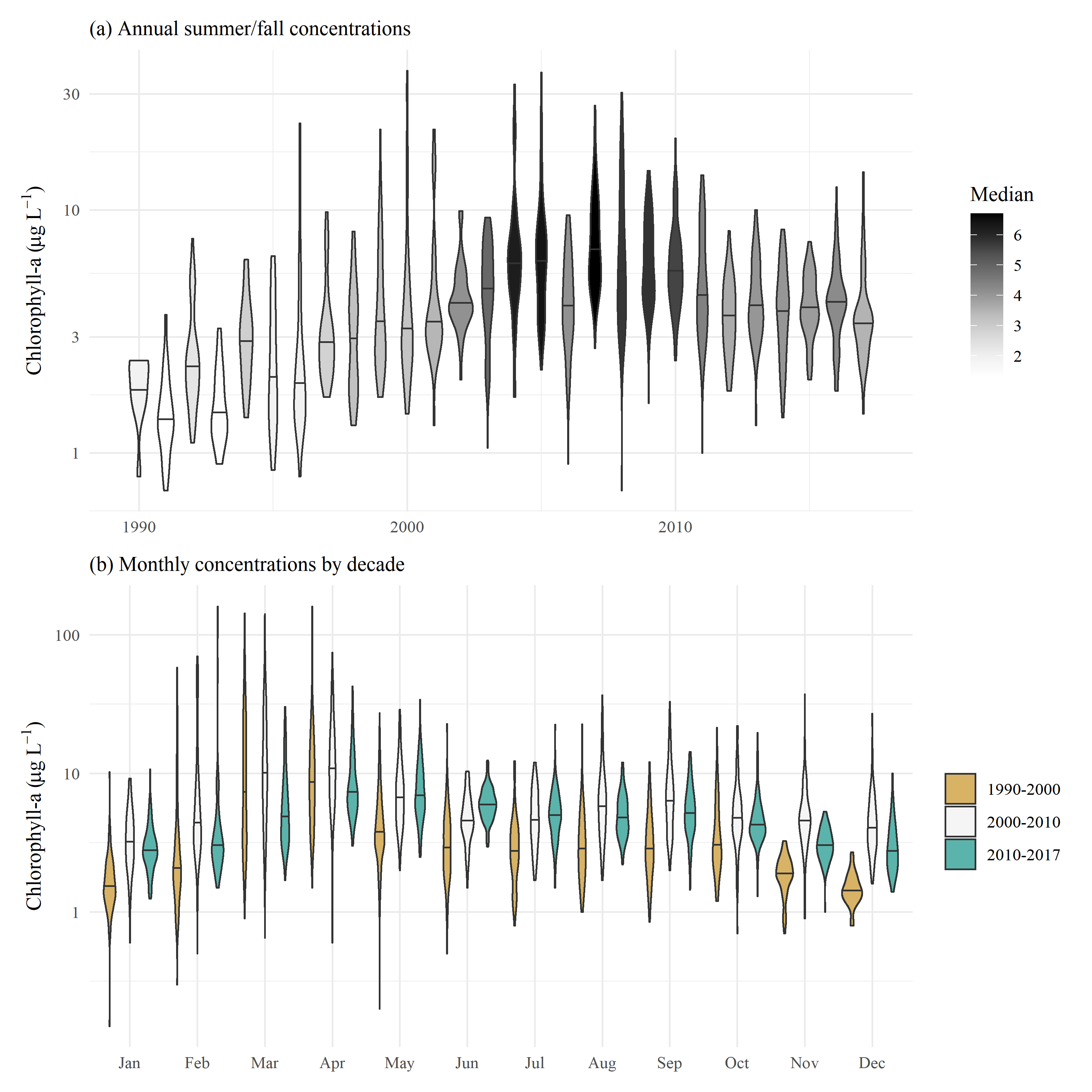


Figure 1: Observed chlorophyll concentrations for all stations in central and south San Francisco Estuary (18-36, Figure 2), with (a) annual summer/fall concentrations (Aug - Dec) and (b) monthly concentrations by decade.

