Scientific Basis to Assess the Effects of Nutrients on San Francisco Bay Beneficial Uses

Prepared for:
San Francisco Bay Regional Water Quality Control Board
Contract 11-151-120
Contract Manager: Naomi Feger

Martha Sutula
Southern California Coastal Water Research Project, Costa Mesa CA

David Senn
San Francisco Estuary Institute, Richmond, CA

October 2015
Technical Report 864

EXPERT WORKGROUP

Gry Mine Berg
Applied Marine Sciences, Santa Cruz, CA

Suzanne Bricker
NOAA National Centers for Coastal Ocean Science, Silver Spring, MD

James Cloern
U.S. Geological Survey, Menlo Park, CA

Richard Dugdale
Romberg Tiberon Center, San Francisco State University, Tiberon, CA

James Hagy
U.S. Environmental Protection Agency
Office of Research & Development, Gulf Ecology Division, Gulf Breeze, FL

Lawrence Harding
Department of Earth and Atmospheric Sciences, University of California Los Angeles, CA

Raphael Kudela
Ocean Sciences Department, University of California, Santa Cruz, CA

Executive Summary

San Francisco Bay (SFB) has long been recognized as a nutrient-enriched estuary; however, until recently, it has exhibited resistance to symptoms of nutrient overenrichment due to a number of factors such as high turbidity, strong tidal mixing, and grazing by bivalves. Recent observations have reinforced the need to identify numeric water quality objectives and management actions to protect SFB from the potential effects of nutrient over-enrichment. The purpose of this work was to develop a quantitative framework, hereto referred to as an *assessment framework,* to assess eutrophication in the SFB, based on indicators of dissolved oxygen (DO), phytoplankton biomass (chlorophyll-a), gross primary productivity, the prevalence of harmful algal blooms (HAB) and toxins.

A group of experts in the ecology of SFB, as well as international experts in assessment frameworks (AF) and nutrient criteria, worked in concert to define core principles for the AF. These principles include the geographic scope, recommended Bay segmentation of subembayments for assessment, and the protocols and recommended spatial and temporal frequency of monitoring that would support use of the framework to assess nutrient effects on SFB. A quantitative scheme was developed to classify SFB subembayments in tiers of ecological condition, from very high to very low, based on risk of potential adverse effects of nutrient overenrichment and eutrophication. Decisions on classification bins were supported by a combination of existing literature and guidance, quantitative analyses of existing SFB data from the USGS research program, and expert best professional judgment. Analyses of two decades of phytoplankton species composition, chlorophyll-a, and dissolved oxygen (DO), and 3 years of toxin data from solid phase adsorption toxin tracking (SPATT) samplers were used to support decisions on the AF and demonstrated: 1) significant increases in chlorophyll-a, declines in DO, and a high prevalence of HAB species and toxins across most SFB subembayments and 2) strong linkage of increasing chlorophyll-a to declining DO and HAB abundance. Statistical approaches were used to define thresholds in chlorophyll-a relating to increased risks of HABs and declining DO. These thresholds were used, in combination with expert best professional judgment, to develop an AF classification scheme. A qualitative summary of uncertainty associated with each indicator was made for the purpose of focusing future research, monitoring, and modeling on AF refinement.

The AF is intended to provide a decision framework for quantifying the extent to which SFB is supporting beneficial uses with respect to nutrients. This AF is comprised of three important elements: 1) a set of conceptual models that defines what a problem would look like in SFB, if it occurred, 2) a set of core principles supporting the AF, and 3) classification tables. The AF supports and is supported through the other major science elements. The conceptual models and AF core principles provide a sound scientific foundation for informing modeling and monitoring. Through early interactions with the stakeholder community, these two components of the AF appear to have the greatest consensus and the least "uncertainty."

The classification scheme is a critical element of the AF, because it represents a quantitative and transparent mechanism through which SFB data can be interpreted to assess, nutrient-related beneficial use support. Given its importance, <u>the authors of this document fully acknowledge the uncertainty in the AF classification scheme and need for refinement, through multiple iterations of basic research, monitoring, and modeling</u>. We suggest that the near-term use of the AF classification system be focused on a scientific “test drive” focused on understanding how to collectively use and improve efficiencies for assessment, monitoring and modeling. The “test drive” of the AF can be conducted in tandem with research, monitoring, and modeling to improve the scientific foundation for the AF, aimed at the following six major recommended actions:

1. Improve the scientific basis for nutrient-related segmentation of SFB.

2. Reduce sources of uncertainty in chlorophyll-a, HAB abundance and toxin classification by: 1) Better assessment and characterization of the ecological and human risk of HABs in SFB, 2) Co-location of chlorophyll-a and monitoring of toxins in Bay surface waters, shellfish and SPATT to improve documentation of linkage of chlorophyll-a to HAB toxin concentrations, 3) Expand SPATT samplers to include other toxins and conduct better validation of SPATT toxin data relative to surface waters or mussel toxin tissues, 4) Assemble a scientific workgroup to evaluate and provide recommendations on the chronic effects of HAB toxins, and 5) Improve monitoring through better spatial and temporal coverage of HAB data to link chlorophyll-a to DO.

3. Optimize spatial and temporal sampling of AF indicators to best align quality of the information produced, while balancing costs, logistics, and power to detect trends.

4. Improve the scientific basis for dissolved oxygen classification and monitoring in future iterations of the AF. Current recommendations focus on indicators of phytoplankton. We recommend: 1) synthesis of DO expectations for SFB species types and the seasonal use of specific habitat types (deep channel, shallow subtidal, tidal sloughs, etc.) within SFB subembayments; 2) improved characterization of the diel variability of DO at key points within the deep water and shallow margin habitat of each subembayment in order to better characterize support of species and habitats; and 3) improved mechanistic understanding of the physical and biological factors influencing DO within and between the deep channel and shallow water margin habitat.

5. Include diked baylands, restored salt ponds and tidal sloughs in future iterations of the AF, which is currently focused on open water habitats.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

We wish to thank San Francisco Bay Regional Water Quality Control Board staff and the members of the San Francisco Bay Nutrient Technical Workgroup and Steering Committee for their thoughtful feedback and guidance throughout the study. This report was produced under San Francisco Bay Regional Water Quality Control Board contract to the Southern California Coastal Water Research Project (Agreement Number 11-151-120).

Emily Novick of SFEI, Marcus Beck (US EPA ORD), and Karen McLaughlin (SCCWRP) provided assistance with data management and analyses. We are extremely grateful for their assistance.

This report should be cited as:
Sutula, M. and D. Senn. Scientific basis for assessment of nutrient impacts on San Francisco Bay. Technical Report 864. Southern California Coastal Water Research Project. Costa Mesa, CA. www.sccwrp.org

1 INTRODUCTION

1.1 Background and Purpose

The San Francisco Bay Regional Water Quality Control Board (Water Board) is developing nutrient water quality objectives for San Francisco Bay. Water Board staff favor an ecological risk assessment approach (EPA 1998), in which ecological response indicators (e.g. change in algal abundance and assemblage, dissolved oxygen) are used as the endpoints to assess whether the San Francisco Bay (SFB) is supporting designated uses. A model would then be used to link those endpoints to nutrients and other factors that comprise management options to (e.g. best management practices). In this risk-based approach, nutrients are considered a resource that should be managed at levels that support SFB beneficial uses. The key is managing nutrients at levels that pose a low risk of adverse effects, while ensuring the system doesn't become nutrient-limited. This approach is consistent with that being used for nutrient objective development for other waterbodies in California, including other estuaries (SWRCB 2014).

The process of selecting appropriate endpoints begins with a synthesis of science and the development of a framework for interpreting the endpoints that is ultimately based on policy decisions by the Water Board, taking into consideration advice from its advisory groups. In this document, we refer to the product of scientific synthesis as a **nutrient assessment framework (AF)**, defined as a structured set of decision rules that specify how to use monitoring data to categorize specific subembayments of SFB from very high to very low ecological condition, using indicators that have a direct linkage to nutrients and support of SFB beneficial uses. Thus, while the decision on regulatory endpoints should be informed by science, it is ultimately a policy decision. The ultimate goal of this effort is that the Water Board would propose numeric endpoints for SFB, based on the synthesis of science represented in the AF and feedback from the SFB stakeholders and scientific peer review.

The purpose of this document is to describe the SFB nutrient AF, the scientific synthesis upon which it is based, and key data gaps and recommendations for its future refinement.

1.2 Document Audience, Authorship, and Organization

This report was written to address the information needs of both scientists and technically-oriented decision makers and stakeholders involved in the SFB Nutrient Management Strategy. With that audience in mind, the report assumes a certain baseline familiarity with SFB as well as a basic understanding of the biology, nutrient cycling, biogeochemistry, and physical processes in estuaries. The scientific synthesis supporting this report was developed collaboratively with a team of co-authors consisting of scientists whose areas of expertise cover a range of relevant disciplines and much of whose work has focused on SFB.

This document is organized as follows:

Section 1	Introduction. Purpose, and Organization
Section 2	Context: Detailed Background, Process for AF Development, and Review of Existing Approaches
Section 3	Proposed AF Core Principles and Classification Tables
Section 4	Summary and Recommendations

Appendices Key definitions, supporting literatures reviews and quantitative analyses

2 Context for Framework Development: Detailed Background, Process for Development, and Review of Existing Approaches

2.1 San Francisco Bay: A Brief History and Context for Nutrient Management

SFB encompasses several subembayments of the San Francisco Estuary, the largest estuary in California. SFB is surrounded by remnant tidal marshes, an array of intertidal and subtidal habitats, tributary rivers, the freshwater "Delta" portion of the estuary, and the large mixed-land-use area known as the San Francisco Bay Area. San Francisco Bay hosts an array of habitat types, many of which have undergone substantial changes in their size or quality due to human activities (Conomos (ed.) 1979). Urban residential and commercial land uses comprise a large portion of Bay Area watersheds, in particular those adjacent to Central Bay, South Bay and Lower South Bay. Open space and agricultural land uses comprise larger proportions of the areas draining to Suisun Bay and San Pablo Bay. The San Joaquin and Sacramento Rivers drain 40% of California, including agricultural-intensive land use areas in the Central Valley. Flows from several urban centers also enter these rivers, most notably Sacramento which is ~100 km upstream of Suisun Bay along the Sacramento River.

SFB receives high nutrient loads from 37 public owned wastewater treatment works (POTWs) servicing the Bay Area's 7.2 million people (Association of Bay Area Governments, www.abag.ca.gov). Several POTWs carry out nutrient removal before effluent discharge; however, the majority are designed to have secondary treatment without additional N or P removal. Nutrients also enter SFB via stormwater runoff from the densely populated watersheds that surround SFB. Flows from the Sacramento and San Joaquin Rivers deliver large nutrient loads, and enter the northern estuary through the Sacramento/San Joaquin Delta.

SFB nutrient loads and ambient nutrient concentrations are among the highest of the U.S. estuaries (2012), However, SFB has long been considered relatively immune to its high nutrient loads. For example, the first San Francisco Bay Regional Basin Plan from 1975 stated that only limited treatment for nutrients was necessary because the system was considered to be light-limited (SFRWQCB, 1975). Research and monitoring over the last 40 years have identified several factors that impart SFB with resilience to high nutrient loads, i.e., control on phytoplankton production (e.g., see Cloern and Jassby 2012; Cloern et al., 2007), including high turbidity, strong tidal mixing, and abundant filter-feeding clam populations.

However, recent studies indicate that the response to nutrients in SFB is changing. These shifts in nutrient responses may be triggered by one or more recently documented changes in SFB, including shifts in the timing and extent of freshwater inflow and salinity intrusion, decreasing turbidity, restructuring of plankton communities, and reduced metal contamination of biota, and food web changes that decrease resistance of the estuary to nutrient pollution (Cloern and Jassby 2012).

Since 1969, a USGS research program has supported water-quality sampling in the San Francisco Bay. This program collects monthly samples between the South Bay and the lower

Sacramento River to measure salinity, temperature, turbidity, suspended sediments, nutrients, dissolved oxygen and chlorophyll-a. The USGS data, along with sampling conducted by the Interagency Ecological Program (IEP), provide coverage for the entire San Francisco Bay-Delta system. Although these data are critical to our current understanding of the Bay-Delta Estuary, the USGS program is a research program and, thus, is not intended to serve as a comprehensive SFB nutrient monitoring program.

The Nutrient Strategy highlights the need to lay the groundwork for a regionally supported, long-term monitoring program that should be organized in such a way as to collaborate with ongoing research efforts to provide the information that is most needed to support management decisions in the Bay.

The technical approach underpinning the SFB Nutrient Management Strategy is compatible with a major statewide initiative, led by the California State Water Resources Control Board (State Water Board), to develop nutrient water quality objectives for the rest of the State's estuaries www.swrcb.ca.gov/water_issues/programs/nutrient_objectives/ .

2.2 SFB Nutrient Management Strategy: Management Questions, Major Work Elements, and Linkage to AF

To address growing concerns that SFB's response to nutrients is changing and that conditions may be trending toward adverse impacts due to elevated nutrient loads, the Water Board worked collaboratively with stakeholders to develop the San Francisco Bay Nutrient Management Strategy (herein referred to as "the Strategy"; SFRWQCB 2012), which lays out an approach for gathering and applying information to inform management decisions. The Strategy identified four overarching management questions:

- Is SFB currently experiencing nutrient-related impairment, or are there signs of future impairment?
- What are appropriate guidelines for identifying a problem?
- What nutrient loads can the Bay assimilate without impairment of beneficial uses?
- What are the contributions of different loading pathways, and how do they vary in importance as a function of space and time?

To address these management questions, the Strategy identified five major work elements:

- Conceptual model development, scientific synthesis and basic research
- Nutrient assessment framework
- Modeling
- Monitoring and special studies
- Characterization of nutrient loads, sources and major pathways

This report consists of the proposed AF and the analyses and literature that supported its development. Other major elements exist and are in various stages of progress (http://sfbaynutrients.sfei.org/).

The nutrient AF is intended to provide a decision framework for quantifying the extent to which SFB is supporting beneficial uses with respect to nutrients. It also is integral to the other major elements by:

- Defining monitoring requirements (the core indicators, spatial and temporal frequency of sampling) needed to support routine assessments of SFB
- Identifying a set of management endpoints that should constitute the output of SFB water quality models that will improve the mechanistic understanding of the linkage of nutrients to adverse outcomes in SFB
- Contributing to key science needs and analyses needed to further refine the AF

This last bullet point is a critical product of this effort, as the authors of this document fully acknowledge the considerable uncertainty in the AF classification scheme and need for refinement, through multiple iterations of basic research, monitoring, and modeling.

2.3 Conceptual Approach, Desired Attributes of a Nutrient AF and Process for Development

Conceptual Approach to AF Development

Nutrient objectives are scientifically challenging because nutrients are required to support life and the assessment of how much is "too much" is not straightforward. Typical paradigms used to set thresholds for toxic contaminants do not apply, in part because the adverse effects of nutrient over-enrichment are visible at orders of magnitude below recognized toxicity thresholds for unionized ammonia and nitrate. In addition, the effects of nutrient discharges often occur via indirect exposure pathways, which are spatially and temporally disconnected from their points of discharge.

The conceptual approach for AF development is anchored in an ecological risk assessment approach (EPA 1998), which consists of multiple ecological response indicators (e.g., algal abundance and assemblage, dissolved oxygen) as endpoints to assess whether SFB is supporting beneficial uses (Tetra Tech 2006). A hydrodynamic and water quality model is then used to link those assessment endpoints to nutrients and other factors that comprise management options (e.g., best management practices). In this risk-based approach, nutrients are considered a resource that should be managed at levels to maintain SFB designated uses, while maintaining a low risk of adverse effects. If the nutrients present – regardless of actual magnitude – have a low probability of impairing uses, then water quality standards can be considered met. This approach is consistent with EPA guidance for nutrient criteria development (e.g., cause-effect approach; EPA 2001) and with guidance being used by the State Water Board for nutrient objective development for other waterbodies in California (SWRCB 2014), including other estuaries (Sutula 2011).

This ecological risk-based approach has two important advantages. First, it offers a more direct linkage with beneficial uses and is generally thought to lend itself to a more precise diagnosis of adverse effects. Second, the alternative approaches, such as stress-response or reference-based approaches, are particularly problematic in estuaries. SFB and other estuaries within California are highly variable in how they respond to nutrient loading, due to differences in physiographic setting, salinity regime, frequency and timing of freshwater flows, magnitude of tidal forcing, sediment load, stratification, residence time, denitrification, etc. This combination of "co-factors" results in differences in the dominant primary producer communities (i.e., phytoplankton, macroalgae, benthic algae, submerged aquatic vegetation, emergent macrophytes). It also creates variability in the pathways that control how nutrients cycle within the estuary. At times, these co-factors can play a larger role in mitigating estuarine response to nutrient loads or concentrations, blurring or completely obscuring a simple prediction of primary productivity limited by nutrients.

Thus, the Water Board is working to develop an AF based on the following key tenets:

1. *Ecological response indicators (e.g., dissolved oxygen, primary producer abundance, productivity and assemblages) should provide a more direct risk-based linkage to beneficial uses than to nutrient concentrations or loads.* The AF should be based on assessing eutrophication (or other adverse effects), rather than nutrient over-enrichment per se.
2. *A weight-of-evidence approach with multiple indicators can produce a more robust assessment of eutrophication.* Wherever possible, the use of multiple indicators in a "weight-of-evidence" approach provides a more robust means to assess ecological condition and determine impairment. This approach is similar to the multimetric index approach, which defines an array of metrics or measures that provide limited information on biological status on an individual basis, but when integrated, serve to inform overall biological condition.
3. *Models can be used convert response indicators to site-specific nutrient loads or concentrations.* A key premise of the NNE framework is the use of models to convert numeric endpoints, based on ecological response indicators, to site-specific nutrient goals appropriate for permitting and TMDLs. A key feature of these models is that they account for site-specific co-factors, such as light availability, temperature, and hydrology that modify the ecological response of a system to nutrients. Thus, nutrient forms and ratios are not an explicit element of the AF, but become linked to assessment endpoints through modeling of ecological processes.

Desirable Attributes of an AF

The goal of the nutrient AF is to provide a structured set of decision rules that specify how to use monitoring data to categorize specific subembayments of SFB, from very high to very low ecological condition, using indicators that have a direct linkage to nutrients and support of SFB beneficial uses.

To achieve this goal, a nutrient AF for SFB should offer the following features:

- The AF should employ indicator(s) that have a strong linkage to Bay beneficial uses. This linkage should be scientifically well-supported and easily communicable to the public.
- One or more primary indicators of the AF should have a predictive relationship with surface water nutrients and/or nutrient loads to the Bay.
- The AF should employ the indicator(s) that classify the Bay subembayments from very high ecological condition to very low ecological condition. It should be explicit as to how

the magnitude, extent, and duration of the effects cause the subembayments to be classified differently.

- The AF should be spatially explicit for different subembayments of the Bay and different habitat types (deep vs. shallow subtidal), as warranted by the ecological nature of response to nutrients.
- The AF should specify what appropriate methods are used to measure the indicator and the temporal frequency and spatial density of data required to make that assessment.
- It should provide guidance on how the data should be analyzed to categorize the Bay subembayments.

Methodology Used to Develop AF

The methodology used to develop the AF consisted of five main steps:

1. **Empanel a team of scientific experts to guide AF development.** These experts represented a diverse body of knowledge of SFB hydrology, estuarine ecology and nutrient biochemistry, as well as expertise in nutrient criteria and AF development. This team is listed as contributing authors on this document.
2. **Review existing approaches to nutrient AF development.** A white paper was completed identifying candidate indicators and metrics, summarizing existing literature for how those indicators have been used to assess ecological condition, and recommending a suite of options to consider for further exploration (Appendix 1).
3. **Identify AF core principles,** including geographic scope and key habitats, key indicators and recommended measures, and the spatial and temporal frequency of sampling required for assessment.
4. **Analyze existing data to develop supporting information to develop a classification scheme.** Existing data were utilized to test out existing classification schemes and to quantify relationships between key variables of interest. These analyses are summarized in Section 3, and additional methods and supporting information are provided in Appendix 2.
5. **Develop AF classification scheme and quantify/describe major uncertainties**. Existing literature and supporting analyses were used to develop the AF classification scheme. For each indicator, uncertainties corresponding to classification "bins" were summarized. Key science needs required for the refinement of the classification scheme and core principles were summarized.

Testing the AF with existing or newly collected monitoring data, and further refinement based on monitoring and modeling, are steps envisioned for the AF in subsequent phase(s) outside the scope of this document.

2.4 Review of Existing Frameworks to Assess the Effects of Nutrient Over-Enrichment on Estuaries

We reviewed the existing regulatory and non-regulatory approaches to the assessment of the effects of nutrient over-enrichment in estuarine waterbodies worldwide in order to consider an appropriate approach to AF development (see white paper in Appendix B). A wide variety of methodologies exist (Table 2.1). All of the conceptual models reviewed focused on ecological impacts (i.e., eutrophication), rather than on nutrients' direct effects on ecological condition (i.e., toxicity).

The white paper (Appendix B) arrived at the following conclusions:

- **The eutrophication AFs reviewed have a common set of conceptual models.** These conceptual models show linkages to nutrients and relevant co-factors, as well as the risk pathways of "impairment" of ecosystems services and beneficial uses. These pathways of impairment include **(**1) increased harmful algal blooms, which can produce toxins that adversely affect both human health and aquatic life, (2) hypoxia and anoxia triggered by frequent algal blooms, which change the long-term balance of organic matter cycling and accumulation within an estuary (Nixon 1995) and can adversely affect habitat and aquatic life, (3) shifts in the dominance assemblages and size class of phytoplankton, which lead to degradation of food quality for estuarine consumers, including commercial and recreational fisheries, and (4) overabundance of algae, which results in reduced light availability for benthic primary producers (e.g., seagrass).

- **A common set of response indicators are used, focusing on dissolved oxygen and primary producers (e.g., Bricker et al. 2003, Zaldivar et al. 2008), that link to these major conceptual models.** Among primary producer indicators used, phytoplankton biomass (water column chlorophyll-a) is the most common (Table 2.1). The frequency and magnitude of harmful algal blooms and toxin concentrations have also been used, either directly as an indicator or indirectly using chlorophyll-a as a proxy for the increased probability of occurrence of HAB events. Phytoplankton assemblage has been used in assessment of ecological condition, but only in estuaries that can use a reference approach to defining the envelope of reference assemblages. Where TN and TP are used (typically in regulatory programs), they have been determined as a proxy for primary productivity either through statistical or process modeling to primary producer numeric targets (e.g., regulatory programs such as Chesapeake Bay and Florida), or through a reference water body approach (Andersen et al. 2011).

