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 2020-04-13 12:45:44

# Abstract

# Introduction

# 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. The estuary is subdivided into six sub-embayments having distinct geomorphological characteristics: Suisan Bay, San Pablo Bay, North Central Bay, Central Bay, South Bay,and Lower South Bay. Major riverine 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-dominanted (salinity ranging from 0 to 15 ppt) with seasonal freshwater pulses primarily occurring in the winter and spring during the rainy season and subsequent snowmelt in the upper watershed. The southern subembayments are marine-dominated with salinity ranging from 5 to 35 ppt depending on diurnal variation in the tidal cycle and effluent discharge from wastewater treatment plants. South Bay 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 conctributes also contributes to nutrient loading in the SFE with annual nutrient export estimated as approximately 30 thousand kg dy of Nitrogen from the Delta.

Nitrogen and phospohorus 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 has occurred since the 1970s. The resistance of SFE to eutrophication has been attributed to several factors, including elevated susepended sediments that reduce light penetration in the water column, regular exchange 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 recent harmful algal blooms (HABs), increases in summer-fall chlorophyll concentrations, and low dissolved oxygen concentrations were observed beginning in 1999.

Recent renewed interest in the effects of elevated nutrient concentrations in South Bay have been motivated by

* SF south Bay dataset, USGS monitoring data, show observed data summer/fall IQR plots (is the trend significant? when did it start or stop being significant?)

Analysis focused on near-surface (depth = 1-3m) chl-a data from Central Bay, South Bay, and Lower South Bay (stations 18-36, Figure 1), collected biweekly to monthly through long-term monitoring by USGS (Cloern and Schraga, 2016; Schraga et al.2017). For this initial work, we selected the time period 1990-2017 because it represented a suitable balance among three factors relevant to testing the statistical approaches: a. sufficient length of record; b. consistent biweekly-monthly sampling; and c. a diverse set of stations, across multiple subembayments, for which a and b were satisfied. While sampling frequency varied somewhat overtime or by station, all data were treated as individual date/station/concentration records within the statistical models (no spatial or temporal binning or averaging).

## GAM model format

* GAM general format and structures (gam0, gam1, gam2, gam6)
* Model application to SF time series and model evaluation, including transformation for response variables (why Box-Cox?)
* Model predictions (complete time series, annual trend) and back-transformation estimates

## Long-term and seasonal trend analysis

* Secondary methods for trend analysis, estimate of seasonal values, hypothesis tests including mixed meta-analysis

## Sensitivity analysis

* Sensitivity analysis - do trend estimates or hypothesis test results change by model type?

# Results

# Discussion

# Conclusions

# Acknowledgments

# Figures

# Tables

# References