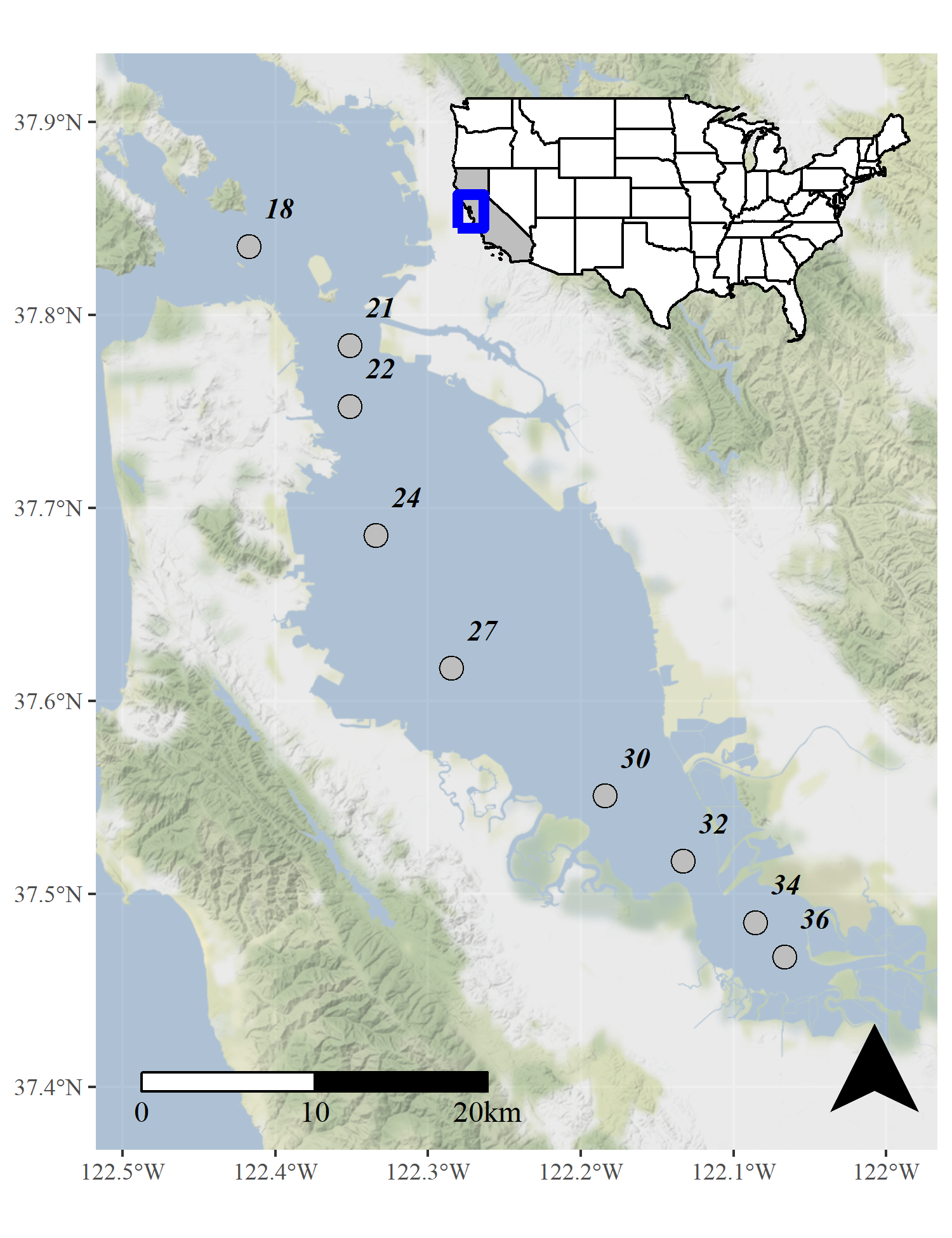


Figure 2: Station locations in the central and south San Francisco Estuary used for analysis. See Table 1 for station descriptions. Full dataset described in Schraga et al. ([2020](#ref-Schraga20)).