- **Among non-regulatory AFs (Bricker et al. 2003, Zaldivar et al. 2008), estuarine subembayments are binned into multiple condition classes, representing a disturbance gradient of high to low ecological condition (e.g., Zaldivar et al. 2008) or trophic state (Bricker et al. 2003).** These condition classes are developed through a combination of scientific data analyses and expert best professional judgment.

- **There is some degree of convergence on the thresholds or ranges represented within the various classification scheme, particularly for chlorophyll-a (see white paper, Appendix B).** This suggests consensus among experts who developed these frameworks that the ranges representing condition classes correspond to real ecosystem decline. That said, two points are worth mentioning. First, there is great variability in the temporal statistic (e.g., annual average, season max, 90th percentile) used to make the assessment. Second, the differences in the ranges, while small, represent large differences in estuarine productivity, especially on annual timescales.

- **Inherent in these AFs are key differences in temporal statistic, spatial density of data used to make an assessment and, in some cases, the way that multiple indicators are combined into a single score (Table 2.2).** These details are less obvious, but can have large effects on scoring (McLaughlin et al. 2013).

Table 2.1 Methods of eutrophication assessment and examples of biological and physico-chemical indicators used and integration capabilities (pressure-state and overall; modified from Borja et al. 2009). From Ferreira et al. 2011.

Method Name	Biological indicators	Physico-chemical indicators	Nutrient load related to impairments	Integrated final rating
TRIX[b]	Chl	DO, DIN, TP	no	yes
EPA NCA Water Quality Index[a]	Chl	Water clarity, DO, DIN, DIP	no	yes
ASSETS[e]	Chl, macroalgae, seagrass, HAB	DO	yes	yes
TWQI/LWQI[c]	Chl, macroalgae, seagrass	DO, DIN, DIP	no	yes
OSPAR COMPP[g]	Chl, macroalgae, seagrass, phytoplankton indicator species	DO, TP, TN, DIN, DIP	yes	yes
WFD[f]	phytoplankton, Chl, macroalgae, benthic invertebrates, seagrass,	DO, TP, TN, DIN, DIP, water clarity	no	yes
HEAT[d]	Chl, primary production, seagrass, benthic invertebrates, HAB, macroalgae	DIN, DIP, TN, TP, DO, water clarity	no	yes
IFREMER[h]	Chl, seagrass, macrobenthos, HAB	DO water clarity, SRP, TP, TN, DIN, sediment organic matter, sediment TN, TP	no	yes
STI[i]	Chl, Primary Production	DIN, DIP	no	no

[a] USEPA, 2005, 2008.
[b] Vollenweider *et al.*, 1998.
[c] Giordani *et al.*, 2009.
[d] HELCOM, 2009.
[e] Bricker *et al.*, 1999, 2003, 2007.
[f] Devlin, pers.Com.
[g] OSPAR, 2002, 2008.
[h] Souchu *et al.*, 2000.
[i] Ignatiades, 2005.

Table 2.2. Summary of approaches used for assessment of eutrophication applicable to shallow and deepwater unvegetated subtidal habitat. Adapted from Devlin et al. 2011.

		UK WFD	OSPAR	TRIX	ASSETS	EPA NCA	TWQI/LWQF	HEAT	IFREMER
Grouping of Variables	Caus-ative Factors	Nutrient Load	DIN and DIP concentration, ratios, and loads	DIN and TP concentration	DIN and DIP loads	DIN, DIP conc	TN, TP, DIN and DIP conc.	DIN and DIP	PO4, NOX, NH4, TN, TP
	1ary effects	Chl-a, PP indicator species, seasonal changes in cell abundance of diatoms/dinoflagellates, SAV, macroalgae	Chl-a, PP indicator species, macroalgae, microphytobenthos, SAV	Chl-A	Chl-a macroalgae	water clarity, chl-a	Chl a, SAV, macroalgae	Chl a, water clarity, SAV,	Chl a, turbidity
	2ary effects	DO	DO, zoobenthos and/or fish kills, organic carbon	DO	Nuisance/toxic blooms	DO	DO	Benthic invertebrates	DO percent saturation
	Other		Algal toxins						
Temporal sampling framework		Annual chla and DO, winter DIN, monthly PP groups	Growing season chl-a (Mar-Sept), Winter DIN, summer DO	Annual	Annual	One sample per year (per station) within summer index period	Results can be derived based on one time or multiple periods	Growing season chl-a (Mar-Sept), Winter DIN, summer DO	Annual
Spatial sampling framework		Sampling in estuaries and nearshore defined by salinity, reported by waterbody	Sampling defined by salinity in estuaries, nearshore	Sampling mostly in larger offshore systems; results reported by region	Sampling in salinity zones, synthesized to waterbody, region, national, with reporting at all levels	Sampling is regional, synthesized to national level, reported at regional and national level	For shallow, benthic PP dominated. Can be applied to single stations or groups of stations.	Sampling defined by salinity in Baltic Sea	For shallow, benthic PP dominated. Can be applied to single stations or groups of stations.
Assessment of indicators		Deviation from reference conditions	Deviation from reference conditions	Placement on scale from 1-10 TRIX units	Deviation from reference conditions	Deviation from reference conditions	Deviation from reference condition	Deviation from reference condition	Deviation from reference
Comb-ination Method		Indicator scores are averaged within an indicator group. Final score gives classification status	One out, all out for individual categories and overall classification	Linear combo of logarithm of variables modified by scaling coefficient	Scores of avg. primary and secondary indicators combined in a matrix	Indicators assessed individually. WQI based on % of samples in 4 categories.	TWQI scores combined as the sum of weighted quality values for individual indicators.	One out, all out for individual categories and overall classification	One out all out

3 Framework to Assess The Effects of Nutrients on San Francisco Bay Beneficial Uses

3.1 AF Core Principles

Geographic Scope and Focal Habitats

The geographic scope for the SFB AF is defined by the Golden Gate Bridge as the oceanward boundary, and Broad Slough in the Sacramento River as the upstream boundary, which is just upstream of Winter Island (the boundary between the San Francisco and Central Valley Water Boards; Figure 3.1).

SFB is comprised of deep and shallow water subtidal habitats and intertidal wetlands, and remnant tidal marshes (Figure 3.1). Deepwater and shallow subtidal habitats are the focus of this AF.

Although diked baylands, restored salt ponds, and tidal sloughs also are present in SFB and are important, they are excluded in this initial assessment work. That said, preliminary data indicate that these habitats may be in questionable ecological condition (Topping et al. 2009, SFEI 2014a); thus, we recommend development of an AF targeting these habitats in a subsequent phase of framework development.

Segmentation

SFB has six subembayments with very different physical, biogeochemical, and biological characteristics that shape their individual responses to nutrients. For this reason, the AF should be spatially explicit for these regions (herein referred to as subembayments) of SFB, as warranted by the ecological nature of response to nutrients.

The physical features in SFB provide natural breakpoints for segmentation, as documented by Jassby et al. (1997) for chlorophyll-a, TSS and salinity. These breakpoints or subembayment boundaries are also obvious in other ecological data. The SFB Regional Monitoring Program (RMP) uses a segmentation scheme that differs slightly from that of Jassby et al. (1997); this segmentation scheme was derived based on a variety of different contaminant and environmental gradients not necessarily relevant for nutrients.

For the AF and supporting analyses, we used subembayment classification based on Jassby et al. (1997; Table 3.1., Figure 3.1). That said, we strongly recommend reanalysis of existing data in the Jassby et al. (1997) methodology, using newly available and relevant ecological data, to finalize this segmentation scheme.

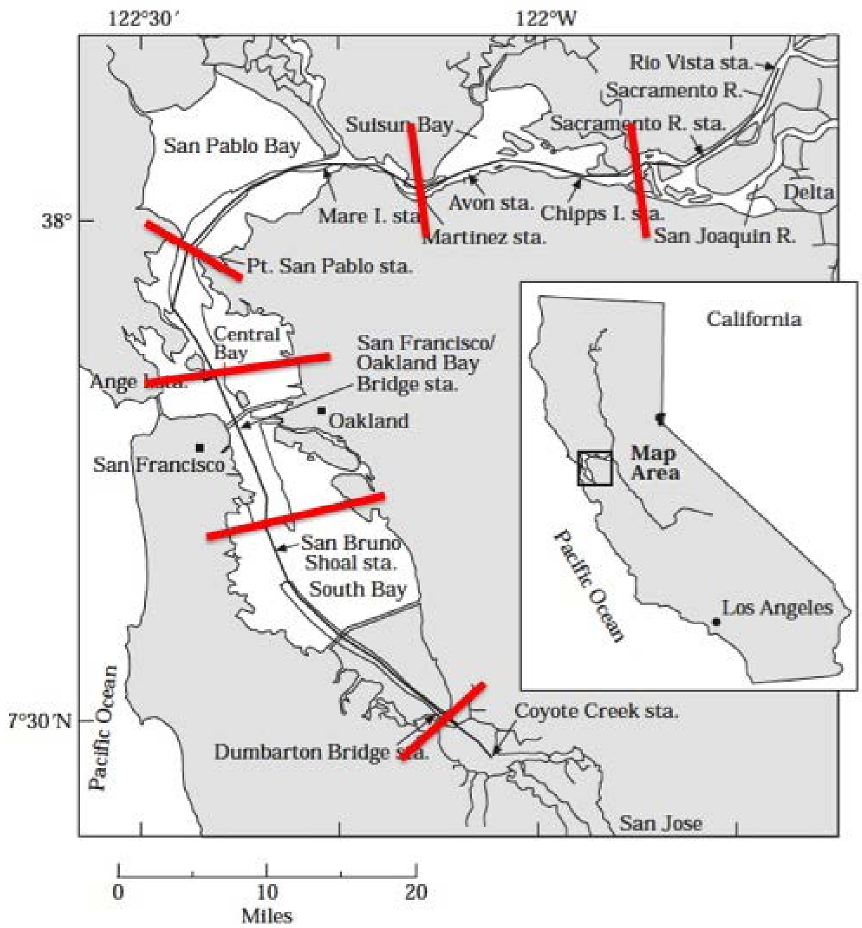


Figure 3.1 Map of SFB showing geographic scope of AF, focal habitats and subembayment boundaries. Subembayment names are designated on the map.

Table 3.1. Size and locations of boundaries defined by preliminary AF classification scheme (from Jassby et al. 1997).

Stratum no.	Description	Size ± SD (km)	Northing (km)	Easting (km)
1	South of Dumbarton Br.	6·9 ± 0·3	<151·4	
2	Dumbarton Br. to San Bruno Shoal	23·3 ± 0·6	151·4–165·3	
3	San Bruno Shoal to Angel I.	28·7 ± 0·7	165·3–188·8	
4	Angel I. to Mare I.	37·3 ± 2·2	≥188·8	<564·6
5	Mare I. to Martinez	13·1 ± 1·1	≥188·8	564·6–574·5
6	East of Martinez	51·8 ± 1·7	≥188·8	≥574·5

Key Indicators and Linkage to SFB Beneficial Uses

A core principle of the AF is the use of several indicators as multiple lines of evidence for potential adverse impacts (Figure 3.2), assuring a more robust assessment of the ecological condition of SFB subembayments. In the SFEI 2014b report, experts arrived at consensus regarding what undesirable conditions would plausibly manifest in SFB in response to adverse nutrient-related impacts – and how each undesirable state would impact beneficial uses (Table 3.2). The undesirable states were divided into seven categories that represent specific examples extending from more general adverse impact pathways (Figure 3.2).

The undesirable states can be measured by six key indicators representing the multiple lines of evidence within this AF:

1. Phytoplankton biomass (as chlorophyll-a)
2. Gross and net phytoplankton production (hereto referred to collectively as GPP)
3. Harmful algal bloom species abundance
4. HAB toxin concentrations
5. Phytoplankton assemblage, expressed as phytoplankton food quality, percent of biovolume < 0.5 microns, and other metrics of community change
6. Dissolved oxygen

The remainder of this section is devoted to analyzing the seven undesirable states and the role that the six condition indicators can play in assessing these undesirable conditions.

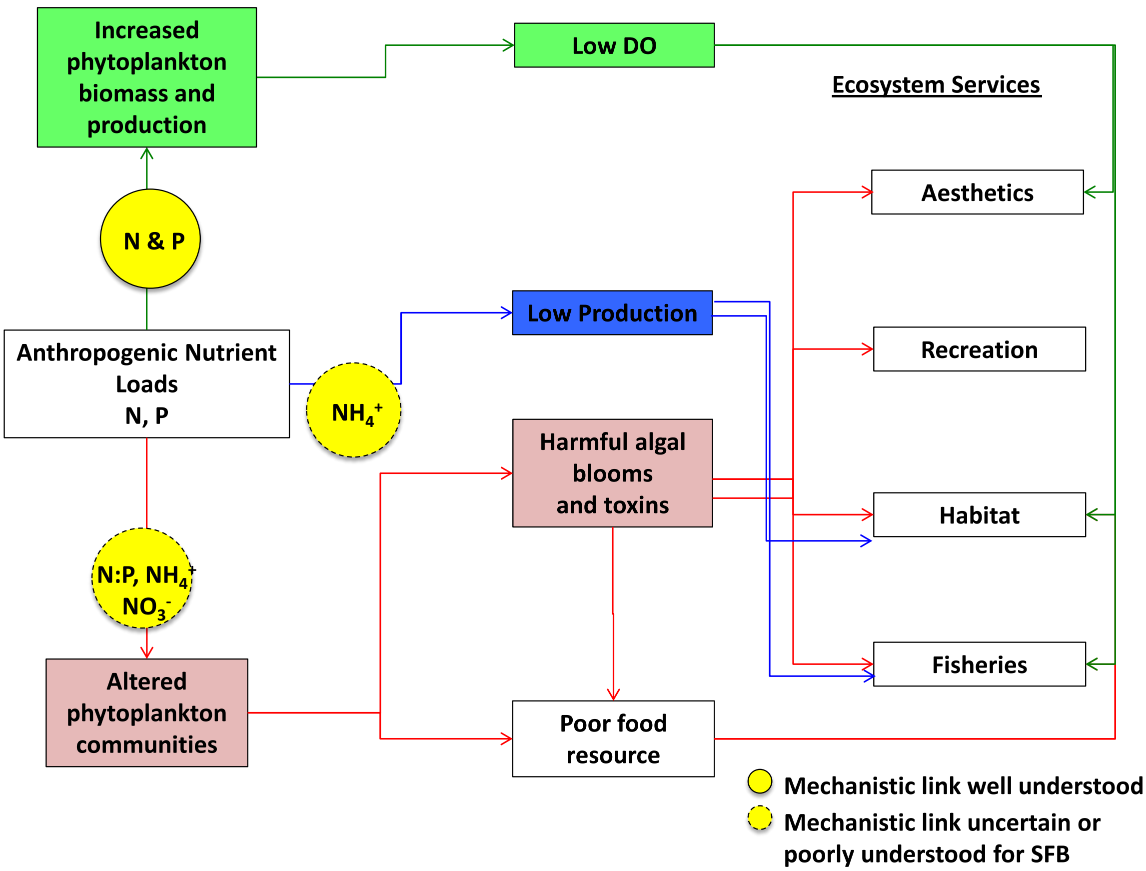


Figure 3.2 Potential adverse impact pathways: linkages between anthropogenic nutrient loads and adverse ecosystem response. The shaded rectangles represent indicators that are recommended for measurement along each pathway to assess condition. From SFEI 2014b).

Table 3.2 Plausible undesirable states and link to beneficial uses (adapted from SFEI 2014b).

Undesirable State (S)	Rationale or Link to Beneficial Uses
S1. High Phytoplankton Biomass and Productivity High phytoplankton biomass and productivity of sufficient magnitude, duration, and spatial extent that it impairs beneficial uses due to direct or indirect effects (S2-S3). This could occur in deep subtidal or in shallow subtidal areas.	Direct effects on noncontact water recreation (REC2) due to aesthetics via odors and surface scum. Other main concern is through increased organic matter accumulation causing low dissolved oxygen (S2-S3) and proliferation of pathogenic bacteria, leading to degraded contact and noncontact water recreation (REC1 and REC2).
S2 and S3. Low Dissolved Oxygen ***Deep subtidal:*** Low DO in deep subtidal areas of the Bay, over a large enough area and below some threshold for a long enough period of time that beneficial uses are adversely affected. ***Shallow/margin habitats***: DO in shallow/margin habitats below some threshold, and beyond what would be considered "natural" for that habitat, for a period of time that it impairs beneficial uses.	Fish kills, die-off of beneficial benthos, loss of critical habitat that result in lowered survival or spawning/reproductive success or recruitment success of fish and beneficial benthos. These consequences directly affects EST, RARE, etc. beneficial uses.
S4. HAB Abundance and Algal Toxins ***HABs and toxins:*** Occurrence of HABs and/or related toxins at sufficient frequency or magnitude of events that habitats reach an impaired state, either in the source areas or in areas to which toxins are transported. ***NABs:*** Occurrence of nuisance algal blooms with sufficient frequency and magnitude that they impair beneficial uses; for example, similar to the red tide bloom in Spring 2004	***HABs and toxins*:** Passive or active uptake of toxins, or ingestion of HAB-forming species and accumulation of toxins. Ingestion of bioaccumulated toxins is harmful to both wildlife and humans through consumption of toxins via shellfish or fish. Skin contact and inhalation can also be problematic. ***NABs*:** Some species are considered HABs for reasons other than toxins (e.g., directly impairing biota at very high levels, e.g., coating fish gills, birds wings, rapid biomass production leading to low DO). Impaired aesthetics, surface scums, discoloration, odors. These adverse effects directly impact EST, WILD, SHELL, RARE, and COMM beneficial uses.
S5. Low Phytoplankton Biomass and Productivity Low phytoplankton biomass in Suisun Bay or other habitats due to elevated NH_4^+, which would exacerbate food supply issues.	Suisun Bay is considered a food limited system, and low levels of phytoplankton biomass and productivity may contribute to impairment in this highly altered system. These adverse effects directly impact EST, SHELL, RARE, and COMM beneficial uses.
S6. Suboptimal Phytoplankton Assemblages that Impact Food Quality Nutrient-related shifts in phytoplankton community composition, or changes in the composition of individual cells (N:P), that result in decreased **phytoplankton food quality**, and have cascading effects up the food web.	Phytoplankton primary production is the primary food resource supporting food webs in SFB. Changes in the dominant assemblages and their relative size fractions would impact food quality. These adverse effects directly impact EST, SHELL, RARE, and COMM beneficial uses.
S7. Other Nutrient-Related Impacts Other direct or indirect nutrient-related effects that alter habitat or food web structure at higher trophic levels by other pathways. Several additional nutrient-related impacts on food webs in the northern estuary have been proposed that are not captured by S1-S6.	

High phytoplankton biomass and primary productivity (S1, Table 3.2) can have direct effects on REC2 in SFB via nuisance scums and odors.

However, among the most common and problematic impairments due to high phytoplankton biomass is **low dissolved oxygen** (S2 and S3, Table 3.2) in subtidal areas that results through metabolism of phytoplankton-derived organic matter by oxygen-consuming microorganisms (e.g., Figure 3.3). Because aquatic organisms rely on DO for survival, growth and reproduction, the consequences of sub-optimal DO in SFB include die-offs or low production of fish and benthos and loss of critical habitat due to lowered survival or spawning/reproductive success or recruitment success (Figures 3.4). These adverse effects directly impact EST, SHELL, RARE, and COMM beneficial uses.

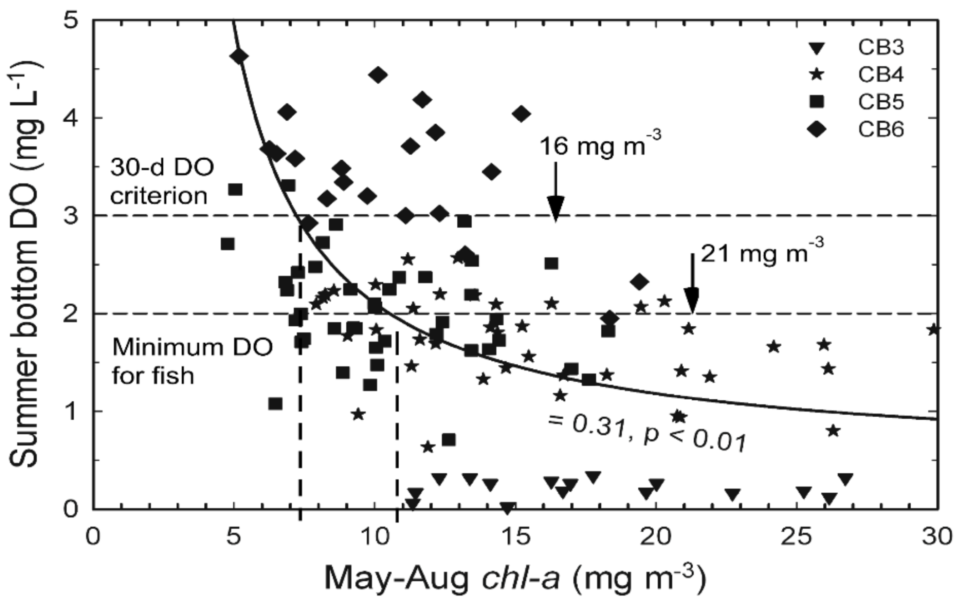


Figure 3.3. Example of dissolved oxygen as a function of chlorophyll-a in Chesapeake Bay. From Harding et al. 2013. Scientific bases for numerical chlorophyll criteria in Chesapeake Bay. *Estuaries and Coasts* doi:10.1007/s12237-013-9656-6

Elevated nutrient concentrations, or changes in relative abundance of nutrient forms, could increase the frequency with which **harmful algal blooms (HAB) and algal toxins** (S4, Table 3.2) occur, including abundance, duration, and spatial extent. Algal toxins, such as microcystin and domoic acid, bioaccumulate and can exert toxicity to consumers at all levels of the food web, including humans. Some HAB exudates also exert direct toxicity (e.g., skin contact). High nutrient loads may also increase the frequency of so-called nuisance algal blooms (NABs), which are not toxic but may degrade aesthetics due to surface scums or odors. Elevated phytoplankton biomass is typically correlated with increased probability of HABs (and NABs) and toxins (e.g., Figure 3.5).

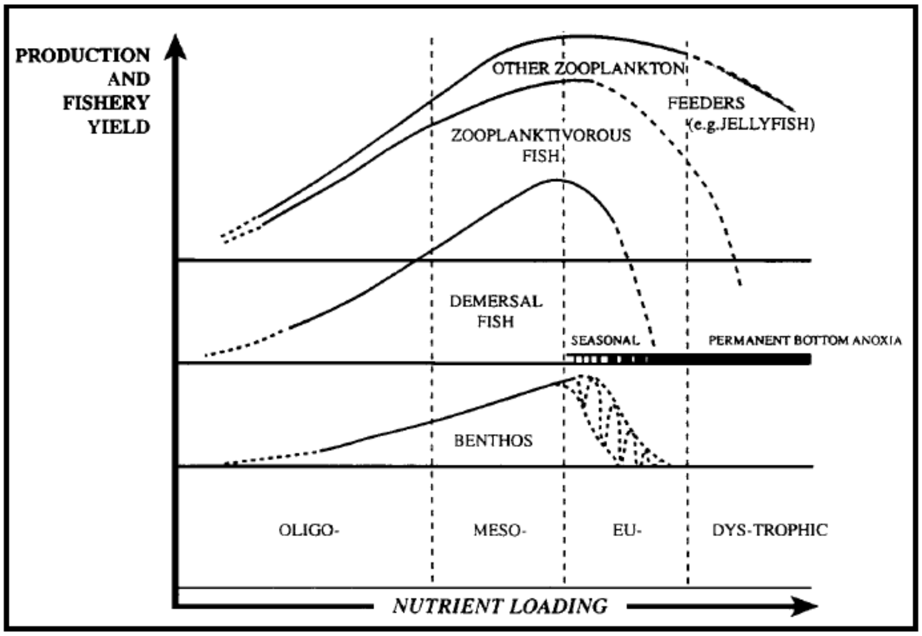


Figure 3.4. Comparative evaluation of fishery response to nutrients along continuum of oligotrophic, mesotrophic, eutrophic and dystrophic states of primary productivity (Nixon 1995). Although higher nutrient inputs initially increase the productivity of fisheries, ecological systems worldwide show negative effects as nutrient loading increases and hypoxic or anoxic conditions develop. Each generic curve in the lower half of the figure represents the reaction of a species guild to increasing nutrient supplies. From Diaz and Solow (1995).

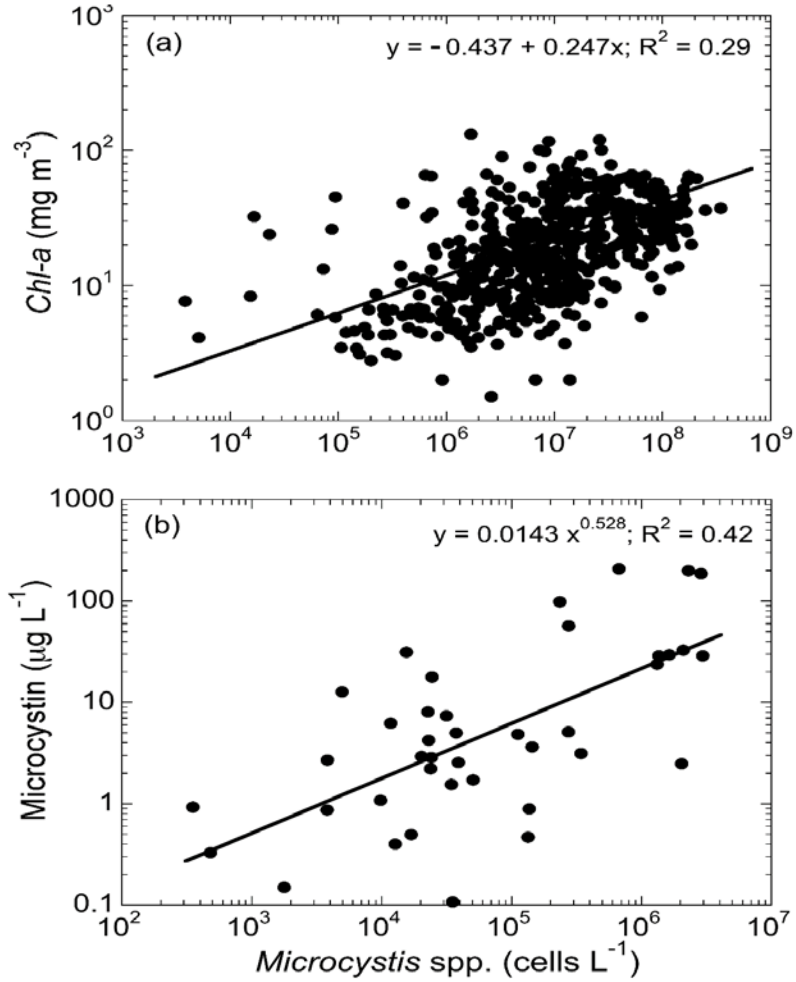


Figure 3.5. Example of relationships between chlorophyll-a, cyanobacteria Microsystis spp. abundance, and toxin concentrations, From L. W. Harding et al. 2013. Scientific bases for numerical chlorophyll criteria in Chesapeake Bay. ***Estuaries and Coasts*** **doi:10.1007/s12237-013-9656-6**

A number of factors can lead to **low phytoplankton biomass and productivity** (S5, Table 3.2) and **suboptimal phytoplankton assemblages that impact food quality** (S6, Table 3.2), a phenomenon marked by a shift in phytoplankton community composition away from assemblages found under minimally disturbed conditions, toward smaller, suboptimal compositions that do not adequately sustain organisms at higher trophic levels.

Two metrics have been discussed for measuring adverse changes to phytoplankton communities:

1) **Fraction of small-sized phytoplankton**: Fisheries yields are correlated to phytoplankton biomass (e.g., biovolume) and primary productivity (Friedland et al. 2012; Figure 3.6). When the portion of picophytoplankton (< 5 microns) grows, the result is a comparatively lower trophic transfer of energy and carbon up the food web (e.g., Figure 3.6) than is seen with other phytoplankton, which results in lower fisheries yields.

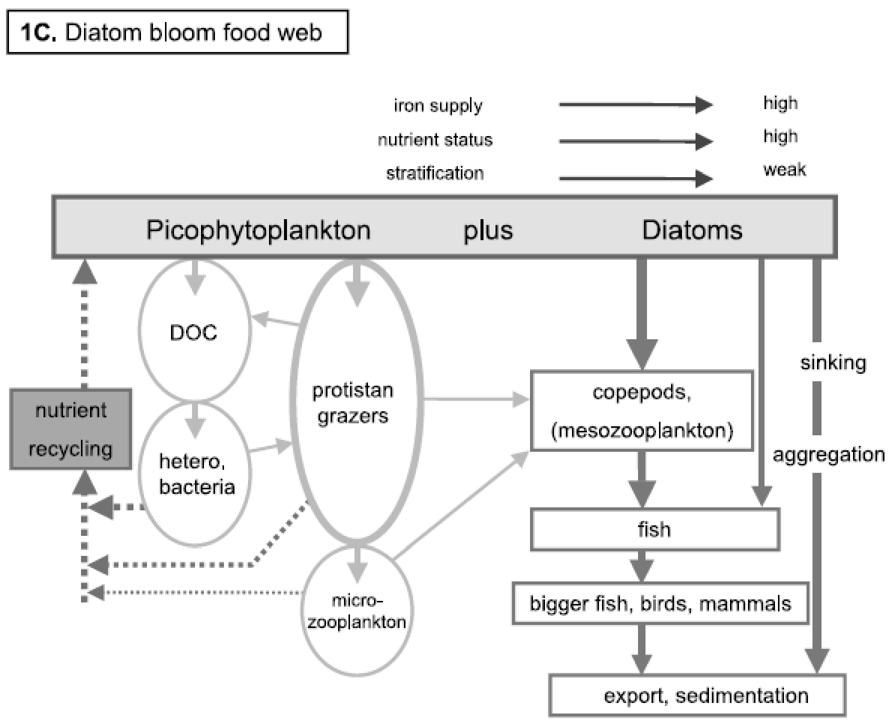


Figure 3.6. Example of a marine food web showing the complex pico-phytoplankton and diatom food web structure in diatom-dominated blooms. For simplicity, the regeneration paths are shown only on the left side of the figure (Source: Barber and Hisock 2006).

2) **Index of phytoplankton food quality**: This index utilizes data on phytoplankton composition to characterize the "food quality" that phytoplankton represent in supporting productivity of upper trophic levels. This is a key pathway to link phytoplankton composition to beneficial uses, such as commercial and recreationally important fisheries (i.e., EST, COMM, RARE). The concept of a phytoplankton food quality index is based on laboratory experiments showing that growth efficiency of crustacean zooplankton is highest when they are fed algae enriched in highly unsaturated fatty acids (cryptomonads and diatoms), and lowest when fed algae poor in these essential fatty acids (e.g., cyanobacteria; Brett and Müller-Navarra 1997).

Based on Galloway and Winder (2015), the fatty-acid food quality index (FQI) can be computed from the average composition of long chained essential fatty acids (LCEFA) at the algal taxonomic group level (Park et al. 2003, Galloway and Winder 2015).

The scale of the index (0–1; Equation 1) is defined by calculating the relative quality of each algal group (AGi) compared to the maximal LCEFA content of all AG:

Equation 1. FQI = AGcy*Pcy + AGgr*Pgr + Agdi*Pdi + AGcr*Pcr

where the FQI is the biovolume weighted average of the AGi for each individual group, and Pcy, Pgr, Pdi, and Pcr are the proportions of phytoplankton biovolume in a sample contributed by cyanobacteria, green algae, diatoms, and cryptomonads. Figure 3.7 shows the separation in AGi by phytoplankton taxonomic group. The concept has recently been applied to phytoplankton composition data collected by the USGS in the Lower Sacramento River through Suisun Bay from 1992 to 2014 (Cloern et al. 2015).

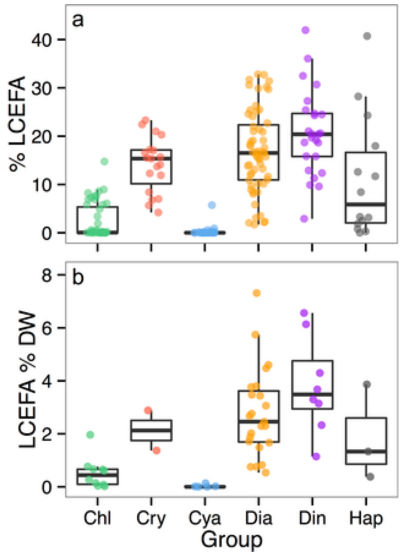


Figure 3.7. From Galloway and Winder 2015. Boxplots of species averages of Σ long-chain essential fatty acids (LCEFA) in six major phytoplankton groups. (a) Shows the percent total fatty acids (% FA) dataset, consisting of 208 averages from 666 raw profiles. (b) Shows the percentage of algal dry weight (FA % DW) dataset, consisting of 55 averages from 105 raw profiles. Group name abbreviations follow Fig 1. The heavy line is the median, box boundaries are the 25th and 75th percentiles, and whiskers extend to the most extreme value within 1.5*IQR (interquartile range). The y-axis is set to show the extent of whiskers; thus, some extreme outliers are not plotted (outliers were included in calculation of average group LCEFA).

We propose that a number of metrics for phytoplankton community composition be deployed in routine assessments of SFB. In addition to tracking HAB abundance and toxin concentrations, phytoplankton metrics should be developed with the intent to create classification schemes in the future, if warranted, as these metrics (in combination with chlorophyll-a and GPP, discussed in more depth in Section 3.2) can give a more robust understanding of SFB condition and ecological change.

One final note: Nutrient forms and ratios are not explicitly considered as metrics within the present AF, although they will most certainly be included within the framework of monitoring and mechanistic modeling. The reason is that while several authors have hypothesized that high nutrient concentrations, elevated NH_4^+, or altered N:P are currently adversely impacting food webs in SFB (Table 3.1, S6; Dugdale et al., 2007; Parker et al., 2012a,b; Dugdale et al., 2012), scientific consensus is lacking on the importance of these hypothesized pathways relative to other controls on phytoplankton production and community composition.

3.2 Protocols, Temporal and Spatial Frequency Recommended for Measurement of Key Indicators

An important attribute of an AF is clarity in the methods used to measure the indicators, as well as the temporal and spatial frequency in which they should be measured in order to make an assessment. Table 3.3 provides a list of six key indicators and the specific analytes associated with each. This table is not inclusive of the longer list of parameters required for data interpretation or for other Nutrient Strategy program elements. The SFB Monitoring Strategy (SFEI 2014c) provides a more comprehensive picture of those data needs, as well as specific recommendations on protocols for measurement of key indicators.

DO and metrics of phytoplankton quantity and quality are the two principal groups of indicators proposed for the SFB nutrient AF. The Water Board's basin plan already contains numeric objectives for DO, and Water Board staff has expressed interest in reviewing the existing DO objectives.

Table 3.3 Recommended indicators, analytes and basis for classification scheme.

Indicator	Analyte	Basis for Classification Scheme
Dissolved oxygen	Dissolved oxygen as % saturation and concentration	SF Water Board Basin Plan (2016)
Phytoplankton biomass	Water column chlorophyll-a	Analysis of existing data (Appendix C)
Depth integrated, annual gross and net primary production	Chlorophyll-a, photic depth and surface irradiance, recalibrated on a frequency to be determined by direct measures of GPP (per Cole and Cloern 1984)	Nixon (1995)
HABs abundance (Alexandrium spp, cyanobacteria[1], Pseudo-nitzchia spp., Dinophysis spp.)	Genus and/or species cell counts and biovolume	Existing state, federal or international guidance—Appendix C for specifics by HAB species
HAB toxin concentrations		Existing state, federal or international guidance
Phytoplankton composition	Genus and/or species cell counts	No classification scheme proposed.
	% of Biovolume < 0.5 microns	
	Phytoplankton Food Quality Index (Galloway and Winder 2015)	

[1] Cyanobacteria of interest include, but are not limited to, Cylindrospermopsis spp., Anabaena spp., Microcystis spp., Planktothrix spp., Anabaenopsis spp., Aphanizomenon spp., Lyngbya spp., Raphidiopsis spp., Oscillatoria spp., and Umezakia spp.

Review of the science supporting SFB DO objectives is beyond the scope of this initial phase of AF development. Thus, the present recommendations focus on phytoplankton indicators.

Until further work is undertaken to consider and refine DO objectives and/or optimize sampling, assessments of DO are assumed to occur at the same frequency and location as those for the phytoplankton indicators.

Because dissolved oxygen, phytoplankton biomass, productivity and phytoplankton composition are all extremely variable across both time and space, the following two sections outline recommendations regarding the temporal and spatial elements of the AF and how to align them with the monitoring program to optimize capturing this variability, while also balancing costs, logistics and power to detect trends.

Temporal Scales of Interest and Recommended Frequency

For phytoplankton indicators, four temporal components are of interest for documenting ecosystem change (Figures 3.8 and 3.9):

- Magnitude of spring blooms
- Emergence and magnitude of fall blooms
- Elevated baseline occurring during non-blooms periods (typically during June-September)
- Interannual variability and trends

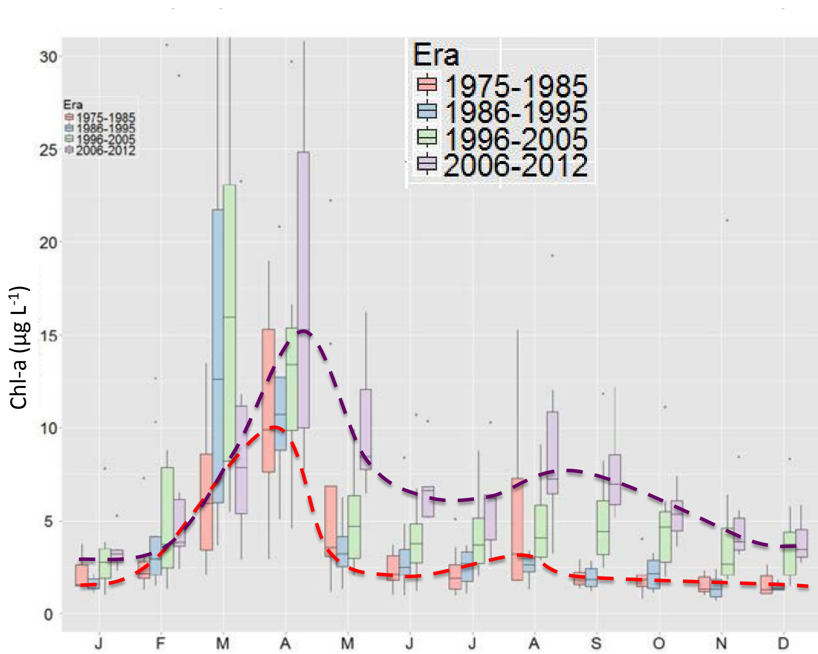


Figure 3.8. 10-year rolling average chlorophyll-a by month of the year in Lower South Bay, illustrating the four elements of interest in phytoplankton variability: (1) spring bloom, (2) fall bloom, (3) elevated baseline during non-bloom periods, and (4) interannual variablility. Source: Jim Cloern, USGS

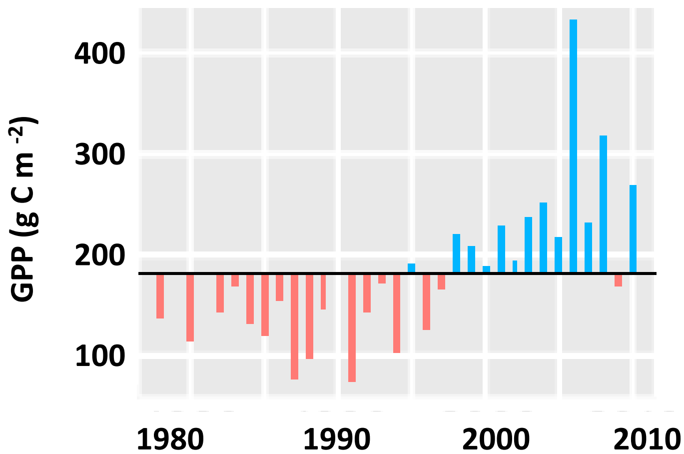


Figure 3.9. Trends in estimated annual GPP over time. From Cloern and Jassby (2012). Drivers of change in estuarine-coastal ecosystems: discoveries from four decades of study in San Francisco Bay. Rev. Geophys., 50, RG4001, doi:10.1029/2012RG000397.

Considering this variability, we recommend a sampling frequency of no less than monthly via ship-based sampling, with weekly sampling possible in order to better characterize bloom events.

Spatial Elements and Minimum Recommended Density

To adequately capture spatial gradients, we recommend sampling that encompasses (1) the SFB subembayments defined by Jassby et al. (1997), (2) both deep-channel parts and shallow parts of the Bay, (3) vertical gradients in the water column, either as grabs with depth or conductivity-temperature-depth (CTD) profiles, and (4) both the upstream, oceanic boundary conditions, as well as other potential "seed" sources of HABs, e.g., salt ponds.

We used best professional judgment to recommend preliminary placement of ship-based transects, water quality stations and moorings by subembayment (Figure 3.10). These locations should be considered provisional, subject to funding availability and optimization in concert with other nutrient strategy components that require monitoring (e.g., model development, etc.). Locations of historic USGS stations are preserved to maintain continuity of the long-term data set. Additional stations were added while balancing the logistics and cost of ship-based sampling. No stations are placed in tidal sloughs and restored salt ponds; consideration of monitoring in these habitats should be undertaken in a subsequent phase of AF development. Additional data analyses have been recommended to optimize the placement of stations (Senn et al. 2014).

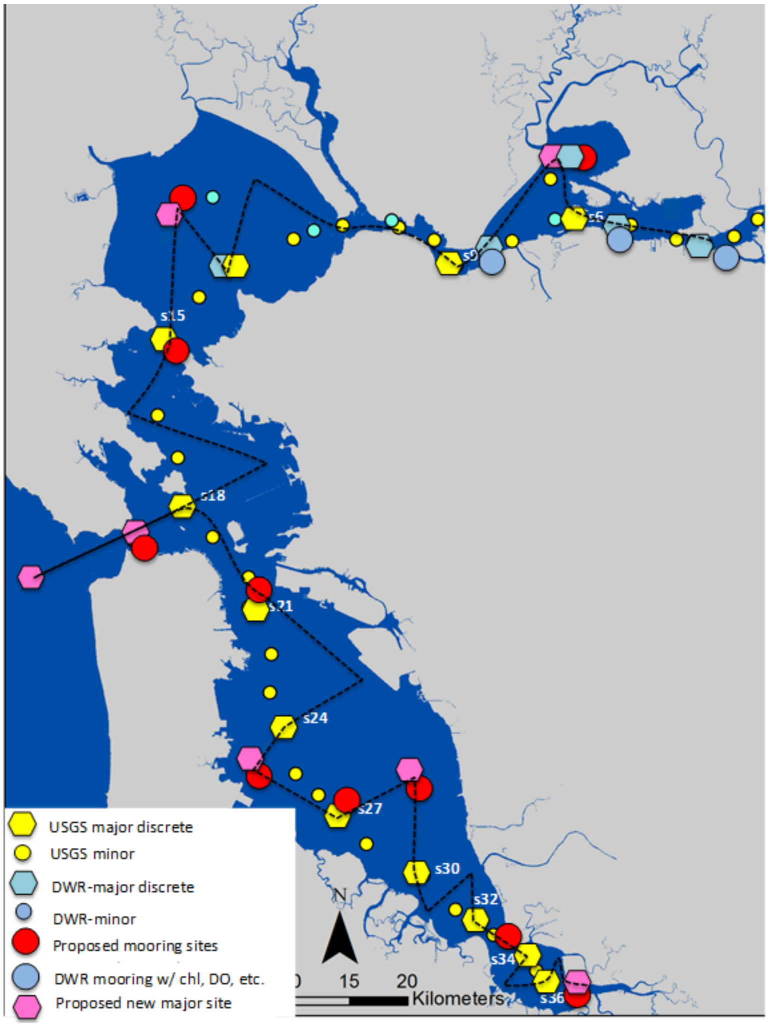


Figure 3.10. Recommendation of sampling stations representing minimum effort needed to support ambient nutrient assessment of SFB subembayments. Locations should be considered provisional, subject to funding availability and optimization in concert with other nutrient strategy components that require monitoring (e.g., model development, etc.).

3.3 Proposed AF Classification Tables, Justification, and Sources of Uncertainty

As noted above, we have proposed classification frameworks for five of the six indicators of SFB ecological condition: phytoplankton biomass (chlorophyll-a), gross primary productivity, HABs abundance, HABs toxins, and dissolved oxygen (Table 3.3).