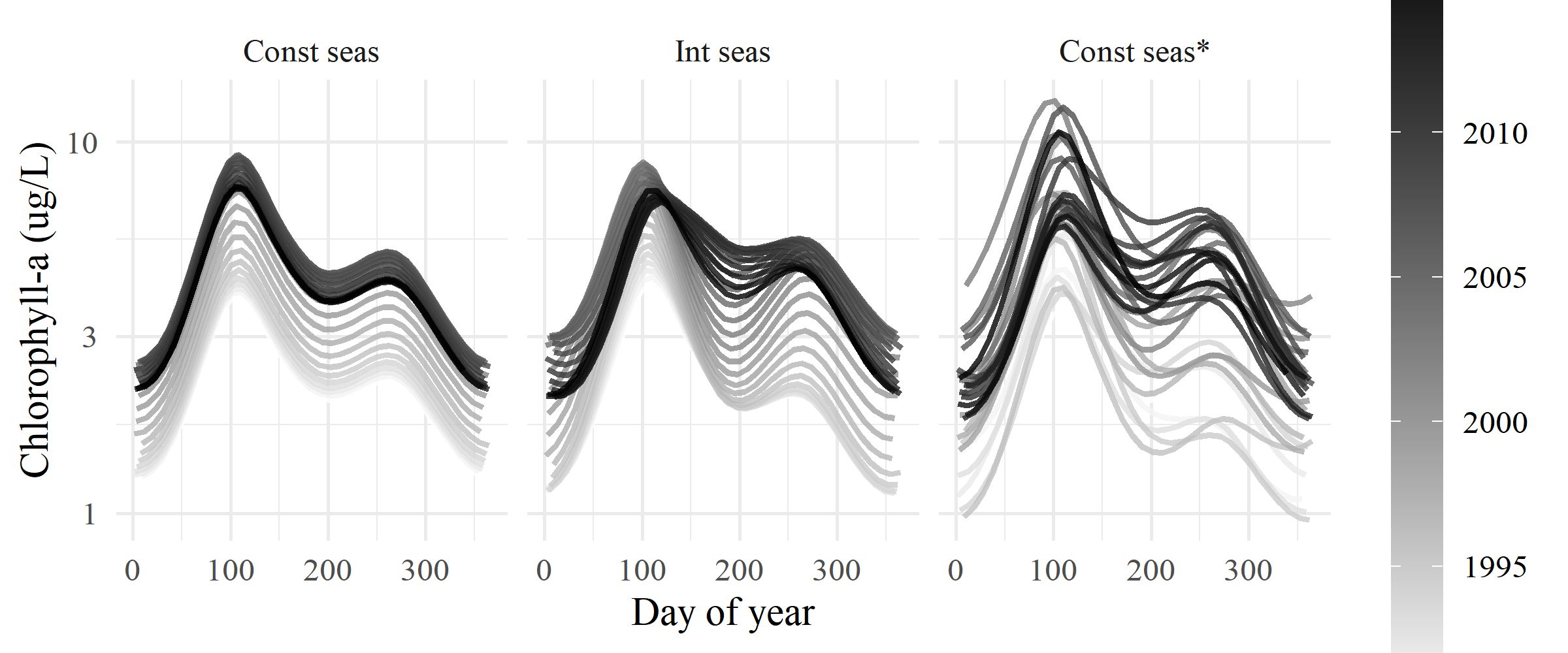


Figure 3: General output for each of the GAM structures for constant season, interactive season, and constant season\*. Day of year (Julian) is on the x-axis and each curve represents a different fitted year for chlorophyll concentration.

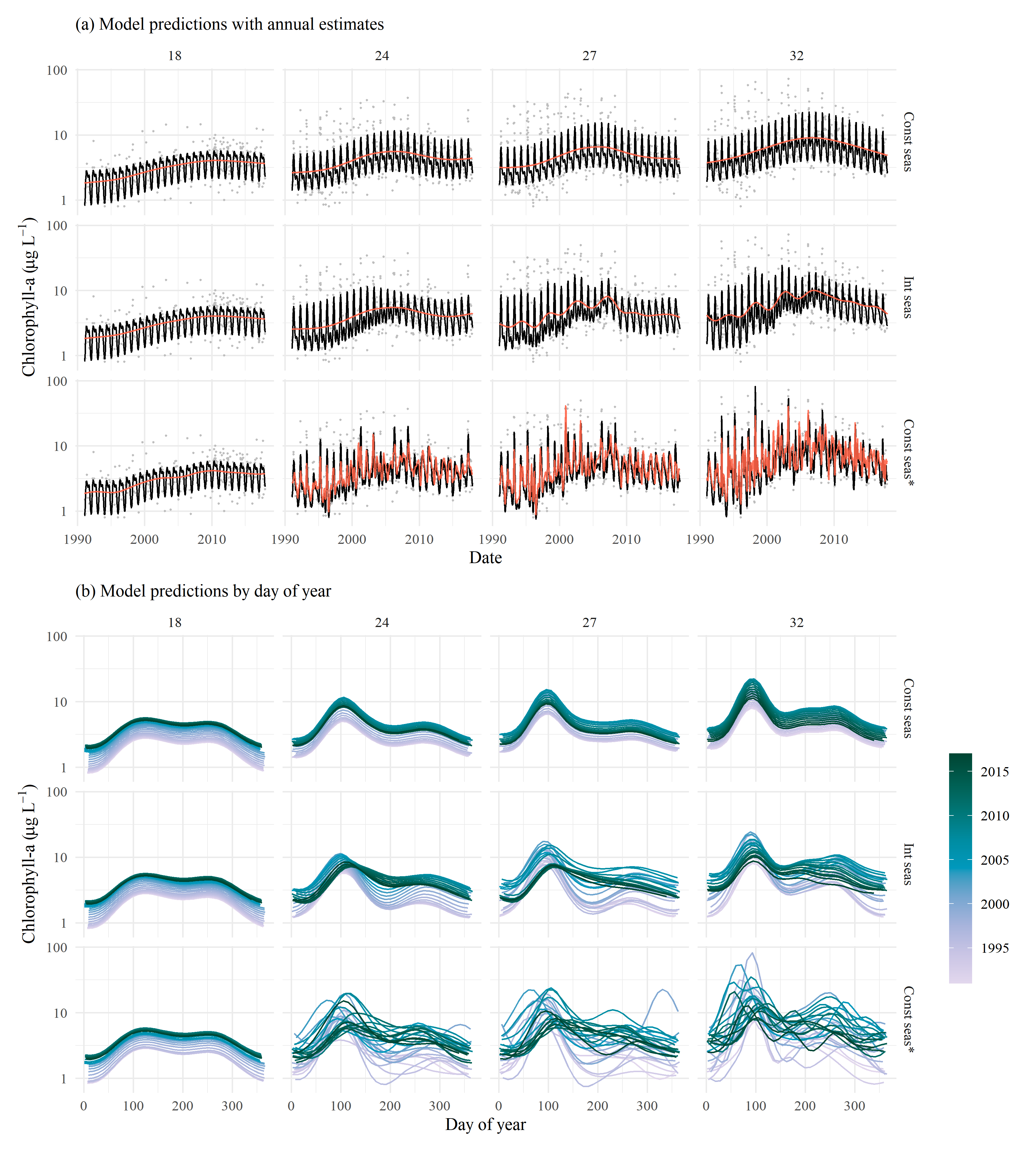


Figure 4: GAM predictions for select stations from north to south (18, 24, 27, 32) for each of the GAM structures for constant season, interactive season, and constant season\*. The results show (a) predictions across the time series with annual estimates (red line) and (b) predictions by day of year with one smooth per year. Observed data in (a) are shown with the gray points. Station locations are in Figure 2.