For the sixth indicator – phytoplankton community composition – we explored two metrics that could be used to assess adverse changes (Section 3.1), and also made recommendations regarding temporal and spatial considerations (Section 3.2), but are stopping short of proposing a classification table for phytoplankton community composition.

Among the other five indicators, dissolved oxygen already has a classification table in use in SFB, and we recommend that the next step be a review of the need to refine the Basin Plan DO objectives (Section 3.2).

Our approach to developing classification tables for the four remaining indicators consisted of separating Bay subembayments into categorical bins of ecological condition, from high to low, based on indicators that are linked to ecosystem services (i.e., beneficial uses). An intent was made to be as explicit as possible on the precise metrics used to measure the indicators, as well as the temporal and spatial density of data required to make assessments and to specify how the data would be used to report on status and trends.

Existing guidance and the results of the quantitative analyses were synthesized, using expert opinion, into a classification scheme to assess ecological condition for multiple subembayments of SFB for each of the four indicators. For each indicator, a scheme was developed to parse SFB subembayments into a maximum of five ecological condition states (very high, high, moderate, low, very low), analogous to ecological condition frameworks developed for the European Union Water Framework Directive (Zaldivar et al. 2008). Existing guidance and quantitative analyses were used to inform the “thresholds” that define the range of values within each bin.

For most indicators, guidance exists in the form of established WQOs, state, federal or international guidance, or published studies that form the scientific foundation for their use in a classification scheme. For chlorophyll-a, we lacked confidence that an expert-derived existing guidance developed for estuaries around the world (e.g., Zaldivar et al. 2008) could be applied, without question, to SFB. For this reason, analyses of existing data were used to investigate the linkage between chlorophyll-a and potential pathways of impairment, detailed in Appendix C. Quantitative analyses and existing published guidance were supplemented by best professional judgment to address key data gaps and describe uncertainty and level of confidence in the classification.

For the purpose of reporting on status and trends, we recommend that classification occurs annually by subembayment, thus characterizing the spatial extent if the results are viewed on the whole for SFB or for each subembayment. The AF was designed to be applied using a data set that includes a minimum of monthly, ship-based discrete samples and CTD profiles, with spatial resolution given in Figure 3.10 (Senn et al. 2014a).

The following sections describe development of classification tables for each of the four indicators: phytoplankton biomass (chlorophyll-a), gross primary productivity, HABs abundance, and HABs toxins (the two HABs indicators are merged into one section). The final section offers recommendations regarding the future of indicator development work for dissolved oxygen.

Phytoplankton Biomass (Chlorophyll-a)

Chlorophyll-a has formed a cornerstone of standardized approaches to assess eutrophication (Bricker et al., 2003, Zaldivar et al. 2008) and to support regulatory water-quality goals in estuaries (Harding et al., 2013) because it is a well-recognized indicator that integrates nutrient loadings and represents adverse effects to ecosystems. Decisions based on quantitative endpoints can be based on deviations from "reference" conditions, or on quantitative relationships with ecosystem impairments (e.g., Harding et al. 2013). In SFB, records of chlorophyll-a prior to human disturbance are not available, complicating development of reference chlorophyll-a ranges. An extensive, multi-decadal dataset is available to explore quantitative relationships between chlorophyll-a and potential pathways of adverse effects, as a means for establishing chlorophyll-a endpoints.

We analyzed a multi-year dataset that included chlorophyll-a (1993-2014), phytoplankton species composition (1993-2014), DO (1993-2014), and algal toxins (2012-2014) to (1) explore trends in HAB abundance, toxins, and DO concentrations and their relationships with chlorophyll-a, and (2) quantify chlorophyll-a thresholds and related uncertainty that correspond to categories of "protected" and "at risk" in the context of current DO WQOs and HAB alert levels. Quantile regression and conditional probability analysis were used to identify thresholds of chlorophyll-a, corresponding to categories of increasing risk in the context of current DO WQOs (SFRWQCB 2015) and HAB alert levels (Appendix C).

We found that HAB toxins and species can be routinely detected in SFB subembayments. Increased occurrences of HAB species and declining DO were correlated with increased chlorophyll-a over the 20-year period. Monthly chlorophyll-a "thresholds" corresponding to increased risk of HABs were identified, aggregating across all subembayments. The analyses were also sufficiently robust to estimate chlorophyll-a thresholds relating to DO for South Bay and Lower South Bay. Taken together, these analyses were used to support a preliminary set of chlorophyll-a assessment thresholds aimed at defining a gradient of ecological condition (from low to high risk) for increased HAB events and low DO in SFB subembayments.

Classification of chlorophyll-a linked to HABs is based on a monthly timescale because the HAB alert guidance is based on acute risk. In contrast, classification based on the linkage to dissolved oxygen was based on the mean concentration of monthly values from February to September, the time period in which biomass has been observed to be changing over the last two decades in SFB. This difference in temporal statistic reflects a more contemporaneous linkage between chlorophyll-a and HABs, as compared to the lagged response of organic matter production and the eventual increased potential for DO depletion. For DO, the differences in classification by subembayment reflect regional differences in hydrogeographic factors affecting DO dynamics.

Classification of Chlorophyll-a Linked to HABs. Categorization of monthly mean chlorophyll-a is directly linked to the outcome of quantile regressions and CPA relating the acute risk of HABs as a function of increased chlorophyll-a (Table 3.4, Appendix C: Figures 6-8). The highest category of ecological condition is defined by monthly mean chlorophyll-a values < 13 mg m^{-3}, which represents a baseline probability of ~0.39 to 0.4 for HAB abundance and ~0.3 for domoic acid and microcystins. Ecological condition is downgraded as monthly values in the range of 13-25 mg m^{-3} show increased probabilities of exceeding HAB alert values to up to 0.44 for HAB abundance and 0.6 or greater for toxins. Chlorophyll-a concentrations in the range of 40 mg m^{-3} represent a 0.5 to 0.68 probability of a HAB event; while there are only two data points for toxins between 20-60 mg m^{-3}, the CPA suggests a probability of 0.6-0.7 within this range of chlorophyll. Occurrence of HABs on a more frequent basis represents a potentially chronic exposure to toxins (e.g. Ger et al. 2009; Goldstein et al. 2008), and thus, condition is downgraded as the annual frequency of occurrence in monthly samples increases (Table 3.4).

For context, on a Bay-wide scale, 13 mg m^{-3} corresponds to the 90^{th} percentile of monthly surface chlorophyll-a over the last 20 years. On a sub-embayment scale, Central, North Central, San Pablo and Suisun Bay stations were below 13 mg m^{-3} for greater than 95% of the time over the last 20 years. The range of chlorophyll-a at Lower South Bay and South Bay stations was slightly higher. The ranges were below 13 mg m^{-3} 74% and 85% of the time, respectively, in Lower South Bay and South Bay, and below 25 mg m^{-3} 88% and 93% of the time, respectively (Figure S3, supplemental materials in Appendix C).

Table 3.4. Chlorophyll-a Classification Table Linked to HAB Abundance, Based on Annual Frequency of Occurrence in Monthly Samples. Classification should be applied to each subembayment.

Subembayment Monthly Mean Chlorophyll-a Linked to HAB Abundance (µg L^{-1})	**Ecological Condition Based on Annual Frequency of Occurrence in Monthly Samples**			
	1 of 12	**2-3**	**4-6**	**6+**
≤ 13	Very high	Very high	Very high	Very high
>13 – 25	Good	Moderate	Moderate	Low
>25 – 40	Moderate	Moderate	Low	Very Low
>40 – 60	Moderate	Low	Very Low	Very Low
>60	Low	Very low	Very low	Very low

Classification of Chlorophyll-a Linked to DO. While chlorophyll-a was negatively correlated with DO in all subembayments, only in South Bay and Lower South Bay were these relationships consistently significant to quantify thresholds supporting classification decisions. Conceptually, the mechanism resulting in an expected negative relationship between summer DO and February-September mean chlorophyll-a is that high primary production during this time scale is expected to promote increased abundance of planktonic and benthic detritus, which during summer leads to an increasing probability of net ecosystem heterotrophy (Caffrey 2003). In some areas of San Francisco Bay, and at some times in all subembayments of the Bay, biological effects on DO are dominated by physical processes such as fluvial transport, stormwater and treated wastewater inputs, water exchange between subembayments, and mixing or exchange between habitats within a subembayments (Smith and Hollibaugh, 2006). The modulating factors are generally very important in both Central and Suisun Bays, which are most proximal to and have greater exchange with the coastal ocean and the Delta, respectively. It may still be possible to establish chlorophyll-a thresholds at which DO will begin to decline to unacceptable levels in the Central and North SFB subembayments, using other modeling approaches than what was employed by Sutula et al. (in prep, Appendix C).

In developing a chlorophyll-a classification scheme linked to DO for South and Lower South Bays, we relied principally on the predicted chlorophyll-a thresholds produced from quantile regressions of DO concentration that represent a range of ecological condition, from 7 to 4 mg L^{-}

[1] (Table 3.4, Appendix C: Tables 1-2). We note that the three-month median percent saturation WQO of > 80% is ~ 7 mg L^{-1} at summertime mean temperature and salinity in South SFB. According to the proposed European Union Water Framework Directive (EU WFD) for classification of estuarine waters based on DO (Best et al. 2007), 5.7 mg L^{-1} at marine salinities is equivalent to 7 mg L^{-1} in freshwater criteria, with chronic values considered to be supportive of salmonid reproduction and survival, which is not a designated use in South SFB. Thus, the "very high" tier of 7.0 mg L^{-1} is roughly equivalent to meeting the three-month median percent saturation objective, while the "moderate" condition category has 90% probability that the 5 mg L^{-1} concentration objective would be met (Table 3.5). This approach is comparable, though with higher expectations, than is used in Best et al. (2007). Without specific analyses that clarify the seasonal and habitat-specific DO acute and chronic criteria required to support beneficial uses, we have more heavily weighted our DO classification bins to align with existing SFB WQOs. We used the lower 95% confidence interval of the predicted 0.1 Tau quantile of February to September mean chlorophyll-a (Sutula et al., in prep, Appendix C) as the basis for the classification bin, because it gives greater confidence that chlorophyll-a falls above the predicted lower end of the classification bin.

Table 3.5. Chlorophyll-a Classification Table Based on Risk of Falling Below DO Water Quality Objectives, Based on Annual February-September Mean Chlorophyll-a, for South Bay and Lower South Bay only.

Classification of ecological condition based on mean February - September chlorophyll-a (mg m^{-3}) linked DO benchmarks - South Bay and Lower South Bay Only		
Category	**Lower South Bay**	**South Bay**
Very high)	≤23	≤14
High		>25 - 32
Moderate	>23 - 35	>32 - 44
Low	>35 - 51	>44 - 58
Very Low	>51	>58

In South Bay, quantile regression results provided in Appendix C suggest that a February to September mean chlorophyll-a of 13-16 mg m^{-3} is "protective" of the three-month median DO percent saturation WQO (80% or ~7 mg L^{-1} at summertime mean temperature and salinity in South SFB). At a February-September mean of 13 mg m^{-3}, 90% of the DO is predicted to be above 7 mg L^{-1}, while at 42 mg m^{-3}, 90% of the DO is predicted to be above 5.0 mg L^{-1} (Appendix C: Table 2). Ninety-five percent of the February-September mean chlorophyll-a measured at South Bay sites over the 20-year record is below 14 mg m^{-3} (Appendix C: Figure A4), reflecting the fact that primary production in combination with physics in the deep channel habitat of South Bay promotes largely normoxic conditions – greatly improved from the periods of hypoxia recorded prior to implementing advanced wastewater treatment in the 1970s (Cloern and Jassby, 2012). Uncertainty in this classification is low (see 95% confidence intervals, Appendix C: Table 2), given the significance of the quantile regression. However, we note that existing data were limited to ship-based data that do not capture a diel curve, contributing to uncertainty that existing relationship does not capture true DO minima. These analyses should be repeated with continuous DO data that better characterizes physical and biological exchanges with the shallow water margin habitat. Such data do not exist and we recommend that they be collected.

CPA and quantile regressions were also used to support a chlorophyll-a classification scheme for Lower South Bay, albeit with more uncertainty than for South Bay. The reasons for this greater uncertainty are two-fold. First, biological and physical exchanges between Lower South Bay and the adjacent shallow margin habitats are unquantified. While CPA analyses could only be used to suggest a threshold in which the subembayment is "at risk" of falling below the 80% percent saturation WQO (~13 mg m^{-3}), neither CPA nor quantile regression could be used to derive a chlorophyll-a value that would be "protective" of the percent saturation WQO. It is likely that an additional source of DO water < 80% saturation (from either the tidal slough or restored salt ponds) is exchanging with Lower South Bay deep channel habitat. These margin habitats have been documented to routinely fall below 5 mg L^{-1} DO on diel timescales (Thebault et al, 2008; Shellenbarger et al, 2008, SFEI 2014a). Considering that these intertidal habitats rich in organic carbon may have natural sources of low DO water, the expectations for DO in these habitats and their physical and biological exchanges with Lower South Bay need to be considered in setting appropriate expectations for Lower South Bay deep channel habitat (Sutula et al. 2012, Bailey et al. 2014). Second, it is noteworthy that while these data show that Lower South Bay is meeting the 3-month median DO saturation objective only 72% of the time, it is above 5 mg L^{-1} 97% and above 5.7 mg L^{-1} 90% of the time over the past 20 years, with 95% of the February to September mean chlorophyll-a less than 25 mg m^{-3}. Best et al. (2007) have proposed > 5.7 mg L^{-1} as a benchmark to represent the highest ecological condition category for estuaries assessed under the European Union Water Framework Directive. Given this, it will be helpful to review the science supporting existing DO WQOs in SFB specifically with respect to both deep water and shallow margin habitats, as is currently being done for Suisun Marsh as part of development of a DO TMDL (Bailey et al. 2014).

Major Sources of Uncertainty in Chlorophyll-a Classification. Overall, uncertainty exists in this proposed chlorophyll-a classification framework and our ability to quantify that uncertainty is constrained. Five major types of uncertainties exist in the chlorophyll-a framework linked to HABs and DO impairment pathways: (1) significance of the ecological and human risk of HABs in SFB, (2) linkage of chlorophyll-a to HAB cell counts, rather than toxin concentrations, as the foundation for the risk paradigm; SPATT toxin data were used to supplement the analyses, but the calibration of SPATT relative to particulate or mussel toxin tissues is still ongoing and should be a continued management focus, (3) uncertainty in the risk to aquatic life, since the HAB alert levels are focused on risk to human health rather than aquatic life, (4) uncertainty in capturing risks of chronic exposure to HABs, stemming from the fact that alert levels are based on acute toxin exposure, (5) the underlying mechanism of the correlation between February-September chlorophyll-a and summer DO, and (6) appropriate DO expectations for shallow water margins, tidal sloughs and intertidal wetland habitat, and portions of the SFB open water habitat that are strongly linked to the margins (e.g. LSB).

Our classification tables for chlorophyll-a are somewhat distinct from the other indicators in that they rely on relationships with other SFB attributes (e.g. HAB abundance and DO). We know from other long-term observational programs that changes can also include shifts in the efficiency with which nutrients are assimilated into algal biomass (Riemann et al. 2015). SFB's high nutrient concentrations imply a potential to produce phytoplankton biomass at levels that impair water quality. To illustrate this point we computed median concentrations of dissolved inorganic nitrogen (DIN) and chl-a a across four subembayments of the estuary (Appendix C: Table 3). We then computed potential chl-a as the sum of measured chl-a plus the quantity of chl-a that could be produced if all remaining DIN was assimilated into phytoplankton biomass, assuming a conversion factor of 1 g chl-a per mol N (Eppley et al. 1971). If this potential is realized then the median chl-a concentrations in all Bay subembayments would increase by an order of magnitude. Given the uncertainty in SFB's trajectory amidst global change, it is this potential for high biomass production that motivates establishment of chl-a thresholds to support nutrient management in SFB. Though we like to think of these relationships as fixed, in reality, these chl-a thresholds can change as fundamental drivers such as oceanic exchange, top-down grazing, light limitation, etc. that control the nature of the relationship between chl-a, HAB cell density and DO can change with climate variability and climate change, (Cloern et al. 2014, Riemann et al. 2015).

This point underscores the critical need to continuously reevaluate these relationships through a long-term consistent monitoring program in SFB. A consistent monitoring program would go a long way to reduce some of the remaining uncertainties in the existing data, given the large data gaps and inconsistent available data between sites, for the analyses conducted here (Sutula et al, (in prep), Appendix C).

Gross and Net Primary Production

Annual GPP is proposed as an AF indicator, to be measured via an empirical method utilizing chlorophyll-a, photic depth, surface irradiance (per Cole and Cloern 1984), recalibrated with specified direct, discrete measures of GPP (e.g., Cloern et al. 2014). GPP is complementary to chlorophyll-a, which does not provide a direct measure of the internal supply rate of biological oxygen demand, nor the rate of turnover of phytoplankton carbon. Annual GPP would be assessed based on the identical temporal and spatial data collected to support chlorophyll-a.

Decisions on classification thresholds for GPP were based on Nixon (1995), who proposed definitions of the trophic state of estuaries as oligotrophic (< 100 g C m^{-2} yr^{-1}), mesotrophic (100-300 g C m^{-2} yr^{-1}), eutrophic (>300-500 g C m^{-2} yr^{-1}), and hypereutrophic (> 500 g C m^{-2} yr^{-1}). For the purposes of assessment of SFB subembayments, we collapsed these into three categories (Table 3.6). Hypereutrophic represents the boundary between moderate and low/very low ecological condition (>500 g C m^{-2} yr^{-1}). Oligotrophic and mesotrophic are combined into one category (very high/high ecological condition), expressly to avoid categorizing very low production values as indicative of very high ecological condition, since some level of production is considered important.

Nixon did not specify a method for measurement of GPP; Cloern et al. (2014) documented how differences among methodologies can have a large impact on estimated GPP. We propose confirming proposed GPP classification boundaries using the SFB water quality model, once calibrated for DO, in order to provide an additional confirmation of these proposed classification thresholds.

Table 3.6. Gross Primary Productivity Classification Table Based on Annual Rate (g m^{-2} yr-^{1}). Classification should be applied to each subembayment.

Category	Gross Primary Productivity (g m^{-2} yr-^{1})
Very high/High	≤300
Moderate	>300 - 500
Low/ Very Low	≥ 500

Major Sources of Uncertainty in Classification of GPP. The greatest source of uncertainty in the proposed GPP classification is the lumping of highly oligotrophic GPP into the highest category. We acknowledge that, while it would be desirable to identify some level of GPP that is too low, the Expert Workgroup felt that we did not have the scientific basis to determine at what level that is. This remains a source of uncertainty in this classification. Another source of uncertainty is the use of an indirect approach to estimate GPP. Although other sources of uncertainty in estimates of GPP exist (e.g. short term pulses missed by monthly sampling programs, Gallegos and Neele, 2015), we feel that if these indirect estimates are calibrated on a frequent basis with direct measures, this uncertainty will be constrained.

HAB Abundance and Toxins

Classification of HAB cell counts and toxins is based on the assumption that values exceeding thresholds or alert levels used in comparable systems (Table 3.7), or trends of increasing occurrence, are evidence of reduced water quality. This is consistent with findings from the U.K. Undesirable Disturbance Study Team (Tett et al. 2007) and is supported by recent syntheses examining the relationship between HABs and coastal water quality (Heisler et al. 2008; Anderson et al. 2008).

Table 3.7. Potential HABs from San Francisco Bay, and alert levels used in other regions.

Organism	***Alert Level (cells/L)***	***Reference***
Alexandrium spp.	Presence	http://www.scotland.gov.uk/Publications/2011/03/16182005/37
Blue-Green Algae	100,000	WHO, 2003; California Guidance (OEHHA, 2012)
Dinophysis spp.	100-1,000	http://www.scotland.gov.uk/Publications/2011/03/16182005/37; Vlamis et al. 2014
Heterosigma akashiwo	500,000	Expert opinion
Karenia mikimotoi	500,000	Expert opinion
Karlodinium veneficum	500,000	Expert opinion
Pseudo-nitzschia	10,000-50,000	Cal-HABMAP ; Shumway et al. 1995; Anderson et al. 2009

The classification scheme assumes data collection similar to the USGS monitoring program data described above, and includes regular (monthly) monitoring of phytoplankton species and total (particulate and dissolved) toxin from the top 2 m of the water column using grab samples, deployment of SPATT or similar integrative samplers as part of Bay-wide surveys, and targeted collection of tissue samples from bivalves and marine mammals. For the assessment, the expert working group assumed maximum toxin concentration and maximum cell abundance by Bay subembayment would be used as a metric because of the potential risk to human and ecosystem health, and the likelihood of undersampling given the relatively coarse temporal and spatial scales. As with the classification scheme for chlorophyll-and DO, we consider this initial set of recommendations to be hypotheses that should undergo further testing and refinement when more data are available.

Classification of HAB Toxins. Guidance for toxins is currently restricted to domoic acid, microcystins, and paralytic shellfish toxins (PSTs) since those three classes of toxins are both persistent and regulated in the State of California. The scheme could be extended to other toxins given sufficient information about acceptable levels. Since existing guidance is based on acute exposure or Tolerable Daily Intake (e.g. World Health Organization guidelines for microcystins), we did not include a "duration" of exposure, and consider chronic effects to be an area of emerging concern (e.g., Ger et al. 2009; Goldstein et al. 2008; Hiolski et al. 2014) that should be considered as more data become available.

For toxin concentrations, progressions among classification bins are treated the same, based on existing alert levels, where we classify 50% of the regulatory closure level as a "warning level" and the closure limit as a (regulatory) action level. Ecological condition states are therefore: non-detect to 10% of the warning level, 10-100% of the warning level, above the warning level and below an action level, and above an action level. Since there is no direct correlation between SPATT toxin concentrations and grab sample concentrations, we assigned categories based on historical data from the region, corresponding to those categories and based on comparison of SPATT with grab and tissue samples (Lane et al. 2010; Kudela 2011). We acknowledge that this is a weak point of the classification scheme and a major source of uncertainty, but the advantages of SPATT for routine monitoring (Mackenzie et al. 2004) outweigh these concerns.

Tables 3.8, 3.9, and 3.10 provide classification schemes for microcystins, domoic acid, and saxitoxins. Note that SPATT is not routinely used for saxitoxins and has been omitted from Table 3.10. For microcystins, water concentrations are based on OEHHA 2012 guidance, which sets the alert level for recreational contact, domestic animals, and livestock at 0.8 ppb for microcystins LR, RR, YR, and LA. For mussel tissue, values are based on WHO guidance of 0.04 µg/kg body weight per day, assuming 100 g consumption of tissue and a 60 kg individual; it is assumed that these values can be scaled to other organisms. Tables 3.9 and 3.10 provide the same classification scheme for domoic acid and paralytic shellfish toxins. Alert levels are based on California Department of Public Health guidelines for tissue of 20 ug/g for domoic acid and 80 ug/100g for PSTs for protection of human health. For all three toxins, annual assessment of ecological condition would be based on the lowest rating for the year to provide the most protective classification.

Table 3.8. Toxin Classification Table for Microcystin. Classification should be applied to each subembayment. If multiple occurrences in different media (particulate, SPATT, tissue) are detected within a subembayment on an annual basis, the lowest rating for the year should be applied.

Toxin Concentration	**Ecological Condition Based on Annual Frequency of Occurrence in Monthly Samples**			
	1 of 12	2-3	4-6	6+
Particulate concentration				
Non-detect	Very high	Very high	Very high	Very high
Detectable, but < 0.8 ppb	High	Moderate	Moderate	Low
0.8 - 20 ppb	Moderate	moderate	Low	Very Low
>20 ppb	Low	Very Low	Very Low	Very Low
SPATT				
Below the warning level <100 ng/g)	Very high	Very high	Very high	Very high
100-250 ng/g	Moderate	Low	Very low	Very Low
>250 ng/g	Low	Very Low	Very Low	Very Low
Mussel Tissue				
Non-detect	Very high	Very high	Very high	Very high
Detectable, but < 12 ng/g	High	Moderate	Moderate	Low
12-24 ng/g	Moderate	moderate	Low	Very Low
> 24 ng/g	Low	Very Low	Very Low	Very Low

Table 3.9. Toxin Classification Table for Domoic Acid. Classification should be applied to each subembayment. If multiple hits in different media (particulate, SPATT, tissue) are detected within a subembayment on an annual basis, lowest rating for the year should be applied.

Toxin Concentration	**Ecological Condition Based on Annual Frequency of Occurrence in Monthly Samples**			
	1 of 12	2-3	4-6	6+
Particulate concentration				
Non-detect	Very high	Very high	Very high	Very high
0-100 ug/L	High	Moderate	Moderate	Low
100 - 1000 ug/L	Moderate	moderate	Low	Very Low
> 1000 ug/L	Low	Very Low	Very Low	Very Low
SPATT				
<30 ng/g	Very high	Very high	Very high	Very high
30-75 ng/g	Moderate	Low	Very low	Very Low
>75	Low	Very Low	Very Low	Very Low
Mussel Tissue				
Non-detect	Very high	Very high	Very high	Very high
< 10 ppm	High	Moderate	Moderate	Low
10-20 ppm	Moderate	moderate	Low	Very Low
> 20 ppm	Low	Very Low	Very Low	Very Low

Table 3.10. Toxin Classification Table for Paralytic Shellfish Toxins. Classification should be applied to each subembayment. If multiple hits in different media (particulate, SPATT, tissue) are detected within a subembayment on an annual basis, lowest rating for the year should be applied.

Toxin Concentration	Ecological Condition Based on Annual Frequency of Occurrence in Monthly Samples			
	1 of 12	2-3	4-6	6+
Particulate Concentration				
Non-detect	Very high	Very high	Very high	Very high
Detectable	Low	Very low	Very low	Very Low
Mussel Tissue				
Non-detect	Very high	Very high	Very high	Very high
< 40 µg/100 g	High	Moderate	Moderate	Low
40-80 µg/100 g	Moderate	moderate	Low	Very Low
> 80 µg/100 g	Low	Very Low	Very Low	Very Low

Classification of HAB Abundance. The classification scheme for presence of HAB organisms is based on a similar metric as for toxins (Table 3.11). An alert level is defined based on existing monitoring programs, and condition is graded based on expert opinion relative to those alert levels. For *Alexandrium* specifically, because all monitoring programs consider presence of *Alexandrium* to be a potential impairment, only three cell abundance categories are used (not detected, detected at up to 100 cells/L, and more than 100 cells/L). For BGA, the criteria are restricted to stations or locations where salinity is less than or equal to 2, and the alert level is based on OEHHA 2012 guidance of 1E6 cells/mL (i.e., scum-forming blooms). Given the prevalence of BGA toxins in SFB (Appendix C-Figure 3), more conservative cell abundances were chosen for transitions from high Very High to Very Low condition compared to an alert threshold of 1E6 cells/mL.

Table 3.11. HAB Abundance Classification Table. Classification should be applied to each subembayment. If multiple HABs are detected within a subembayment on an annual basis, lowest rating for the year should be applied.

Cell Count By Taxonomic Group	Ecological Condition Based on Annual Frequency of Occurrence in Monthly Samples			
	1 of 12	2-3	4-6	6+
Cyanobacteria[1]. Applies at salinities ≤ 2 ppt.				
Absent to < 20,000 cells per ml	Very high	Very high	Very high	Very high
20,000 – 10^5 cells per ml	High	Moderate	Low	Very Low
10^5 – 10^7 cells per ml	Moderate	Low	Very Low	Very Low
> 10^7 cells per ml	Low	Very Low	Very Low	Very Low
Pseudo-nitzchia spp.				
<100 cells per l	Very high	Very high	Very high	Very high
100 to 10,000 cells per l	High	High	Moderate	Low
10,000 -50,000 cells per l	Moderate	Low	Low	Very Low
> 50,000 cells per l	Low	Very Low	Very Low	Very Low
Alexandrium spp.				
Non detect	Very high	Very high	Very high	Very high
Detectable to < 100 cells	High	Moderate	Low	Very low
>100 cells	Low	Very low	Very low	Very Low

[1] Cyanobacteria include: Cylindrospermopsis, Anabaena, Microcystis, Planktothrix, Anabaenopsis, Aphanizomenon, Lyngbya, Raphidiopsis, Oscillatoria, and Umezakia

Uncertainty Associated with HAB Abundance and Toxin Classification. There are three major sources of uncertainty associated with the classification of HAB abundance and toxin concentrations. The first source derives from the use of existing guidance on cell counts and toxin concentrations. Standard guidelines have not been adopted at the State or federal level. Second, while HABs represent a palatable risk to human and ecological threat in SFB, uncertainty exists in the significance of that threat. For humans, the uncertainty lies in the level of risk given the amount of contact and noncontact recreation that occurs, as well as consumption of shellfish from SFB. Improved data on the concentrations of toxins in mussel tissue and shellfish consumption survey may help to better quality that risk. For aquatic organisms, this risk is difficult to characterize, particularly because existing guidance is oriented towards human health rather than ecological endpoints and on acute rather than chronic exposure to toxins. Because of the high baseline of HAB occurrence in SFB, uncertainty about values corresponding to this pathway of chronic exposure becomes a significant concern. The third source of uncertainty is the inclusion of SPATT-derived toxins in the classification scheme. SPATT as a tool has not undergone rigorous calibration. Because of its utility as a monitoring tool, calibration of SPATT relative to particulate or mussel toxin tissues should be a continued management focus.

Dissolved oxygen

Dissolved oxygen (DO) is considered to be keystone indicator within the AF. DO is necessary to sustain the life of all aquatic organisms that depend on aerobic respiration and, thus, it has a direct linkage to aquatic life and beneficial use protection (see Sutula et al. 2012 for comprehensive review). Eutrophication produces excess organic matter that fuels the development of hypoxia and, in some cases, anoxia as that organic matter is respired (Diaz 2001). Low dissolved oxygen (DO) has direct effects on the reproduction, growth and survival of pelagic and benthic fish and invertebrates (USEPA 2000, Bricker et al. 2003, Best et al. 2007). The response of aquatic organisms to low dissolved oxygen will depend on the intensity of hypoxia, duration of exposure, and the periodicity and frequency of exposure (Rabalais and Harper 1992). Thresholds for assessment of effects of DO are derived from criteria deemed to be protective of the most sensitive species from acute (timescales of days) and chronic (time scales of weeks to months) exposures to low dissolved oxygen.

In this work, we chose explicitly to defer work on a classification scheme for DO, citing the need to prioritize the development of classification for phytoplankton related indicators and the fact that DO objectives already exist for SFB. The following recommendations are intended to encourage future discussion of DO classification schemes for SFB, given that no scheme is being proposed at this time.

Existing DO WQOs exist for SFB, based on a combination of DO concentration and percent saturation objectives. The SFB Water Board staff is considering revising the Basin Plan to allow for deviation from these numeric objectives in Suisun Marsh (Howard et al. 2014) and is entertaining a similar undertaking for shallow margin and intertidal habitats in South and Lower South Bay. Once this has been established, modeling could be used to refine expectations for the deep channel habitats of South SFB. Considering that these intertidal habitats rich in organic carbon may have natural sources of low DO water, and may experience natural conditions of low DO, the expectations for DO in these habitats and their physical and biological exchanges with open water habitat need to be considered in setting appropriate expectations for the deep channel habitat.

One question that should be addressed in future iterations of the SFB AF is the need to develop a DO AF that captures a fuller gradient in condition than expressed through binary classification associated with meeting established WQOs (i.e., above or below established objectives). Best et al. (2007) have proposed a DO classification scheme for European Union Water Framework Directive (EU-WFD) based on observed impacts of hypoxia on benthic and demersal fauna, as well as expert opinion, that is targeted to be relevant in a wide range of estuarine environments (Vaquer-Sunyer and Duarte 2008). The thresholds proposed by Best et al. (2007) are similar to those calculated for California species, including those found in SFB (5.7 mg L^{-1} as chronic-effects criteria protective of 95% of the non-salmonid population and 2.8 mg L^{-1} as acute effects criteria; Sutula et al. 2012). For salmonids, Sutula et al. (2012) calculated 6.3 mg L^{-1} as chronic effects criteria and 4.0 mg L^{-1} as acute effects criteria, but notes that the effects data used to calculate these criteria were based of freshwater exposure studies. Thus, applying fixed criteria to habitats that represent a continuum along a salinity gradient can be problematic. The Best et al. (2007) thresholds have the advantage of incorporating the effects of salinity on oxygen solubility and, thus, can reconcile a threshold protective of all life history stages for salmonids from 7 mg L^{-1} in freshwater to 5.7 mg L^{-1} at marine salinities. The ASSETS upper threshold of 5.0 mg L^{-1} is roughly equivalent to this threshold but does not take into account salinity (Bricker et al. 2003). Both ASSETS and EU-WFD (Bricker et al. 1999, 2003) utilize the 5^{th} and 10^{th} percentile, respectively, to integrate over time, similar to the SFB Basin Plan calculation of 10% frequency of non-compliance. The use of the percentile approach integrates the duration and frequency of low DO events and doesn't distinguish between high frequency short duration events and low-frequency but long-duration events. The effect of these two examples can be very different on biota, depending the timing and number of reproductive cycles in the year, number per brood, etc.

Estuarine subtidal habitat and associated intertidal margin habitats are prone to development of density-driven stratification, precluding diffusion and mixing of oxygen to bottom waters (Largier et al. 1991, 1996). Sutula et al. (2012) note that natural hypoxia in bottom waters of stratified estuaries is an issue for interpretation of existing Water Boards' DO objectives. Stacey (2015, Appendix D) analyzed the frequency of stratification events in South Bay; he found that: (1) salinity-stratification most often occurs during periods of peak freshwater flow to SFB (winter-spring), (2) duration of stratification seldom persists for periods greater than two weeks due to tidal mixing associated with spring tides, and (3) observed periods of low DO in South Bay do not typically coincide with stratification events. Incursions of low DO water into SFB is possible when oceanic deep waters upwell at the mouth of SFB (J.E. Cloern, personal communication). Although these are currently rarely observed, it is possible that these events will occur with increased frequency due to rising coastal hypoxia (Booth et al. 2013).

Finally, in the first phase of AF development, we chose not to recommend a prescribed monitoring program for DO. Such recommendations were outside the scope of our current effort, yet we believe that this is an important issue – one that should be coupled to a better characterization of the seasonal DO requirements of the most sensitive species and their important habitats in SFB. Future science plans related to DO should address this important aspect.

3.4 AF Indicators as Multiple Lines of Evidence

A core principle of the AF is that it be comprised of several indicators that should be used as multiple lines of evidence in the determination of overall ecological condition. In this preliminary AF, we have chosen not to specifically address combining each indicator into a multi-metric index, pending refinement of the classification through improved monitoring, modeling and other research. However, we can offer some simple guidance on the relative weight that these indicators can be given in view of their status and relative degree of associated uncertainty. This relative importance, presented as multiple lines of evidence, can be revised as uncertainties are reduced and our understanding of risk to beneficial uses from each impairment pathway improves.

Three indicators should be given strong weight in motivating management attention the near term, given their strong linkage to beneficial uses: (1) dissolved oxygen, (2) HAB toxins, particularly if found to be accumulating to levels of concern in shellfish or other aquatic organisms, and (3) gross and net primary productivity. We note that DO already serves as an independent line of evidence, as it is already in the SFB Water Board Basin plan.

HAB abundances should be given moderate weight in motivating management action. For HAB abundances, this weight could be refined pending better characterization of HAB risk in SFB.

Chlorophyll-a should be given moderate weight in motivating nutrient management action in the short term, because of the considerable uncertainty in the linkage of chlorophyll-a with HAB toxins and DO, particularly in shallow margins with SFB. The trend in chlorophyll-a should be given as much weight as the absolute magnitude. However, given the importance of the linkage of chlorophyll-a and GPP with nutrient loads, reduction in the uncertainty surrounding chlorophyll-a classification should be a high priority in the SFB Nutrient Science Plan.

Finally, for metrics of phytoplankton composition, emphasis should be on research and data visualization to communicate the ecological significance of trends over time. We would expect that a classification system for phytoplankton food quality index should be forthcoming after a period of piloting and demonstration in SFB. However, poor phytoplankton food quality, as well as other shifts in phytoplankton composition, can be driven by factors other than nutrients. For this reason, this indicator will likely serve as a supporting rather than primary line of evidence going into the future.

4 Summary of Findings, Vision for Near-term Use, and Recommendations for AF Refinement

4.1 Summary of Findings

San Francisco Bay has long been recognized as a nutrient-enriched estuary; however, it has exhibited resistance to some of the classic symptoms of nutrient overenrichment, such as high phytoplankton biomass and hypoxia, due to a number of factors such as high turbidity, strong tidal mixing, and grazing that limit organic matter accumulation within the estuary. These observations have reinforced the need to identify numeric WQOs or a specific implementation plan for the existing narrative objective to protect the estuary from the potential effects of nutrient over-enrichment, especially following recent documentation of shifts in the timing and extent of freshwater inflow and salinity intrusion, decreasing turbidity, restructuring of plankton communities, elimination of hypoxia and reduced metal contamination of biota, and food web changes that decrease resistance of the estuary to nutrient pollution.

In this study, we utilized an expert workgroup to develop a quantitative framework to assess eutrophication in the SFB, based on indicators of phytoplankton biomass (chlorophyll-a), gross primary productivity, the prevalence of harmful algal blooms (HAB) and toxin, and DO. Experts defined core principles including geographic scope, recommended Bay segmentation, linkage of key indicators to beneficial uses, and the protocols and recommended spatial and temporal frequency of monitoring that would support a core assessment of nutrient effects on SFB.

We discussed a quantitative scheme to classify SFB subembayments in tiers of ecological condition, from very high to very low, based on risk to adverse effects of nutrient overenrichment and eutrophication. Decisions on classification bins were supported by a combination of existing literature and guidance, quantitative analyses of existing SFB data from the USGS research program, and expert best professional judgment.

Analyses of two decades of phytoplankton species composition, chlorophyll-a, and dissolved oxygen (DO), as well as three years of toxin data from solid phase adsorption toxin tracking (SPATT) samplers, were used to demonstrate (1) significant increases in chlorophyll-a, declines in DO, and a high prevalence of HAB species and toxins across most SFB subembayments, and (2) strong linkage of increasing chlorophyll-a to declining DO and HAB abundance. Statistical approaches were used to define thresholds in chlorophyll-a related to increased risks of HABs and low DO. In development of the AF classification scheme, a qualitative summary of uncertainty associated with each indicator was made for the purpose of focusing future research, monitoring, and modeling on AF refinement.

4.2 Vision for Near-Term Use of AF

The nutrient AF is intended to provide a decision framework for quantifying the extent to which SFB is supporting beneficial uses with respect to nutrients. This AF is comprised of three important elements: (1) a set of conceptual models that defines what a problem would look like in SFB, if it occurred, (2) a set of core principles supporting the AF, and (3) classification tables. The AF supports and is supported through the other major elements through:

- Defining monitoring requirements (the core indicators, spatial and temporal frequency of sampling) needed to support routine assessments of SFB
- Modeling to identify a set of management endpoints that should constitute the output of SFB water quality models and improve mechanistic understanding of the linkage of nutrients to adverse outcomes in SFB
- Informing science by identifying analyses needed to further refine the AF and highlighting areas in which monitoring, modeling and core synthesis should be improved

Given this philosophy, we feel that it is important to provide a statement of the appropriate use of the AF, given existing uncertainties.

The conceptual models and AF core principles provide a sound scientific foundation for informing modeling and monitoring. Through early interactions with the stakeholder community, these are the components of the AF that appear to have the greatest consensus and the least "uncertainty."

The classification scheme is a critical element of the AF, because it represents a quantitative and transparent mechanism through which SFB data are interpreted to assess, ultimately, nutrient-related beneficial use support. Given its importance, the authors of this document fully acknowledge the uncertainty in the AF classification scheme and need for refinement, through multiple iterations of basic research, monitoring, and modeling.

We suggest that the near-term use of the AF classification system be focused on a scientific "test drive" that seeks to understand how to collectively use and improve efficiencies for assessment, monitoring and modeling. This "test drive" should also consider whether or how to combine indicator results into multiple lines of evidence, particularly for communication to the public. Finally, this test drive should be conducted in tandem with research, monitoring and modeling to refine the AF.

4.3 Recommendations for Refinement of the AF

From this initial work, a number of recommendations emerge for refining and potentially expanding the AF. Please note that these recommendations have not been prioritized, and that early discussions to incorporate these needs into the SFB Nutrient Management Science Plan have already begun.

1. **Improve scientific basis for nutrient-related segmentation of SFB.** Our recommendation that the preliminary segmentation be based on Jassby et al. (1997) is a departure from the existing subembayments used by the SFB Water Board for assessments and permit-related activities. We strongly recommend reanalysis of existing data to be repeated using the Jassby et al. (1997) methodology, using newly available and relevant ecological data, to finalize this segmentation scheme.

2. **Include diked baylands, restored salt ponds and tidal sloughs in future iterations of this AF.** Deepwater and shallow subtidal habitats are the focus of this AF; diked baylands, restored salt ponds, and tidal sloughs are excluded in this first phase of work. We believe that these shallow water margin habitats are critical components of the SFB ecosystem and should be include in future iterations of the AF.

3. **Include dissolved oxygen classification and recommendations for monitoring in future iterations of the AF.** Current recommendations for AF focus on indicators of phytoplankton. We recommend science and synthesis to accomplish the following:

    a. Improve understanding of what species, representative of different beneficial uses, are the most sensitive to low DO and what are the temporal and spatial scales of their use of SFB subembayments as habitat
    b. Identify DO criteria representing acute and chronic tolerances to low exposure, and individual and population scales
    c. Improve characterization of the diel variability of DO at key points within the deep water and shallow margin habitat of each subembayment in order to better characterize support of species and habitats
    d. Improve mechanistic understanding of the physical and biological factors influencing DO within and between the deep channel and shallow water margin habitat

4. Optimize spatial and temporal sampling of AF indicators to best align quality of the information produced, while balancing costs, logistics, and power to detect trends. Dissolved oxygen, phytoplankton biomass, productivity and phytoplankton composition are all extremely variable across both time and space. The temporal and spatial elements of the AF and the monitoring program must be aligned and optimized to capture this variability in a manner that is also cost-effective. This could be done by conducting an intensive field observation program coupled interpolated with hydrodynamic model simulations, then conducting power analyses to understand how to best capture variability, given real constraints in available resources. Another approach is to invite subject matter experts to provide perspective about how this was done in systems of similar size and complexity (e.g. Chesapeake Bay).

5. **Reduce sources of uncertainty in chlorophyll-a, HAB abundance and toxin classification.** Three major recommendations are given to reduce uncertainty in the chlorophyll-a classification. These include:

   e. Better characterization of the significance of the ecological and human risk of HABs in SFB through more intensive monitoring of subembayments
   f. Co-location of chlorophyll-a, particulate, shellfish and SPATT monitoring to improve linkage of chlorophyll-a to HAB toxin concentrations, rather than cell counts as the foundation for the risk paradigm
   g. Expansion of SPATT samplers to include other toxins, particularly PSTs
   h. A work element to better validate SPATT toxin data relative to particulate or mussel toxin tissues: While this has historically been difficult, precedence exists (Lane et al. 2010), and because SPATT were originally designed for lipophilic toxins (Mackenzie et al. 2004), an obvious next step would also be to analyze SPATT samplers for okadaic acid, dinophysistoxins, and yessotoxins.
   i. Assembly of a scientific workgroup to synthesize scientific understanding of chronic effects of HAB toxins on SFB food webs and human health
   j. Monitoring improvements through better spatial coverage and temporal coverage of data to link chlorophyll-a to DO, focused specifically on South SFB, coupled with improved understanding of DO expectations for shallow water margins, tidal sloughs and intertidal wetland habitat (see Recommendation C above).

6. **Link HABs more specifically to nutrients.** Although deliberately excluded from this analysis, sufficient data exist to develop more complex multidimensional statistical models for harmful algal species and toxins (e.g. Kudela 2012) or to apply existing estuarine and coastal models to SFB (e.g. Lane et al. 2010; Anderson et al. 2009, 2010). This would also more directly link condition to nutrients.

7. **Fund a Nutrient Monitoring Program.** Since 1969, a USGS research program has supported water-quality sampling in SFB. This USGS program collects monthly samples between the South Bay and the lower Sacramento River to measure salinity, temperature, turbidity, suspended sediments, nutrients, dissolved oxygen and chlorophyll a. The USGS data, along with sampling conducted by the Interagency Ecological Program, provide coverage for the entire San Francisco Bay –Delta system. The San Francisco Bay Regional Monitoring Program (RMP) has no independent nutrient-related monitoring program, but instead contributes approximately 20% of the USGS data collection cost. Thus, there is currently an urgent need to lay the groundwork for a locally-supported, long-term monitoring program to provide information that is most needed to support nutrient-related management decisions in the Bay.