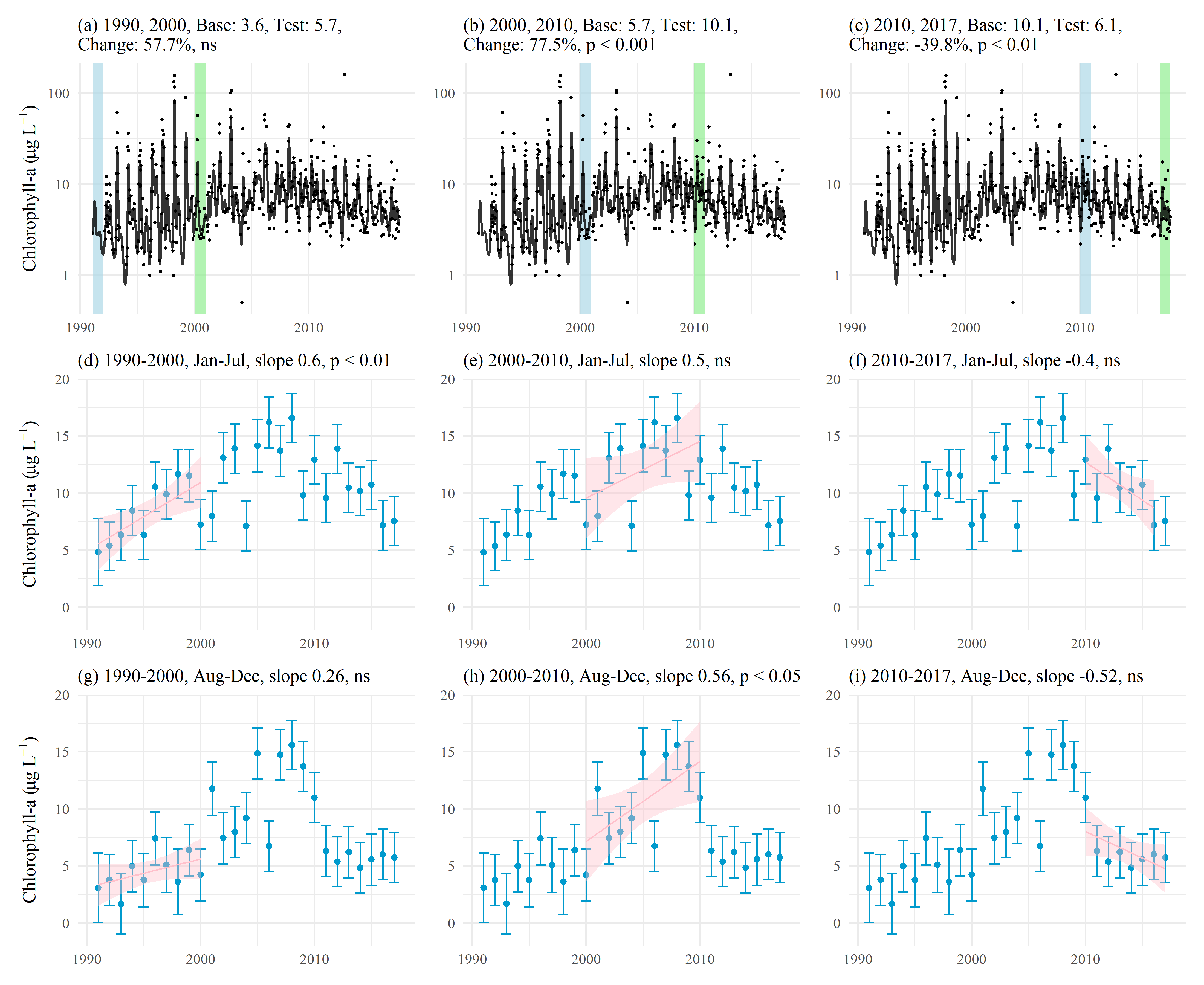


Figure 5: Examples of trend tests using results of GAM predictions for station 34 (Const seas\*). Plots (a), (b), and (c) show estimates of percent change as the difference of means between base and test years for chlorophyll. Plots (d) through (i) show seasonal averages and trend estimates over time. Plot (d), (e), and (f) show trend estimates for January through July and (g), (h), and (i) show trend estimates for August through December. The trend lines in (d) through (i) estimate the rate of change of chlorophyll per year, reported as the slope in the plot title. ns: not significant

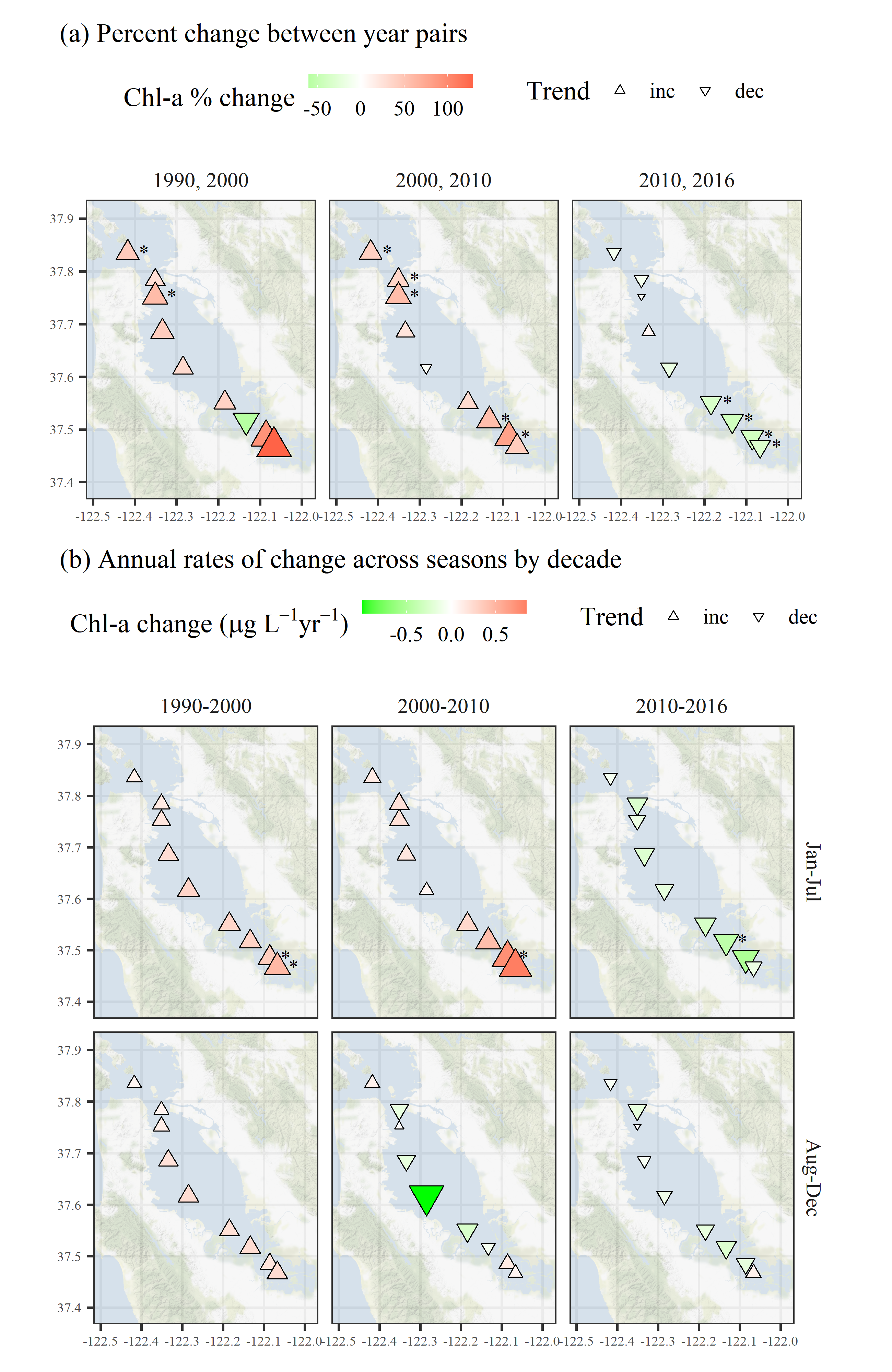


Figure 6: Trend estimates (a) between year pairs and (b) across seasons by decade for chlorophyll at each station. The top plots show percent change comparing the first and last year within a decade and the bottom plots show seasonal estimates of change per year for chlorophyll concentrations for each decade. Point type, shape, and color represent the direction and magnitude of an estimated trend. Trends with are marked with an asterisk. All results are from Const seas\*.