5 LITERATURE CITED

Andersen, Jesper H., Philip Axe, Hermanni Backer, Jacob Carstensen, Ulrich Claussen, Vivi Fleming-Lehtinen, Marko Järvinen et al., 2011. "Getting the measure of eutrophication in the Baltic Sea: towards improved assessment principles and methods." Biogeochemistry 106, no. 2 (2011): 137-156.

Anderson, Donald M., Joann M. Burkholder, William P. Cochlan, Patricia M. Glibert, Christopher J. Gobler, Cynthia A. Heil, Raphael M. Kudela et al., 2008. "Harmful algal blooms and eutrophication: examining linkages from selected coastal regions of the United States." Harmful Algae 8, no. 1 (2008): 39-53.

Anderson, Clarissa R., David A. Siegel, Raphael M. Kudela, and Mark A. Brzezinski., 2009. "Empirical models of toxigenic Pseudo-nitzschia blooms: Potential use as a remote detection tool in the Santa Barbara Channel." Harmful Algae 8, no. 3 (2009): 478-492.

Bailey, H., Curran, C., Poucher, S., Sutula, M., 2014. Science supporting dissolved oxygen objectives for Suisun Marsh. Southern California Coastal Water Research Project Authority Technical Report 830. www.sccwrp.org. 33 p.

Barber, R. T., and M. R. Hiscock., 2006. "A rising tide lifts all phytoplankton: Growth response of other phytoplankton taxa in diatom-dominated blooms." Global Biogeochemical Cycles 20, no. 4 (2006).

Baustian, M.M. and N.N. Rabalais., 2009. Seasonal composition of benthic macroinfauna exposed to hypoxia in the Northern Gulf of Mexico. Estuaries and Coasts 32:975-983.

Best, M. A., Wither, A. W., Coates, S., 2007. Dissolved oxygen as a physico-chemical supporting element in the Water Framework Directive, Mar. Pollut. Bull. 55, 53-64.

Borja, A., J. Bald, J. Franco, J. Larreta, I. Muxika, M. Revilla, J. G. Rodriguez, O. Solaun, A. Uriarte, and V. Valencia., 2009. Using multiple ecosystem components, in assessing ecological status in Spanish (Basque Country) Atlantic marine waters, Marine Pollution Bulletin, 59(1-3), 54-64.

Brett, M.T. and D.C. Muller-Navarra., 1997. The role of highly unsaturated fatty acids in aquatic foodweb processes. Freshwater Biology. 38: 483–499

Bricker, S. B., J. G. Ferreira, and T. Simas, 2003. An integrated methodology for assessment of estuarine trophic status, Ecological Modelling, 169(1), 39-60.

Bricker, S. B., C. G. Clement, D. E. Pirhalla, S. P. Orlando, and D. R. G. Farrow., 1999. National Estuarine Eutrophication Assessment: Effects of Nutrient Enrichment in the Nation's Estuaries, NOAA, National Ocean Service, Special Projects Office and National Centers for Coastal Ocean Science, Silver Springs, MD.

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

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

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

Cloern, J. E. and A. D. Jassby., 2012. Drivers of change in estuarine-coastal ecosystems: discoveries from four decades of study in San Francisco Bay. Reviews of Geophysics 50.

Cloern, J. E., T. S. Schraga, et al., 2005. "Heat wave brings an unprecedented red tide to San Francisco Bay." Eos Transactions of the American Geophysical Union 86(7): 66.

Cloern, J. E., A. D. Jassby, et al., 2007. "A cold phase of the East Pacific triggers new phytoplankton blooms in San Francisco Bay." Proceedings of the National Academy of Sciences 104(47): 18561-18565.

Cloern, J. E., Foster, S. Q., Kleckner, A. E., 2014. Phytoplankton primary production in the world’s estuarine-coastal ecosystems. Biogeosciences 11, 2477–2501. DOI: 10.5194/bg-11-2477-2014

Cole, B. E. and J. E. Cloern., 1984. An empirical model for estimating phytoplankton productivity in estuaries. Marine Ecology Progress Series 36: 299-305.

Conomos, T.J., (ed.). 1979. San Francisco Bay: the urbanized estuary. Investigation into the natural history of San Francisco Bay and Delta with reference to the influence of man. San Francisco, California: Pacific Division of the American Association for the Advancement of Science. 493 p.

Devlin, M., S. Bricker, and S. Painting., 2011. Comparison of five methods for assessing impacts of nutrient enrichment using estuarine case studies, Biogeochemistry, 106(2), 177-205.

Diaz, Robert J., and Andrew Solow., 1999. Ecological and Economic Consequences of Hypoxia: Topic 2 Report for the Integrated Assessment on Hypoxia in the Gulf of Mexico. NOAA Coastal Ocean Program Decision Analysis Series No. 16. NOAA Coastal Ocean Program, Silver Spring, MD. 45 pp.

Dugdale, R., F. Wilkerson, et al., 2012. "River flow and ammonium discharge determine spring phytoplankton blooms in an urbanized estuary." Estuarine Coastal and Shelf Science 115: 187-199.

Dugdale, R. C., F. P. Wilkerson, et al., 2007. "The role of ammonium and nitrate in spring bloom development in San Francisco Bay." Estuarine, Coastal and Shelf Science 73: 17-29.

Eppley, R. W., Carlucci, A. F., Holm-Hansen, O., Kiefer, D., McCarthy, J. J., Venrick, E., Williams, P. M., 1971. Phytoplankton growth and composition in shipboard cultures supplied with nitrate, ammonium, or urea as the nitrogen source. Limnol. Oceanogr. 16, 741-751.

Ferreira, J.G., J.H. Andersen, A. Borja, S.B. Bricker, J. Camp, M. Cardoso da Silva, E. Garcés, A.-S. Heiskanen, C. Humborg, L. Ignatiades, C. Lancelot, A. Menesguen, P. Tett, N. Hoepffner, and U. Claussen., 2011. Overview of eutrophication indicators to assess environmental status within the European Marine Strategy Framework Directive. Estuarine, Coastal and Shelf Science 93: 117–131.

Galloway AWE, Winder M., 2015. Partitioning the Relative Importance of Phylogeny and Environmental Conditions on Phytoplankton Fatty Acids. PLoS ONE 10(6): e0130053. doi: 10.1371/journal.pone.0130053

Gallegos, C. L. and P. J. Neale., 2015. "Long-term variations in primary production in a eutrophic sub-estuary: Contribution of short-term events." Estuarine, Coastal and Shelf Science 162: 22-34.

Glibert, P. M., 2010. "Long-Term Changes in Nutrient Loading and Stoichiometry and Their Relationships with Changes in the Food Web and Dominant Pelagic Fish Species in the San Francisco Estuary, California." Reviews in Fisheries Science 18(2): 211-232.

Glibert, P.M., D.M. Anderson, P. Gentien, E. Graneli, and K.G. Sellner., 2005. The global, complex phenomena of harmful algal blooms. Oceanography 18(2):137–147.

Goldstein, T., J. A. K. Mazet, T. S. Zabka, G. Langlois, K. M. Colegrove, M. Silver, S. Bargu et al., 2008. "Novel symptomatology and changing epidemiology of domoic acid toxicosis in California sea lions Zalophus californianus: an increasing risk to marine mammal health." Proceedings of the Royal Society B: Biological Sciences 275, no. 1632 2008: 267-276.

Harding, L. W., R. A. Batiuk, T. R. Fisher, C. L. Gallegos, T. C. Malone, W. D. Miller, M. R. Mulholland, H. W. Paerl, E. S. Perry, and P. Tango., 2013. Scientific bases for numerical chlorophyll criteria in Chesapeake Bay. Estuaries and Coasts DOI 10.1007/s12237-013-9656-6

Heisler, John, Patricia M. Glibert, JoAnn M. Burkholder, Donald M. Anderson, William Cochlan, William C. Dennison, Quay Dortch et al., 2008. "Eutrophication and harmful algal blooms: a scientific consensus." Harmful algae 8, no. 1 2008: 3-13.

Hiolski EM, Kendrick PS, Frame ER, Myers MS, Bammler TK, Beyer RP, Farin FM, Wilkerson HW, Smith DR, Marcinek DJ, Lefebvre KA., 2014. Chronic low-level domoic acid exposure alters gene transcription and impairs mitochondrial function in the CNS. Aquat Toxicol., 2014, 155: 151-159.

Howard, Meredith D. Armstrong, William P. Cochlan, Nicolas Ladizinsky, and Raphael M. Kudela., 2007. "Nitrogenous preference of toxigenic Pseudo-nitzschia australis (Bacillariophyceae) from field and laboratory experiments." Harmful Algae 6, no. 2 (2007): 206-217.

Jassby, A.D., Cole, B.E., Cloern, J.E., 1997. The design of sampling transects for characterizing water quality in estuaries. Estuarine, Coastal, and Shelf Science 45, 285–302.

Kimmerer, W. J., J. K. Thompson., 2014. "Phytoplankton Growth Balanced by Clam and Zooplankton Grazing and Net Transport into the Low-Salinity Zone of the San Francisco Estuary" Estuaries and Coasts 37 (5): 1202-1218

Kudela, Raphael M., 2011. "Characterization and deployment of Solid Phase Adsorption Toxin Tracking (SPATT) resin for monitoring of microcystins in fresh and saltwater." Harmful Algae 11 (2011): 117-125.

Lane, J.Q., Roddam, C.M., Langlois, G.W., Kudela, R.M., 2010. Application of Solid Phase Adsorption Toxin Tracking (SPATT) for field detection of domoic acid and saxitoxin in coastal California. Limnol. Oceanogr. Methods 8, 645–660.

Lehman, P.W. and S. Waller., 2003. Microcystis blooms in the delta. Interagency Ecological Program for the San Francisco Estuary Newsletter. 16, 18-19. www.water.ca.gov/iep/products/newsletter.cfm

Lehman, P. W., G. Boyer, et al., 2005. "Distribution and toxicity of a new colonial Microcystis aeruginosa bloom in the San Francisco Bay Estuary, California." Hydrobiologia 541: 87-99.

Lehman, P. W., G. Boyer, et al., 2008. "The influence of environmental conditions on the seasonal variation of Microsystis cell density and microcystins concentration in San Francisco Estuary." Hydrobiologia 600: 187-204.

Lehman, P. W., K. Marr, G. L. Boyer, S. Acuna, and S. J. Teh., 2013. "Long-term trends and causal factors associated with Microcystis abundance and toxicity in San Francisco Estuary and implications for climate change impacts." Hydrobiologia 718, no. 1: 141-158.

Lewitus, Alan J; Horner, Rita A; Caron, David A; Garcia-Mendoza, Ernesto; Hickey, Barbara M; Hunter, Matthew; Huppert, Daniel D; Kudela, Raphael M; Langlois, Gregg W; Largier, John L., 2012. Harmful algal blooms along the North American west coast region: History, trends, causes, and impacts. Harmful Algae. 19: 133-159.

Mackenzie, L., V. Beuzenberg, P. Holland, P. Mcnabb, and A. Selwood., 2004. Solid phase adsorption toxin tracking (SPATT): a new monitoring tool that simulates the biotoxin contamination of filter feeding bivalves. Toxicon 44:901-918 [doi:10.1016/j.toxicon.2004.08.020].

K McLaughlin, M Sutula, L Busse, S Anderson, J Crooks, R Dagit, D Gibson, K Johnston, L Stratton., 2013. A regional survey of the extent and magnitude of eutrophication in Mediterranean estuaries of Southern California, USA. Estuaries and Coasts DOI 10.1007/s12237-013-9670-8.

Ning, X., Cloern, J. E. and Cole, B. E., 2000. Spatial and temporal variability of picocyanobacteria Synechococcus sp. in San Francisco Bay. Limnol. Oceanogr., 45, 695–702.

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

NRC. ,2012. Sustainable Water and Environmental Management in the California Bay-Delta. Committee on Sustainable Water and Environmental Management in the California Bay-Delta; Water Science and Technology Board; Ocean Studies Board; Division on Earth and Life Studies; National Research Council. Available at: http://www.nap.edu/catalog.php?record_id=13394.

Null, K.A., Dimova, N.T, et al., 2012. "Submarine Groundwater Discharge-Derived Nutrient Loads to San Francisco Bay: Implications to Future Ecosystem Changes". Estuaries and Coasts. 35: 1299-1315

OEHHA., 2012. Office of Environmental Health Hazard Assessment. Toxicological summary and suggested action levels to reduce potential adverse health effects of six cyanotoxins. May 2012.

Park S., M. Brett, E.T. Oshel, and C.R. Goldman., 2003. Seston food quality and Daphnia production efficiencies in an oligo-mesotrophic Subalpine Lake. Aquatic Ecology 37: 123–136.

Parker, A. E., R. C. Dugdale, et al., 2012a. "Elevated ammonium concentrations from wastewater discharge depress primary productivity in the Sacramento River and the Northern San Francisco Estuary." Marine Pollution Bulletin 64(3): 574-586.

Parker, A. E., V. E. Hogue, et al., 2012b. "The effect of inorganic nitrogen speciation on primary production in the San Francisco Estuary." Estuarine Coastal and Shelf Science 104: 91-101.

Riemann, B., Carstensen, J., Dahl, K., Fossing, H., Hansen, J. W., Jakobsen, H. H., Josefson, A. A., Krause-Jensen, D., Markager, S., Stæhr, P.A., Timmermann, K., Windolf, J., Andersen, J. H., 2015. Recovery of Danish coastal ecosystems after reductions in nutrient loading: A holistic ecosystem approach. Estuar. Coasts doi: 10.1007/s12237-015-9980-0.

Schaeffer, Blake A., James D. Hagy, Robyn N. Conmy, John C. Lehrter, and Richard P. Stumpf., 2012. "An approach to developing numeric water quality criteria for coastal waters using the SeaWiFS satellite data record." Environmental science & technology 46, no. 2 (2012): 916-922.

SFEI 2014a. Synthesis of existing dissolved oxygen data in Lower South Bay.

SFEI 2014b. Conceptual model of nutrient impairment in San Francisco Bay.

SFEI, 2014c. Development Plan for the San Francisco Bay Nutrient Monitoring Program.

SFRWQCB, 2012. San Francisco Bay Regional Water Quality Control Board. San Francisco Bay Nutrient Management Strategy 2012

SFRWQCB, 1975. San Francisco Bay Basin (Region 2) Water Quality Control Plan

SFRWQCB, 2015. San Francisco Bay Basin (Region 2) Water Quality Control Plan, available at www.waterboards.ca.gov/rwqcb2/basin_planning.shtml.

Shellenbarger, G. G., Schoellhamer, D.H., Morgan, T.L., Takekawa, J.Y., Athearn, N.D., and Henderson, K.D., 2008. "Dissolved oxygen in Guadalupe Slough and Pond A3W, South San Francisco Bay, California, August and September 2007." U.S. Geological Survey Open-File Report 2008–1097, 26 p.

Smith S.V., Hollibaugh, J. T., 2006. Water, salt, and nutrient exchanges in San Francisco Bay. Limnol. Oceanogr. 51, 504–517.

Sutula, M., 2011. Review of indicators for development of nutrient numeric endpoints in California estuaries. Technical Report 646. Southern California Coastal Water Research Project. Costa Mesa, CA.

SWRCB, 2014. State Water Resources Control Board. Proposed Workplan for Development of a Nutrient Control Program. www.waterboards.ca.gov/plans_policies/nutrients.shtml

Tetra Tech, 2006. Technical Approach to Develop Nutrient Numeric Endpoints for California. Tetra Tech, Inc. http://rd.tetratech.com/epa/Documents/CA_NNE_July_Final.pdf.

Tett, Paul, Richard Gowen, Dave Mills, Teresa Fernandes, Linda Gilpin, Mark Huxham, Kevin Kennington et al., 2007. "Defining and detecting undesirable disturbance in the context of marine eutrophication." Marine Pollution Bulletin 55, no. 1 2007: 282-297.

Thebault, J., T. S. Schraga, et al., 2008. "Primary production and carrying capacity of former salt ponds after reconnection to San Francisco Bay." Wetlands 28(3): 841-851.

Topping, B. R., Kuwabara, J.S., Athearn, N.D., Takekawa, J.Y., Parchaso, F., Henderson, K.D., and Piotter, S., 2009. Benthic oxygen demand in three former salt ponds adjacent to south San Francisco Bay, California. U.S. Geological Survey Open-File Report 2009-1180, 21 p.

USEPA, 2001. Nutrient Criteria Technical Guidance Manual: Estuarine, Coastal and Marine Waters. Office of Water, U.S. Environmental Protection Agency. EPA 822-B-01-003.

Vlamis A. and P. Katikou., 2014. Climate influence on Dinophysis spp. spatial and temporal distributions in Greek coastal water. Plankton Benthos Res 9(1): 15–31.

Zaldivar, J.-M., Ana Cristina Cardoso, Pierluigi Viaroli, Alice Newton, Rutger de Wit, Carles Ibañez, Sofia Reizopoulou, Francesca Somma, Arturas Razinkovas, Alberto Basset, Marianne Holmer, and Nicholas Murray., 2008, Eutrophication in transitional waters: an overview, Transitional Waters Monographs, 1, 1-78.

APPENDIX A. DEFINITIONS OF KEY TERMS AND SFB BENEFICIAL USES

APPENDIX B REVIEW OF APPROACHES TO ASSESSMENT OF NUTRIENT EFFECTS ON ESTUARIES

APPENDIX C QUANTITATIVE ANALYSES SUPPORTING DECISIONS ON CHLOROPHYLL-A ASSESSMENT ENDPOINTS (SUTULA ET AL. MANUSCRIPT IN PREP FOR SUBMISSION TO A SCIENTIFIC JOURNAL)

APPENDIX D. SUPPLEMENTAL ANALYSES SUPPORTING DISCUSSION OF THE IMPORTANCE OF STRATIFICATION ON THE RELATIONSHIP BETWEEN DISSOLVED OXYGEN AND CHLOROPHYLL-A IN SF BAY (STACEY AND SENN, 2015 TECHNICAL MEMO)

1.1 IMPORTANT DEFINITIONS

For those outside the regulatory world, the distinction between terms like "criteria," "standards", "objectives," and "endpoints" can be confusing. The purpose of this section is to provide definitions of the terms that are linked closely to how the assessment framework could be in used water quality regulation.

Eutrophication: Eutrophication is defined as the acceleration of the delivery, in situ production of organic matter, and accumulation of organic matter (Nixon 1995). One main cause of eutrophication in estuaries is nutrient over-enrichment (nitrogen, phosphorus and silica). However, other factors influence primary producer growth and the build-up of nutrient concentrations, and hence modify (or buffer) the response of a system to increased nutrient loads (hereto referred to as **co-factors**). These co-factors can include hydrologic residence times, mixing characteristics, water temperature, light climate, grazing pressure and, in some cases, coastal upwelling.

***Indicator*:** A characteristic of an ecosystem that is related to, or derived from, a measure of biotic or abiotic variable, that can provide quantitative information on ecological condition, structure and/or function. With respect to the water quality objectives, indicators are the ecological parameters for which narrative or numeric objectives are developed.

***Water Quality Standard*s:** Water quality standards are the foundation of the water quality-based control program mandated by the Clean Water Act. Water Quality Standards define the goals for a waterbody by designating its uses, setting criteria to protect those uses, and establishing provisions to protect water quality from pollutants. A water quality standard consists of three basic elements:

- Designated uses of the water body (e.g., recreation, water supply, aquatic life, agriculture; Table 1.1),
- Water quality criteria to protect designated uses (numeric pollutant concentrations and narrative requirements), and
- Antidegradation policy to maintain and protect existing uses and high quality waters.

Water Quality Criteria: Section 303 of the Clean Water Act gives the States and authorized Tribes power to adopt water quality criteria with sufficient coverage of parameters and of adequate stringency to protect designated uses. In adopting criteria, States and Tribes may:

- Adopt the criteria that US EPA publishes under §304(a) of the Clean Water Act;
- Modify the §304(a) criteria to reflect site-specific conditions; or
- Adopt criteria based on other scientifically-defensible methods.

The State of California's water criteria are implemented as "water quality objectives," as defined in the Water Code (of the Porter Cologne Act; for further explanation, see below).
States and Tribes typically adopt both **numeric** and **narrative** criteria. **Numeric** criteria are quantitative. **Narrative** criteria lack specific numeric targets but define a targeted condition that must be achieved.

Section 303(c)(2)(B) of the Clean Water Act requires States and authorized Tribes to adopt numeric criteria for priority toxic pollutants for which the Agency has published §304(a) criteria. In addition to narrative and numeric (chemical-specific) criteria, other types of water quality criteria include biological, nutrient and sediment criteria.

Water Quality Objectives: The Water Code (Porter-Cologne Act) provides that each Regional Water Quality Control Board shall establish water quality objectives for the waters of the state i.e., (ground and surface waters) which, in the Regional Board's judgment, are necessary for the reasonable protection of beneficial uses and for the prevention of nuisance. The State of California typically adopts both **numeric** and **narrative** objectives. **Numeric** objectives are quantitative. **Narrative** objectives present general descriptions of water quality that must be attained through pollutant control measures. Narrative objectives are also often a basis for the development of numerical objectives.

Numeric Endpoint: Within the context of the ecological risk assessment framework, numeric endpoints are thresholds that define the magnitude of an indicator that is considered protective of ecological health. These numeric endpoints serve as guidance to Regional Boards in translating narrative nutrient or biostimulatory substance water quality objectives. They are called “numeric endpoints” rather than “numeric objectives” to distinguish the difference with respect to State and Regional Water Board policy. Objectives are promulgated through a public process and incorporated into basin plans. Numeric endpoints are guidance that presumably can evolve over time without the need to go through a formal standards development process.

1.2 BENEFICIAL USE DEFINITIONS

TABLE A1.1 DEFINITION OF ESTUARINE BENEFICIAL USES APPLICABLE TO SELECTION OF NUTRIENT ASSESSMENT FRAMEWORK ENDPOINTS IN SF BAY.