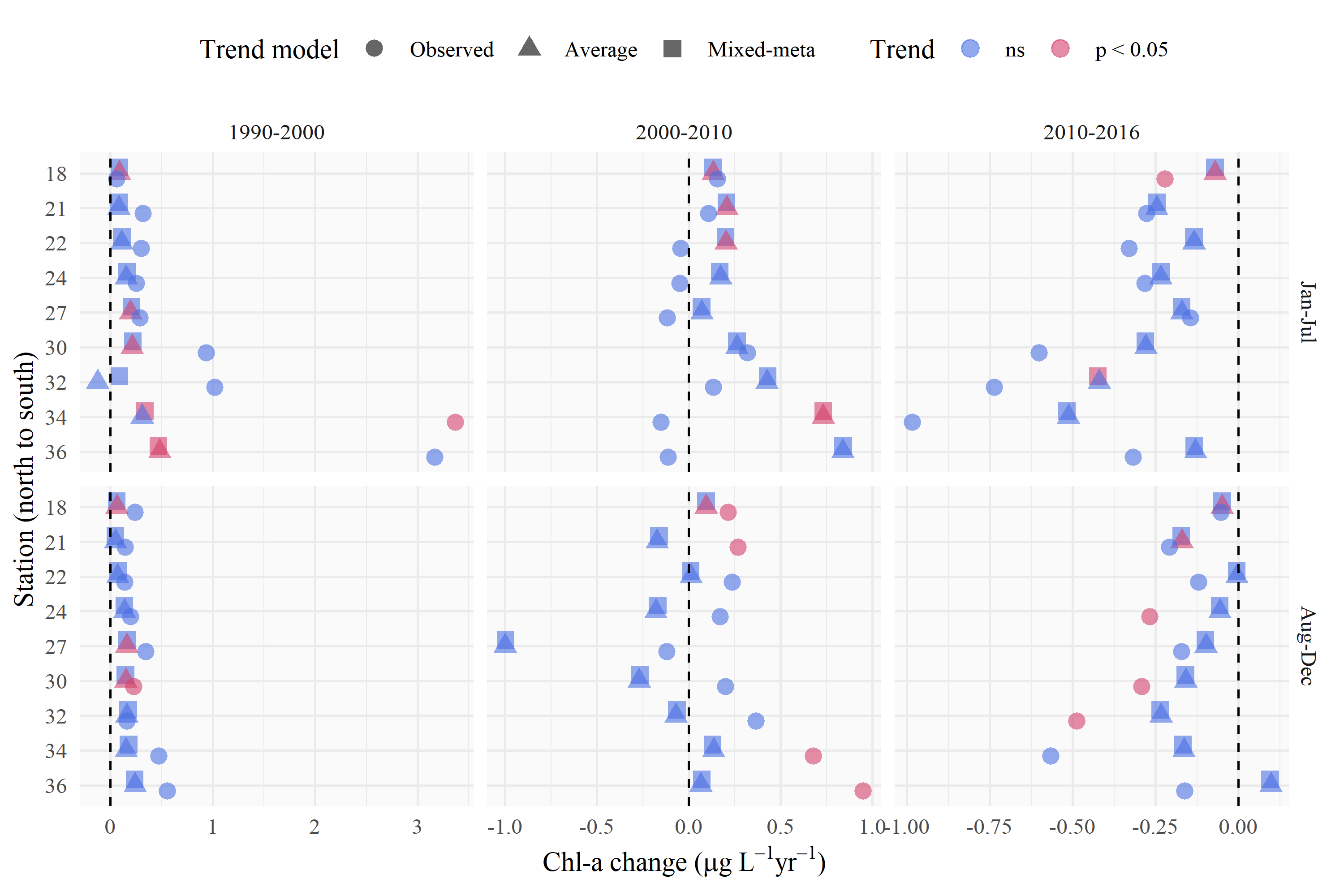


Figure 7: Trend estimate comparisons for three models applied to seasonal averages of chlorophyll in different annual periods at each station. The “observed” trend model is based on a linear fit to the annual averages of chloropyll from the observed data, the “average” trend model is based on a linear fit to the annual averages of chlorophyll from the GAM constant season\* model, and the “mixed-meta” trend model is based on a mixed-meta regression model fit to the annual averages of chlorophyll from the GAM constant season\* model. Values for each model are the slope esimates as annual change per year within each season, with color denoting significant trends.

# Tables

Table 1: Station locations, sample sizes, and median values for salinity, dissolved inorganic nitrogen (DIN, sum of ammonia, nitrate, and nitrite), and chlorophyll. Rows are arranged from north to south.

|  |  |  | Salinity (psu) | | DIN (mg/L) | | Chl-a (ug/L) | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Station | Latitude | Longitude | n | Med. | n | Med. | n | Med. |
| 18 | 37.836 | -122.418 | 373 | 29.9 | 276 | 23.3 | 364 | 3.2 |
| 21 | 37.784 | -122.351 | 535 | 28.4 | 308 | 27.3 | 525 | 4.2 |
| 22 | 37.752 | -122.351 | 537 | 28.3 | 145 | 24.3 | 527 | 3.9 |
| 24 | 37.686 | -122.334 | 555 | 28.0 | 363 | 29.5 | 544 | 4.1 |
| 27 | 37.617 | -122.285 | 557 | 27.2 | 370 | 31.3 | 543 | 4.4 |
| 30 | 37.551 | -122.184 | 568 | 26.6 | 383 | 33.0 | 553 | 5.2 |
| 32 | 37.517 | -122.133 | 550 | 25.8 | 366 | 39.8 | 534 | 6.3 |
| 34 | 37.485 | -122.086 | 511 | 24.9 | 126 | 61.6 | 496 | 6.8 |
| 36 | 37.468 | -122.067 | 444 | 24.0 | 294 | 61.2 | 426 | 6.5 |

Table 2: Summary and details for each of the GAM structures.