Marine Habitat (MAR) - Uses of water that support marine ecosystems including, but not limited to, preservation or enhancement of marine habitats, vegetation such as kelp, fish, shellfish, or wildlife (e.g., marine mammals, shorebirds). **Estuarine Habitat (EST) -** Uses of water that support estuarine ecosystems including, but not limited to, preservation or enhancement of estuarine habitats, vegetation, fish, shellfish, or wildlife (e.g., estuarine mammals, waterfowl, shorebirds) and the propagation, sustenance and migration of estuarine organisms. **Cold Freshwater Habitat (COLD) -** Uses of water that support cold water ecosystems including, but not limited to, preservation or enhancement of aquatic habitats, vegetation, fish or wildlife, including invertebrates. **Warm Freshwater Habitat (WARM) -** Uses of water that support warm water ecosystems including, but not limited to, preservation or enhancement of aquatic habitats, vegetation, fish or wildlife, including invertebrates. **Wildlife Habitat (WILD**) **-** Uses of water that support wildlife habitats including, but not limited to, preservation and enhancement of vegetation and prey species used by wildlife, such as waterfowl. **Rare, Threatened, or Endangered Species (RARE) -** Uses of water that support habitats necessary for the survival and successful maintenance of plant or animal species established under state or federal law as rare, threatened or endangered. **Spawning, Reproduction, and/or Early Development (SPWN**) **-** Uses of water that support high quality aquatic habitats suitable for reproduction and early development of fish. **Migration of Aquatic Organisms (MIGR**) **-** Uses of water that support habitats necessary for migration, acclimatization between fresh and salt water, and protection of aquatic organisms that are temporary inhabitants of water in the region. **Commercial and Sport Fishing (COMM) -** Uses of water for commercial or recreational collection of fish, shellfish, or other organisms including, but not limited to, uses involving organisms intended for human consumption or bait purposes. **Shellfish Harvesting (SHELL**) **-** Uses of water that support habitats suitable for the collection of crustaceans and filter-feeding shellfish (e.g., clams, oysters and mussels) for human consumption, commercial, or sport purposes. **Contact Water Recreation (REC-1) -** Uses of water for recreational activities involving body contact with water, where ingestion of water is reasonably possible. These uses include, but are not limited to, swimming, wading, water-skiing, skin and SCUBA diving, surfing, white water activities, fishing, or use of natural hot springs. **Non-contact Water Recreation (REC-2)** – Uses of water for recreational activities involving proximity to water, but not normally involving body contact with water, where ingestion of water is reasonably possible. These uses include, but are not limited to, picnicking, sunbathing, hiking, beachcombing, camping, boating, tidepool and marine life study, hunting, sightseeing, or aesthetic enjoyment in conjunction with the above activities.

A Review of Scientific Approaches Supporting NNE Assessment Framework Development for San Francisco Bay

May 2013 Version

Prepared for:
San Francisco Regional Water Quality Control Board
Basin Planning and TMDL Unit

Prepared by:

Martha Sutula and Karen McLaughlin
So. California Coastal Water Research Project
3535 Harbor Blvd, Suite 110
Costa Mesa CA 92626

Suzanne Bricker
NOAA National Centers for Coastal Ocean Science
1305 East West Highway, Rm 8110
Silver Spring, MD 20910

David Senn and Emily Novick
San Francisco Estuary Institute
4911 Central Ave, Richmond, CA 94804

Table of Contents

LIST OF FIGURES

LIST OF TABLES

1 INTRODUCTION

1.1 BACKGROUND AND PURPOSE

The California State Water Resources Control Board (State Water Board) is developing nutrient water quality objectives for the State's surface waters, using an approach known as Nutrient Numeric Endpoints (NNE). The NNE is comprised of two components. First, it would establish a suite of numeric regulatory endpoints based on the ecological response of an aquatic waterbody to nutrient over-enrichment (eutrophication, e.g., algal biomass, dissolved oxygen). Second, nutrient-response models would be used to link the ecological response endpoints to site-specific nutrient targets and other potential management controls. The NNE, intended to serve as numeric guidance to translate narrative water quality objectives, is currently under development for all California estuaries (Sutula 2013).

San Francisco Bay represents California's largest estuary (70% by area of estuarine habitat statewide). Because of its size and complexity, State Water Board staff determined that it merits development of site-specific nutrient objectives. The State Water Board and the San Francisco (SF) Water Board have agreed to collaborate on the development of site-specific nutrient objectives for SF Bay and that the SF Water Board will lead on this effort. In 2012, the SF Water Board and its stakeholders jointly developed a strategy to development regulatory endpoints and nutrient-response model for San Francisco Bay.

The process to select NNE regulatory endpoints begins with synthesis of science and ends with policy decisions. In this document, we refer to the product of scientific synthesis as an "**NNE assessment framework**," defined as a structured set of decision rules that specify how to use monitoring data to categorize specific segments of SF Bay with respect to adverse effects on Bay beneficial uses due to nutrient-overenrichment. While the decision on regulatory endpoints should be informed by science, it is ultimately a policy decision. The intention is that the SF Water Board would propose regulatory endpoints for SF Bay, based on the synthesis of science represented in the NNE assessment framework and feedback from the SF Bay stakeholders.

The purpose of this document is to review approaches to developing an NNE assessment framework, based on existing work in the United States and other countries. This document would summarize existing literature for how those indicators have been used to assess ecological condition and recommend a suite of options to consider for further exploration. The intent is that this white paper would be used to initiate discussions via a kick-off meeting with a working group of experts in estuarine eutrophication to: 1) discuss possible approaches and 2) identify the types of analyses of existing data that would support their evaluation. The white paper would also be discussed with SF Bay stakeholders for feedback and comments on approaches as well as identification of additional data sources that could support the evaluation.

Conceptually, the assessment framework builds on work by McKee et al (2011), which reviewed candidate indicators indicative of eutrophication or other adverse effects to Bay beneficial uses, assessed status and trends in these indicators, identified data gaps and recommended next steps. This review served as a starting point for the development of a nutrient management program for San Francisco Bay, spearheaded by the San Francisco RWQCB. Since the publication of the McKee et al. (2011) report, this program has produced an overarching strategy or work plan to guide technical, outreach and policy elements (SFRWQCB 2012) and several technical work products related to addressing data gaps or building on recommendations in the McKee et al. (2011) report (e.g. Senn et al. 2013).

The review recommended developing regulatory endpoints for subtidal habitat based on indicators such as phytoplankton, nutrient concentrations, and dissolved oxygen. Work to review the science supporting dissolved oxygen objectives will be completed separately from this effort; thus assessment framework development will focus on indicators and metrics of phytoplankton and nutrient concentrations. A particular approach to developing this framework is not presumed at the outset; rather the intent is to select the appropriate approach with advice of experts and stakeholders as a part of the process. The assessment framework will also build on recent work, led by SFEI, to develop conceptual models of SF Bay ecological response to nutrient loads and linkage to Bay beneficial use (Senn et al. 2013).

1.2 Development of a Nutrient Assessment Framework for SF Bay: Process and Desirable Attributes

Process

To understand the context for this white paper, it is helpful to understand the process envisioned to develop the SF Bay Nutrient Assessment Framework. We envision this process to consists of 5 steps:

1. Review existing approaches to nutrient assessment framework development
2. Analyze existing data to test applicable approaches
3. Draft assessment framework
4. Test with existing or newly collected monitoring data
5. Refine assessment framework

Philosophically, each step requires the review and input of the stakeholder advisory group.

Review Existing Approaches. The first step in developing an assessment framework is to prepare a white paper summarizing potential approaches that have been used elsewhere in the United State or in other countries. This white paper will identify candidate indicators and metrics, summarize existing literature for how those indicators have been used to assess ecological condition and recommend a suite of options to consider for further exploration. This white paper would also be discussed with SF Bay stakeholders for feedback and comments on approaches as well as identification of additional data sources that could support the evaluation. It will be used to initiate discussions via a kick-off meeting with a working group of experts in to: 1) discuss possible approaches and 2) identify the types of analyses of existing data that would support their evaluation.

Analyses of Existing Data. The next step is to analyze existing data from SF Bay estuary that would support the evaluation of possible approaches to nutrient assessment framework development. Analyses will focus on identifying how data on indicators or combinations of indicators can be used to identify alternative states and how decisions on data aggregation across temporal and spatial scales affects the results of the assessment.

Draft Assessment Framework. Results of the analysis of existing data will be used by the expert working group to draft an nutrient assessment framework for SF Bay. Workgroup participants will to develop the scientific foundation for the assessment framework, specifying to the degree possible: 1) indicators and specific metrics, 2) a number of categories representing "alternative states" from high to low ecological condition and/or beneficial use support and 3) decision rules for how data should be used to categorize the Bay or Bay segment being to the applicable "alternative state."

Test Assessment Framework With Monitoring Data and Refine (As Needed) Assessment Framework . The draft assessment framework will be tested with monitoring data, either existing or newly collected. This effort will be used as an opportunity to make any refinements to the assessment framework. Results of the assessment will be compiled into a Bay “report card” and communicated to the public.

Desirable Attributes of An Assessment Framework

Desirable attributes of an nutrient assessment framework for SF Bay are as follows:

- The assessment framework should employ indicator(s) that have a strong linkage to Bay beneficial uses. This linkage should be scientifically well supported and easily communicable to the public.
- One or more primary indicators of the assessment framework should have a predictive relationship with surface water nutrients and/or nutrient loads to the Bay.
- The assessment framework should employ the indicator(s) classify the Bay segments from very high ecological condition to very low ecological condition. It should be explicit how the magnitude, extent, and duration of the effects that cause the segment to be classified differently.
- The assessment framework should be spatially explicit for different segments of the Bay and different habitat types (deep versus shallow subtidal) as warranted by the ecological nature of response to nutrients.

- The assessment framework should specify what are the appropriate methods used to measure the indicator and the temporal and spatial density of data required to make that assessment.
- It should provide guidance on how the data should be analyzed to categorize the Bay segments.

1.3 IMPORTANT DEFINITIONS

For those outside the regulatory world, distinction between terms like “criteria,” “standards”, “objectives,” and “endpoints” can be confusing. The purpose of this section is to provide definitions of the terms that are linked closely to how the NNE framework will be implemented.

Eutrophication: Eutrophication is defined as the acceleration of the delivery, in situ production of organic matter, and accumulation of organic matter (Nixon 1995). One main cause of eutrophication in estuaries is nutrient over enrichment (nitrogen, phosphorus and silica). However, other factors influence primary producer growth and the build-up of nutrient concentrations, and hence modify (or buffer) the response of a system to increased nutrient loads (hereto referred to as **co-factors**). These co-factors include hydrologic residence times, mixing characteristics, water temperature, light climate, grazing pressure and, in some cases, coastal upwelling.

***Indicator*:** A characteristic of an ecosystem that is related to, or derived from, a measure of biotic or abiotic variable, that can provide quantitative information on ecological condition, structure and/or function. With respect to the water quality objectives, indicators are the ecological parameters for which narrative or numeric objectives are developed.

***Water Quality Standard*s:** Water quality standards are the foundation of the water quality-based control program mandated by the Clean Water Act. Water Quality Standards define the goals for a waterbody by designating its uses, setting criteria to protect those uses, and establishing provisions to protect water quality from pollutants. A water quality standard consists of three basic elements:

- Designated uses of the water body (e.g., recreation, water supply, aquatic life, agriculture; Table 1.1),
- Water quality criteria to protect designated uses (numeric pollutant concentrations and narrative requirements), and
- Antidegradation policy to maintain and protect existing uses and high quality waters.

Water Quality Criteria: Section 303 of the Clean Water Act gives the States and authorized Tribes power to adopt water quality criteria with sufficient coverage of parameters and of adequate stringency to protect designated uses. In adopting criteria, States and Tribes may:

- Adopt the criteria that US EPA publishes under §304(a) of the Clean Water Act;
- Modify the §304(a) criteria to reflect site-specific conditions; or

- Adopt criteria based on other scientifically-defensible methods.

The State of California's water criteria are implemented as "water quality objectives," as defined in the Water Code (of the Porter Cologne Act; for further explanation, see below). States and Tribes typically adopt both **numeric** and **narrative** criteria. **Numeric** criteria are quantitative. **Narrative** criteria lack specific numeric targets but define a targeted condition that must be achieved.

Section 303(c)(2)(B) of the Clean Water Act requires States and authorized Tribes to adopt numeric criteria for priority toxic pollutants for which the Agency has published §304(a) criteria. In addition to narrative and numeric (chemical-specific) criteria, other types of water quality criteria include:

- Biological criteria: a description of the desired biological condition of the aquatic community, for example, based on the numbers and kinds of organisms expected to be present in a water body.
- Nutrient criteria: a means to protect against nutrient over-enrichment and cultural eutrophication.
- Sediment criteria: a description of conditions that will avoid adverse effects of contaminated and uncontaminated sediments.

Water Quality Objectives: The Water Code (Porter-Cologne Act) provides that each Regional Water Quality Control Board shall establish water quality objectives for the waters of the state i.e., (ground and surface waters) which, in the Regional Board's judgment, are necessary for the reasonable protection of beneficial uses and for the prevention of nuisance. The State of California typically adopts both **numeric** and **narrative** objectives. **Numeric** objectives are quantitative. **Narrative** objectives present general descriptions of water quality that must be attained through pollutant control measures. Narrative objectives are also often a basis for the development of numerical objectives.

Numeric Endpoint: Within the context of the NNE framework, numeric endpoints are thresholds that define the magnitude of an indicator that is considered protective of ecological health. These numeric endpoints serve as guidance to Regional Boards in translating narrative nutrient or biostimulatory substance water quality objectives. They are called "numeric endpoints" rather than "numeric objectives" to distinguish the difference with respect to SWRCB policy. Objectives are promulgated through a public process and incorporated into basin plans. Numeric endpoints are guidance that presumably can evolve over time without the need to go through a formal standards development process.

Table 1.1. Definition of estuarine beneficial uses applicable to selection of E-NNE indicators.

Marine Habitat (MAR) - Uses of water that support marine ecosystems including, but not limited to, preservation or enhancement of marine habitats, vegetation such as kelp, fish, shellfish, or wildlife (e.g., marine mammals, shorebirds).

Estuarine Habitat (EST) - Uses of water that support estuarine ecosystems including, but not limited to, preservation or enhancement of estuarine habitats, vegetation, fish, shellfish, or wildlife (e.g., estuarine mammals, waterfowl, shorebirds).

Cold Freshwater Habitat (COLD) - Uses of water that support cold water ecosystems including, but not limited to, preservation or enhancement of aquatic habitats, vegetation, fish or wildlife, including invertebrates.

Warm Freshwater Habitat (WARM) - Uses of water that support warm water ecosystems including, but not limited to, preservation or enhancement of aquatic habitats, vegetation, fish or wildlife, including invertebrates.

Wildlife Habitat (WILD) - Uses of water that support terrestrial ecosystems including, but not limited to, preservation and enhancement of terrestrial habitats, vegetation, wildlife (e.g., mammals, birds, reptiles, amphibians, invertebrates), or wildlife water and food sources.

Rare, Threatened, or Endangered Species (RARE) - Uses of water that support habitats necessary, at least in part, for the survival and successful maintenance of plant or animal species established under state or federal law as rare, threatened or endangered.

Spawning, Reproduction, and/or Early Development (SPWN) - Uses of water that support high quality aquatic habitats suitable for reproduction and early development of fish. This use is applicable only for the protection of anadromous fish.

Migration of Aquatic Organisms (MIGR) - Uses of water that support habitats necessary for migration, acclimatization between fresh and salt water, or other temporary activities by aquatic organisms, such as anadromous fish

Commercial and Sport Fishing (COMM) - Uses of water for commercial or recreational collection of fish, shellfish, or other organisms including, but not limited to, uses involving organisms intended for human consumption or bait purposes.

Shellfish Harvesting (SHELL) - Uses of water that support habitats suitable for the collection of filter-feeding shellfish (e.g., clams, oysters and mussels) for human consumption, commercial, or sport purposes.

Aquaculture (AQUA) - Uses of water for aquaculture or mariculture operations including, but not limited to, propagation, cultivation, maintenance, or harvesting of aquatic plants and animals for human consumption or bait purposes.

Contact Water Recreation (REC-1) - Uses of water for recreational activities involving body contact with water, where ingestion of water is reasonably possible. These uses include, but are not limited to, swimming, wading, water-skiing, skin and SCUBA diving, surfing, white water activities, fishing, or use of natural hot springs.

Non-contact Water Recreation (REC-2) – Uses of water for recreational activities involving proximity to water, but not normally involving body contact with water, where ingestion of water is reasonably possible. These uses include, but are not limited to, picnicking, sunbathing, hiking, beachcombing, camping, boating, tidepool and marine life study, hunting, sightseeing, or aesthetic enjoyment in conjunction with the above activities.

2 Development of Nutrient Numeric Endpoints (NNE) Framework and Nutrient-Response Models in San Francisco Bay: Basic Concepts

2.1 Background for Development of NNEs in Estuaries

U.S. EPA initiated the National Nutrient Management Strategy in 1998 to begin addressing the pervasive impacts of excessive nutrient loading to both fresh and marine waters (Wayland 1998). A primary objective of the strategy was to develop numeric nutrient criteria to measure the progress of the management strategy. EPA issued a series of technical guidance manuals for the development of nutrient criteria.

The "Nutrient Criteria Technical guidance Manual: Estuarine and Coastal Waters" was released by EPA in October 2001. EPA Region IX had already convened the Regional Technical Advisory Group (RTAG) and the State Technical Advisory Group (STRTAG) to serve as a forum for collaboration among stakeholders, agencies, and all nine Regional Water Boards. RTAG and STRTAG focused on the development of nutrient criteria for fresh waters. In 2006 the STRTAG proposed the California Nutrient Numeric Endpoint framework as California's approach to nutrient objectives. The development of nutrient numeric endpoints for fresh waters is preceeding prior to estuaries with the caveat that endpoints for upstream waterbodies would consider potential downstream impacts on estuaries.

Sutula et al. (2007) developed a conceptual framework for development of NNEs in estuaries based on the framework for streams (USEPA 2006). A work plan governing NNE development in estuaries was funded (McLauglin et al. 2009). Results of initial funding and an the work plan to continue NNE development has recently been updated (Sutula 2013). The work plan specifically identifies efforts by the San Francisco RWQCB and the Central Valley RWQCB to establish "site-specific" nutrient objectives for the San Francisco Bay (SFRWQCB 2012) and Delta.

2.2 Approaches to Setting Nutrient Objectives

Nutrient objectives are scientifically challenging. Nutrients are required to support life, but assessment of how much is "too much" is not straightforward. Typical paradigms used to set thresholds for toxic contaminants do not apply, in part because adverse effects of nutrient over enrichment are visible at orders of magnitude below recognized toxicity thresholds for ammonium and nitrate.

US EPA guidance on nutrient objective development generally recommends three means to set nutrient criteria (USEPA 2001): 1) reference approach, 2) empirical stress-response approach, and 3) cause-effect approach. The reference waterbody approach involves characterization of the distributions of nutrient in "minimally disturbed" waterbodies. Nutrient concentrations are chosen at some statistical percentile of those reference waterbodies. The empirical stress-response approach involves establishing statistical relationships between the causal or stressor (in this case nutrient concentrations or loads) and the ecological response (changes in algal or aquatic plant biomass or community structure, changes in sediment or water chemistry (e.g., dissolved oxygen, pH). The cause-effect approach involves identifying the ecological responses of concern and mechanistically modeling the linkage back to nutrient loads and other co-factors controlling response (e.g., hydrology, grazers, denitrification, etc.).

SWRCB staff and USEPA Region 9 staff evaluated these three approaches for setting nutrient objectives in California waterbodies and determined that, while it may choose to ultimately incorporate some elements of all approaches into California's strategy for setting nutrient objectives, it would rely most heavily on the cause-effect approach. There were several reasons for this. First, the cause-effect approach has a more direct linkage with beneficial uses and is generally thought to lend itself to a more precise diagnosis of adverse effects. Second, the alternative approaches require a tremendous amount of data not currently available in such a large state. Third, the reference approach is particularly problematic because it automatically relegates a certain percentage of the reference sites to an "impaired" status. In addition, for many waterbody types, minimally disturbed reference sites are largely unavailable. Fourth, statistical stress-response relationships can be spurious, or have lots of unexplained variability (i.e., poor precision). This poor precision is translated to a larger margin of safety required (more conservative limits) for load allocations and permit limits. While waterbody typology, to some degree, can assist in explaining some of this variability, it cannot completely remove the concern. Thus, while simpler than the cause-effect approach, the empirical stress-response will result in more false negative and false positive determinations of adverse effects, and in the end will be more costly to the public.

For estuaries, reliance on the cause-effect approach is strongly suggested, because in the majority of circumstances, the reference or empirical stress-response approaches are simply untenable. Estuaries within California are highly variable in how they respond to nutrient loading due to differences in physiographic setting, salinity regime, frequency and timing of freshwater flows, magnitude of tidal forcing, sediment load, stratification, residence time, denitrification, etc. This combination of "co-factors" results in differences in the dominant primary producer communities (i.e., phytoplankton, macroalgae, benthic algae, submerged aquatic vegetation, emergent macrophytes). It also creates variability in the pathways that control how nutrients cycle within the estuary. At times, these co-factors can play a larger role in mitigating estuarine response to nutrient loads or concentrations, blurring or completely obscuring a simple prediction of primary productivity limited by nutrients (e.g., Figure 2.1). For example, in estuaries such as San Francisco Bay, synthesis of existing data by Cloern and Dugdale (2010) have clearly shown that surface water nutrient concentrations do not correlate with measures of primary productivity, in part because of important co-factors that override simple nutrient limitation of primary production.

2.3 Key Tenets of the NNE Approach

The NNE framework for California waterbodies is basely largely on the cause-effect approach. The intent of the NNE framework is to control excess nutrient loads to levels such that the risk or probability of impairing the designated uses is limited to a low level. If the nutrients present – regardless of actual magnitude – have a low probability of impairing uses, then water quality standards can be considered met.

The framework has three organizing principals (USEPA 2006):

1. *Ecological response indicators provide a more direct risk-based linkage to beneficial uses than nutrient concentrations or loads alone. Thus the NNE framework is based on the diagnosis of eutrophication or other adverse effects and its consequences rather than nutrient over enrichment per se.*

Except in some cases, such as unionized ammonium causing toxicity, nutrients themselves do not impair beneficial uses. Rather, ecological response to nutrient loading causes adverse effects that impair uses. Instead of setting objectives solely in terms of nutrient concentrations, it is preferable to use an analysis that takes into account the risk of impairment of these uses. The NNE framework needs to target information on ecological response indicators such as dissolved oxygen, surface water phytoplankton and harmful algal bloom (HAB) biomass (e.g., chlorophyll-a, water clarity), macroalgal biomass and percent cover, benthic algal biomass (sediment chlorophyll-a) and submerged aquatic vegetation (SAV) density and percent cover, and aesthetics (e.g., foul odors, unsightliness). These ecological response indicators provide a more direct risk-based linkage to beneficial uses than the ambient nutrient concentrations or nutrient loads. Given this approach, it is critical that tools be developed that link the response indicators back to nutrient loads and other co-factors and management controls (hydrology, etc.).