|  |  |  |
| --- | --- | --- |
| GAM | Summary | Details |
| Constant season | Nonlinear trends seasonality (constrained knots) | Treats year and doy with separate smoothers with a smooth non-linear trend through time, such that the seasonal peaks are similar across years, but interannual change is evident |
| Interactive season | Nonlinear trends wiht seasonality, plus interaction | Treats cyear and doy same as Const seas; and Within-year seasonal fluctuations that vary across years but in a constrained pattern compared to Const seas\* (ti(decYear, doy, bs = c("tp", "cc"))) |
| Constant season\* | Nonlinear trends with seasonality (unconstrained knots) | Variant of Const seas with much higher values of k, allowing the splines to follow greater fluctuations both within and across years. |

Table 3: Model performance statistics for each station and GAM structure (Table 2). Results are for models of chlorophyll-a using BoxCox and logarithmic transformations. Three GAMs were fit to each station and performance is summarized as AIC (from Const seas), GCV, and R-squared values. Probability values (“P-val”) show the results of Analysis of Variance comparisons between all BoxCox or logarithmic models at a station. The p-value for a row refers to the comparison of the model in the current row to the one preceeding (Const seas is the base model and does not have a p-value). A significant p-value indicates the model in the current row provides a statistically significant difference in explaining chlorophyll variance from the model in the preceding row. ns: not significant at , \* , and \*\* .

|  |  | Delta AIC | | Delta GCV | | R-squared | | P-val | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Station | GAM | BoxCox | Log | BoxCox | Log | BoxCox | Log | BoxCox | Log |
| 18 | Const seas | 0.0 | 0.0 | 0.000 | 0.000 | 0.49 | 0.51 | - | - |
|  | Int seas | -0.6 | -3.0 | -0.001 | -0.000 | 0.49 | 0.53 | ns | \* |
|  | Const seas\* | -0.9 | -0.5 | -0.001 | -0.000 | 0.50 | 0.51 | ns | \* |
| 21 | Const seas | 0.0 | 0.0 | 0.000 | 0.000 | 0.48 | 0.47 | - | - |
|  | Int seas | -60.0 | -57.9 | -0.022 | -0.005 | 0.55 | 0.54 | \*\* | \*\* |
|  | Const seas\* | -150.2 | -146.0 | -0.033 | -0.008 | 0.69 | 0.68 | \*\* | \*\* |
| 22 | Const seas | 0.0 | 0.0 | 0.000 | 0.000 | 0.44 | 0.43 | - | - |
|  | Int seas | -47.9 | -46.1 | -0.015 | -0.004 | 0.50 | 0.48 | \*\* | \*\* |
|  | Const seas\* | -47.1 | -46.0 | -0.010 | -0.003 | 0.55 | 0.54 | \* | \* |
| 24 | Const seas | 0.0 | 0.0 | 0.000 | 0.000 | 0.45 | 0.44 | - | - |
|  | Int seas | -86.6 | -79.7 | -0.027 | -0.008 | 0.54 | 0.52 | \*\* | \*\* |
|  | Const seas\* | -166.6 | -158.2 | -0.033 | -0.010 | 0.67 | 0.66 | \*\* | \*\* |
| 27 | Const seas | 0.0 | 0.0 | 0.000 | 0.000 | 0.42 | 0.40 | - | - |
|  | Int seas | -114.4 | -98.1 | -0.035 | -0.012 | 0.54 | 0.52 | \*\* | \*\* |
|  | Const seas\* | -243.0 | -209.2 | -0.049 | -0.016 | 0.71 | 0.68 | \*\* | \*\* |
| 30 | Const seas | 0.0 | 0.0 | 0.000 | 0.000 | 0.47 | 0.45 | - | - |
|  | Int seas | -81.8 | -70.3 | -0.020 | -0.009 | 0.56 | 0.53 | \*\* | \*\* |
|  | Const seas\* | -253.0 | -243.7 | -0.037 | -0.017 | 0.74 | 0.72 | \*\* | \*\* |
| 32 | Const seas | 0.0 | 0.0 | 0.000 | 0.000 | 0.43 | 0.41 | - | - |
|  | Int seas | -48.6 | -43.6 | -0.015 | -0.006 | 0.50 | 0.47 | \*\* | \*\* |
|  | Const seas\* | -343.1 | -311.5 | -0.053 | -0.022 | 0.77 | 0.75 | \*\* | \*\* |
| 34 | Const seas | 0.0 | 0.0 | 0.000 | 0.000 | 0.37 | 0.39 | - | - |
|  | Int seas | -23.8 | -16.0 | -0.008 | -0.003 | 0.41 | 0.41 | \*\* | - |
|  | Const seas\* | -186.4 | -209.2 | -0.034 | -0.017 | 0.66 | 0.69 | \*\* | \*\* |
| 36 | Const seas | 0.0 | 0.0 | 0.000 | 0.000 | 0.40 | 0.41 | - | - |
|  | Int seas | -41.1 | -22.2 | -0.013 | -0.005 | 0.45 | 0.44 | - | - |
|  | Const seas\* | -248.1 | -253.6 | -0.042 | -0.026 | 0.73 | 0.75 | \*\* | \*\* |

Table 4: Seasonal trends from Const seas\* results by station and decade. Results show the annual slope estimate as chlorophyll change per year for each seasonal period across years (JFM: January, February, March; AMJ: April, May, June; JAS: July, August, September, OND: October, November, December). Significant trends based on mixed-meta regression analysis are indicated with asterisks at , \* , and \*\* .

|  |  | Chlorophyll change (ug/l) per year | | | |
| --- | --- | --- | --- | --- | --- |
| Station | Years | JFM | AMJ | JAS | OND |
| 18 | 1990-2000 | 0.09 | 0.09 | 0.08 | 0.06 |
|  | 2000-2010 | 0.15 | 0.14 | 0.12 | 0.09 |
|  | 2010-2016 | -0.08 | -0.07 | -0.06 | -0.05 |
| 21 | 1990-2000 | 0.08 | 0.06 | 0.12 | 0.03 |
|  | 2000-2010 | **0.27\*** | 0.24 | 0.06 | -0.19 |
|  | 2010-2016 | -0.28 | -0.27 | -0.2 | -0.14 |
| 22 | 1990-2000 | 0.14 | 0.11 | 0.1 | 0.07 |
|  | 2000-2010 | 0.23 | 0.21 | 0.12 | -0.03 |
|  | 2010-2016 | -0.22 | -0.17 | -0.06 | 0.01 |
| 24 | 1990-2000 | 0.11 | 0.11 | 0.21 | 0.11 |
|  | 2000-2010 | **0.3\*** | 0.22 | 0.03 | -0.2 |
|  | 2010-2016 | -0.32 | -0.29 | -0.12 | -0.04 |
| 27 | 1990-2000 | 0.15 | 0.16 | 0.26 | 0.15 |
|  | 2000-2010 | **0.34\*** | 0.2 | -0.49 | **-0.59\*** |
|  | 2010-2016 | -0.15 | -0.16 | -0.15 | -0.11 |
| 30 | 1990-2000 | 0.21 | 0.19 | 0.22 | 0.15 |
|  | 2000-2010 | **0.53\*** | 0.33 | 0.02 | -0.22 |
|  | 2010-2016 | -0.29 | -0.26 | -0.19 | -0.15 |
| 32 | 1990-2000 | -0.84 | -0.32 | 0.19 | 0.2 |
|  | 2000-2010 | **0.52\*** | 0.43 | 0.25 | -0.21 |
|  | 2010-2016 | **-0.41\*** | -0.37 | -0.29 | -0.23 |
| 34 | 1990-2000 | **0.48\*** | 0.32 | 0.26 | 0.21 |
|  | 2000-2010 | **0.72\*\*** | **0.68\*** | **0.56\*\*** | -0.09 |
|  | 2010-2016 | **-0.61\*** | **-0.54\*** | -0.25 | -0.22 |
| 36 | 1990-2000 | **0.51\*** | **0.41\*** | **0.37\*** | 0.25 |
|  | 2000-2010 | 0.58 | 0.73 | 0.57 | -0.1 |
|  | 2010-2016 | -0.16 | -0.11 | 0.01 | 0.07 |

# Supplement

# References

Box, G. E. P., and D. R. Cox. 1964. An analysis of transformations (with discussion). Journal of the Royal Statistical Society (B) 26:211–252.

Cloern, J. E., and T. S. Schraga. 2016. USGS measurements of water quality in San Francisco Bay (CA), 1969-2015: U.S. Geological Survey data release. https://doi.org/10.5066/F7TQ5ZPR.

Cloern, J. E., T. S. Shcraga, E. Nejad, and C. Martin. 2020. Nutrient status of San Francisco Bay and its management implications. Estuaries & Coasts 43:1299–1317.

Hastie, T. J., and R. J. Tibshirani. 1990. Generalized additive models. Page 335. Chapman; Hall, London, New York.

Hyndman, R., G. Athanasopoulos, C. Bergmeir, G. Caceres, L. Chhay, M. O’Hara-Wild, F. Petropoulos, S. Razbash, E. Wang, and F. Yasmeen. 2020. forecast: Forecasting functions for time series and linear models.

Murphy, R. R., E. Perry, J. Harcum, and J. Keisman. 2019. A Generalized Additive Model Approach to evaluating water quality: Chesapeake Bay case study. Environmenal Modelling & Software 118:1–13.

Schraga, T. S., E. S. Nejad, C. A. Martin, and J. E. Cloern. 2020. USGS measurements of water quality in San Francisco (CA), beginning in 2016 (ver. 3.0, March 2020): U.S. Geological Survey data release. https://doi.org/10.5066/F7D21WGF.

Sera, F., B. Armstrong, M. Blangiardo, and A. Gasparrini. 2019. An extended mixed-effects framework for meta-analysis. Statistics in Medicine 38:5429–5444.

Venables, W. N., and B. D. Ripley. 2002. Modern applied statistics with S. Fourth. Springer-Verlag, New York, New York.

Wood, S. N. 2003. Thin-plate regression splines. Journal of the Royal Statistical Society (B) 65:95–114.

Wood, S. N. 2017. Generalized additive models: An introduction with r. Page 476. 2nd editions. Chapman; Hall, CRC Press, London, United Kingdom.