2. *A weight of evidence approach with multiple indicators will produce a more robust assessment of eutrophication.*

When possible, the use of multiple indicators in a “weight of evidence” approach provides a more robust means to assess ecological condition and determine impairment. This approach is similar to the multimetric index approach, which defines an array of metrics or measures that individually provide limited information on biological status, but when integrated, functions as an overall indicator of biological condition (Karr and Chu 1999).

3. *Use of “nutrient-response” models to convert response indicators to site-specific nutrient loads or concentrations.*

A key premise of the NNE framework is the use of models to convert numeric endpoints, based on ecological response indicators, to site- specific nutrient load goals appropriate for assessment, permitting, and TMDLs. A key feature of these models is that they account for site-specific co-factors, such as light availability, temperature, and hydrology that modify the ecological response of a system to nutrients.

2.4 Review of Science Supporting Nutrient Objective Development in San Francisco Bay

McKee et al. (2011) reviewed literature and data relevant to the assessment of eutrophication and other adverse effects of nutrient overenrichment in San Francisco Bay, with the goal of providing information to formulate a work plan to develop NNEs for this estuary. The review had three objectives: 1) Evaluate indicators to assess eutrophication and other adverse effects of anthropogenic nutrient loading in San Francisco Bay, 2) Summarize existing literature in SF Bay using indicators and identify data gaps, and 3) Investigate what data and tools exist to evaluate the trends in nutrient loading to the Bay (McKee et al. 2011).

Recommended NNE Indicators for SF Bay

As noted previously, an NNE assessment framework is the structured set of decision rules that helps to classify the waterbody in categories from minimally to very disturbed, in order to determine if a waterbody is meeting beneficial uses. Development of an assessment framework begins by choosing response indicators, which were reviewed using four criteria: 1) strong linkage to beneficial uses, 2) well -vetted means of measurement, 3) can model the relationship between the indicator, nutrient loads and other management controls, and 4) has an acceptable signal: noise ratio to assess eutrophication.

For San Francisco Bay, indicators varied among four habitat types: 1) unvegetated subtidal, 2) seagrass and brackish SAV, 3) intertidal flats, and 4) tidally muted habitats (e.g. estuarine diked Baylands). Two types of indicators were designated. Primary indicators are those which met all evaluation criteria and would therefore be expected to be a primary line of evidence of the NNE assessment framework for SF Bay. Supporting indicators fell short of meeting evaluation criteria, but may be used as supporting lines of evidence. This terminology is used in order to provide a sense of level of confidence in how the indicators should be employed in a multiple lines of evidence context.

The review found four types of indicators met all evaluation criteria and are designated as primary: dissolved oxygen, phytoplankton biomass, productivity, and assemblage, and cyanobacterial abundance and toxin concentration (all subtidal habitats), macroalgal biomass and cover (intertidal habitat, tidally muted habitats, and seagrass habitats; Table 2.1). Other indicators evaluated met three or fewer of the review criteria and designated as supporting indicators: HAB cell counts and toxin concentration, urea and ammonium (all subtidal), light attenuation and epiphyte load (seagrass/brackish SAV). Ultimately, the real distinction between “primary” and “supporting” and how these classes of indicators would be used as multiple lines of evidence in an NNE assessment is entirely dependent on indicator group and particular applications to specific habitat types. Some primary indicators (e.g. dissolved oxygen) could be stand-alone, while for others such as phytoplankton biomass, productivity and assemblage, the SF Bay Technical Advisory Team recommended using them as multiple lines of evidence, as use of any one alone is likely to be insufficiently robust.

Table. 2.1 Data gaps and next steps for development of an SF Bay NNE assessment framework.

Type	Indicator	Designation	Data Gaps	Recommended Next Steps
Subtidal Habitat	Dissolved oxygen	Primary	Wealth of data exists. Technical Advisory Team does not have expertise to review adequacy of DO objectives. Review did not address dissolved oxygen data in the tidally muted habitats of SF Bay.	Consider update of science supporting Basin Plan dissolved oxygen objectives, if warranted by additional review by fisheries experts. Review could be for entire Bay or limited to the tidally muted areas of the Bay.
	Phytoplankton biomass , productivity, and assemblage	Primary	Need a review of science supporting selection of endpoints. Improved prediction of factors controlling assemblage	Recommend development of a white paper and a series of expert workshops to develop NNE assessment framework for phytoplankton biomass, productivity, taxonomic composition/assemblages, abundance and/or harmful algal bloom toxin concentrations. Recommend augmentation of current monitoring to include measurement of HAB toxin concentrations in water and faunal tissues.
	HAB species abundance and toxin conc.	Cyanobacteria = primary; Other HAB =supporting	Little data on HAB toxin concentrations in surface waters and faunal tissues.	
	Ammonium and urea	Supporting	Lack of understanding of importance of ammonia limitation of nitrate uptake in diatoms on Bay productivity vis-à-vis other factors. Lack of data on urea in SF Bay	Recommend formulation of a working group of SF Bay scientists to synthesize available data on factors known to control primary productivity in different regions in the Bay, and evaluate potential ammonium endpoints. Recommend collecting additional data on urea concentrations in SF Bay via USGS's water quality sampling over a two year period.
	Macrobenthos taxonomy, abundance and biomass	Co-factor	Lack of information on how to use combination of taxonomy, abundance, and biomass to assess eutrophication	Recommend utilization of IE-EMP dataset to explore use of macrobenthos to be used reliably to diagnose eutrophication distinctly from other stressors in oligohaline habitats. This may involve including biomass in the protocol to improve ability to diagnose eutrophication.
Seagrass Habitat	Phytoplankton biomass, epiphyte load and light attenuation	Phytoplankton biomass = primary, epiphyte load and light attenuation = secondary	Poor data availability of data on stressors to SF Bay seagrass beds. Studies needed to establish light requirements for seagrass and to assess effects of light attenuation	Recommend 1) Continued monitoring of aerial extent of seagrass every 3-5 years (currently no further system scale monitoring is planned beyond 2010), 2) studies to establish light requirements for SF Bay seagrass species, 3) development of a statewide workgroup to develop an assessment framework for seagrass based on phytoplankton biomass, macroalgae, and epiphyte load and 4) collection of

Type	Indicator	Designation	Data Gaps	Recommended Next Steps
	Macroalgae biomass and cover	Primary	Data gaps include studies to establish thresholds of macroalgal biomass, cover and duration that adversely affect seagrass habitat	baseline data to characterize prevalence of macroalgal blooms on seagrass beds. Studies characterizing thresholds of adverse effects of macroalgae on seagrass currently underway in other California estuaries should be evaluated for their applicability to SF Bay.
Intertidal Flat Habitat	Macroalgal biomass and cover	Primary	Lack of baseline data on frequency, magnitude (biomass and cover) and duration of macroalgal blooms in these intertidal flats	Recommend collection of baseline data on macroalgae, microphytobenthos and sediment bulk characteristics. Recommend inclusion of SF Bay scientists and stakeholders on statewide workgroup to develop an assessment framework for macroalgae on intertidal flats.
	Sediment nutrients	Supporting		
	MPB taxonomy and biomass	Supporting		
Muted Subtidal Habotat	Macroalgae	Primary	Lack of baseline data on biomass and cover in muted habitat types	Recommend collection of baseline data on macroalgae, dissolved oxygen, phytoplankton biomass, taxonomic composition and HAB species/toxin concentration in these habitat types. Recommendation to develop an assessment framework based on macroalgae, phytoplankton and dissolved oxygen in these habitat types. One component of this discussion should be a decision on beneficial uses that would be targeted for protection and to what extent the level of protection or expectation for this habitat type differ from adjacent subtidal habitat.
	Phytoplankton biomass, assemblage, HAB toxin conc.	Phytoplankton biomass, cyanobacteria = primary; assemblage and other HABs= supporting	Lack of baseline data on biomass and community composition, HAB toxin concentrations	
	Dissolved oxygen	Primary	Some data on dissolved oxygen exist. Unclear what levels of DO required to protect muted habitat beneficial uses	

The use of ammonium as an indicator received review, due to its hypothesized role in limiting phytoplankton primary production via nitrate uptake inhibition in Suisun Bay and the lower Sacramento River. The SF Bay Technical Advisory Team chose to include it as a supporting indicator because the importance of ammonium inhibition of diatom blooms relative to other factors controlling primary productivity Bay wide is not well understood. Additional review and synthesis were recommended, pending currently funded studies, to identify potential ammonium thresholds.

Table 2.1 summarizes data gaps and recommended next steps by McKee et al. (2011) for development of an SF Bay NNE assessment framework by habitat type. Data gaps and recommendations generally fall into four categories: 1) Monitoring to assess baseline levels of indicators of interest where data are currently lacking, 2) Analysis of existing data, 3) Field studies or experiments to collect data required for endpoint development, and 4) Formation of expert workgroups to recommend approach to assessment framework development and synthesize information to be used in setting numeric endpoints.

2.5 Indicators Under Further Consideration for the SF Bay NNE Assessment Framework

The SF Bay Water Board, with advice from stakeholders, chose to prioritize the development of NNE assessment framework for subtidal habitats in SF Bay. Seagrass, intertidal habitat, and diked Baylands are not included in this initial work. For subtidal habitat, McKee et al. (2011) review recommended developing regulatory endpoints for subtidal habitat based on indicators of phytoplankton, nutrient concentrations, and dissolved oxygen. Work to review the science supporting dissolved oxygen objectives will be completed separately from this effort; thus assessment framework development will focus on indicators and metrics of phytoplankton and nutrient concentrations.

Phytoplankton

Phytoplankton are unicellular organisms, which serve a critical ecosystem function of primary production, forming the base of pelagic foodwebs in many aquatic environments. Phytoplankton blooms are a natural phenomenon, typical of spring and summer periods of naturally high primary production which supplies energy to the ecosystem. However, phytoplankton respond rapidly to changes in nutrient concentrations and nutrient enrichment, which can lead to more frequent blooms, of greater intensity, and spatial and temporal extent [*Carstensen et al.*, 2011; *Cloern*, 2001]. Increased biomass is typically the first response to nutrient enrichment, often followed by species shifts, and accumulation of organic matter which results in oxygen depletion in the bottom water of stratified areas [*Cloern*, 1996; 2001; *W M Kemp et al.*, 2005]. Excessive blooms can also increase turbidity such that light penetration through the water column is significantly reduced, thus restricting growth of seagrasses [*Huntington and Boyer*, 2008]. Over production of harmful, toxin producing species can also result in ecosystem effects through poisonings of marine mammals and birds.

Because of their direct link and rapid response to nutrient additions, phytoplankton are considered a primary symptom of eutrophication and have been used extensively as a gauge of ecological condition and change [*Bricker et al.*, 2003; *Domingues et al.*, 2008]. Phytoplankton is used as an indicator or water quality element in various forms in a number of assessment frameworks and is typically considered one of the more robust in terms of establishment of thresholds [*Borja et al.*, 2011].

There are a number of considerations for using phytoplankton as an indicator of eutrophication [*Domingues et al.*, 2008]. Firstly, the establishment of reference condition for water quality may be difficult in systems for which there is no historical data. Secondly, there is a lack of guidance on sampling frequency, and for several water quality frameworks, the proposed frequency is insufficient to assess phytoplankton succession and may even preclude the detection of algal blooms. Finally, the use of chlorophyll-a as a proxy for biomass may overlook blooms of pico- and small nanoplankton, and overestimate the importance of large microphytoplankton because cellular chlorophyll-a content is often species-specific [*Domingues et al.*, 2008].

Phytoplankton Biomass (Chlorophyll-a Concentration, Bloom Intensity and Frequency

Chlorophyll-a is measured as a way to estimate the active phytoplankton biomass and is used extensively as an indicator of eutrophic condition for estuarine waters. Chlorophyll is the green pigment in all plants and Chlorophyll-a is the most common type of chlorophyll. Plants use chlorophyll to capture sunlight for photosynthesis. Chlorophyll-a concentrations are often highest just below the surface, not at the surface of the water.

Chlorophyll-a can be measured in several ways: discrete measures, continuous measurements via data sonde, and remote sensing. Discrete samples of chlorophyll-a are measured by filtering a known amount of sample water through a glass fiber filter. The filter paper itself is used for the analysis. The filter is ground up in an acetone solution and either a fluorometer or spectrophotometer is used to read the light transmission at a given wavelength, which in turn is used to calculate the concentration of chlorophyll-a. Continuous measurements in the field are made with a fluorometer probe mounted to a data sonde or similar logging device. The in situ water is exposed to light of a single wavelength. Some substances in the water sample, including chlorophyll-a, will give off light, or fluoresce, in response to the light. The amount of light emitted by the chlorophyll-a is measured and used to calculate the chlorophyll-a concentration. Field fluorometers must be calibrated routinely against discrete samples for accuracy. Chlorophyll-a is also measured remotely by satellite. Satellites measure the color of seawater to determine the amount of chlorophyll present. The ocean color is often blue, but the satellite can detect very small changes in the ocean color as a result of the chlorophyll in phytoplankton. Satellite measurements need to be compared to discrete measurements to calibrate the satellite measurements.

Phytoplankton blooms are expected to increase in frequency, duration and spatial extent as water bodies continue to experience nutrient over enrichment [*Bricker et al.*, 2003]. Bloom duration can be directly quantified using continuous monitoring data. Frequency and spatial extent are typically assessed heuristically in the field and binned into groups (periodic versus episodic for frequency, and high, moderate, low and very low for spatial coverage) [*Bricker et al.*, 2003].

Phytoplankton Productivity

Primary production is the process by which autotrophic organisms "fix" inorganic carbon using solar energy to carry out metabolic processes and build cellular material. Production in marine waters is influenced by the supply of nutrients, light, temperature, flow regime, turbidity, zooplankton grazing and toxic substances. Low rates of annual primary production may indicate low susceptibility to enrichment while high rates of annual primary production represent higher susceptibility, possibly resulting in symptoms associated with undesirable disturbance [*Cloern*, 2001; *Devlin et al.*, 2007a; *S J Painting et al.*, 2007].

This productivity is typically measured using ^{14}C radiolabeling to measure the rate of carbon uptake over a defined area or volume. The method is based on the assumption that biological uptake of ^{14}C-labelled dissolved inorganic carbon (DIC) is proportional to the biological uptake of the more commonly found ^{12}C DIC. In order to determine uptake, one must know the concentration of DIC naturally occurring in the sample water, the amount of ^{14}C-DIC added, and the amount of ^{14}C retained in particulate matter (^{14}C-POC) at the end of the incubation experiment [*Steeman-Nielsen*, 1952].

Phytoplankton Taxonomic Composition or Assemblage

Changes in phytoplankton community composition are expected to occur as eutrophication develops in estuarine environments. Shifts may reflect a loss of biodiversity of organisms and a shift towards dominance of one or more species, but they often include increased abundance of opportunistic nuisance and toxic species that result from changing nutrient concentrations and ratios [*Borja et al.*, 2011]. Samples for phytoplankton taxonomy can be collected from whole water or can be collected using one or more phytoplankton nets of targeted mesh size. There are several methods for estimating phytoplankton community composition: identification and cell counts using microscopy, flow cytometry/particle counting, and pigment analysis by HPLC. Each has its own advantages and disadvantages , but all provide some measure of phytoplankton community structure [*R A Anderson*, 2005; *P E Kemp et al.*, 1993].

Harmful Algal Bloom Dominance and Toxin Concentrations

Some algal blooms may include a shift towards dominance of nuisance or toxic species which may have a detrimental impact to biological resources [*Bricker et al.*, 2003]. For example, excessive abundance of small phytoplankton species may clog the siphons of filer feeding bivalves and may cause respiratory irritation to fish. Excessive abundance of toxin producing organisms can result in poisonings of marine mammals and birds. Presence of nuisance and toxic species can be identified by the methods described above in phytoplankton community composition. Algal toxins can be measured on whole water samples using spectrophotometric and HPLC techniques.

Nutrient Concentrations and/or Ratios

Eutrophication is primarily caused by nutrient enrichment leading to increased production of organic matter [*Nixon*, 1995]. Primary producers need nutrients for growth and low concentrations of bioavailable nitrogen and phosphorus will limit primary production. Estuarine nutrient concentrations are highly dynamic and are rapidly transformed by biogeochemical processing. The concentrations of dissolved inorganic nutrients in the water column represents the instantaneous net “remainder” after processing by all other factors. Ambient nutrient concentrations are often correlated with nutrient loading into the systems [*Boynton and Kemp*, 2008; *Conley et al.*, 2000; *Hejzlar et al.*, 2009; *Smith et al.*, 2005]. Though empirical relationships between nutrient concentrations and biological response are dependent on a variety of site specific conditions and are highly variable among systems [*Carstensen et al.*, 2011; *Cloern*, 2001].

Both nitrogen and phosphorus can be limiting either exclusively or in combination (co-limitation). Ambient nutrient concentrations of dissolved inorganic nitrogen (DIN) or dissolved inorganic phosphorus (DIP) are used to determine nutrient limitation, usually with the suggestion that primary production is N-limited for DIN:DIP ratios below 10 and mainly P-limited for DIN:DIP ratios greater than 20 [*L A Anderson and Sarmiento*, 1994; *Klausmeler et al.*, 2004; *Redfield et al.*, 1963]. During blooms, ambient nutrient concentrations may become almost completely consumed, resulting in strong seasonal variability in nutrient concentrations. Changes in estuarine geomorphology also result in wide spatial variability in N- and P-limitation, due to variation in supply, removal, and biogeochemical transformations of nutrients [*Carstensen et al.*, 2011].

Relatively recent shifts in our conceptual understanding of eutrophication [*Cloern*, 2001; *Devlin et al.*, 2007a; *S J Painting et al.*, 2007] indicate that estuaries can have complex responses to nutrient inputs, including both direct and indirect responses, and the role additional factors that moderate ecosystem response. In estuarine systems, factors such as light climate and hydrology, affect the susceptibility of different waterbodies to nutrient enrichment [*S J Painting et al.*, 2007]. Consequently, the presence of high nutrient concentrations should be regarded as a potential cause for concern and may trigger further assessment of biological response indicators. Given the current understanding of the consequences of nutrient enrichment it is clear that, for any given aquatic situation, it is not possible to determine specific nutrient thresholds without reference to the biological response [*Devlin et al.*, 2007a].

3 Review of Existing Assessment Methods/F rameworks

3.1 Regulatory Criteria

A number of states and programs within the U.S. are in the process of developing nutrient criteria or biocriteria to protect waterbodies from nutrient overenrichment. Typically, these criteria are based on three types: 1) TN and TP, 2) water column chlorophyll a and 3) dissolved oxygen. Many programs have established narrative criteria for biological response indicators and are in the process of collecting monitoring data that would support the development of numeric values that are protective for specific estuaries (e.g. Maryland, Maine, and Chesapeake Bay for chlorophyll a). Florida has recently established site-specific TN and TP and chlorophyll a criteria for all the State's estuaries. Table 3.1 summarizes existing TN, TP and chlorophyll a criteria for estuaries and tidal rivers.

Of these states, the criteria promulgated for Florida estuaries and Chesapeake Bay represent the most scientifically well-documented approaches to establishing nutrient and chlorophyll a endpoints (USEPA 2007, USEPA 2010). In both cases, estuarine surface TN and TP criteria were established via modeling linkages with biological endpoints (maintenance of seagrass, maintenance of balanced algal population, dissolved oxygen). Although relevant for nutrient-response modeling of SF Bay, we choose not to include a synthesis of this work in our review. Establishment of chlorophyll a criteria based on maintenance of seagrass, which currently represent less than 3% of subtidal habitat in the Bay, is also not a relevant paradigm for SF Bay. Therefore we summarize the scientific paradigms and approaches used in Florida and for the Chesapeake Bay that relevant for the "maintenance of balanced algal populations."

Table 3.1 Summary of existing chl- a criteria by state for lakes and estuaries. Adapted from U.S. EPA. 2003. Survey of States, Tribes and Territories Nutrients Standards. Washington, DC

State	Chlorophyll a Numeric Criteria in Estuaries (all values in µg L-1 unless otherwise noted)
District of Columbia	Seasonal July 1–September 30 segment average chlorophyll *a* concentration of 25 applied to tidally influenced waters only.
Florida	In unvegetated subtidal habitats, chlorophyll a should not exceed 20 for greater than 10% of the time.
Hawaii	Chlorophyll *a* criteria applying to different locations within Lake Mead ranging from 5–45
North Carolina	Freshwater class C waters and tidal saltwaters: For lakes and reservoirs and other waters subject to growths of macroscopic and microscopic vegetation not designated as trout waters: <40. For lakes and reservoirs and other waters subject to growths of macroscopic and microscopic vegetation designated as trout waters: <15.
Oregon	Chlorophyll *a* criteria for: • Natural lakes which do not thermally stratify: <10 • Natural lakes which do not thermally stratify, reservoirs, rivers and estuaries: <15 (OAR340-041-0019)
Virginia	Site specific seasonal numerical chlorophyll *a* criteria applicable March 1–May 31 and July 1–September 30 for the tidal James River segments JMSTF2, JMSTF1, JMSOH, JMSMH, JMSPH (9 VAC 25-260-310), ranging from 10-23.

Florida

In Florida, the rationale or establishment of chlorophyll a criteria to protect a "balanced algal population" is based on the premise that nutrient-driven effects on algal growth and biomass accumulation can result in more frequent, short term blooms that decrease water clarity, adversely affect aesthetics, recreation, and aquatic life habitat. They specifically cite: 1) the increased harmful algal blooms, which can produce toxins that adversely affect both human health and aquatic life and 2) the effect of frequent algal blooms on the long-term balance of organic matter cycling within an estuary (Nixon 1995), leading to hypoxia or anoxia, which also can adversely affect habitat and aquatic life. Because toxic blooms are a frequent occurrence in Florida estuaries and coastal waters, EPA deemed appropriate the derivation of chlorophyll criteria on the basis of reducing the likelihood of nuisance algal blooms on recreation and recreational uses (Larkin and Adams 2007; Walker 1985).

Specific chl-a concentrations consistent with nuisance conditions were defined in that literature on the basis of <u>trophic state boundaries</u>, <u>user perception studies</u>, and <u>observed impacts</u>. While they acknowledge documentation supporting trophic state chl a thresholds is limited, they cite: 1) Assessment of Estuarine Trophic Status (ASSETS, Bricker et al. 2003), in which low algal bloom conditions were defined as maximum chl-a concentrations < 5 µg/L, medium bloom conditions as maximum chl-a concentrations 5–20 µg/L, high bloom conditions as maximum chl-a concentrations 20–60 µg/L, and hypereutrophic conditions as maximum chl-a concentrations above 60 µg/L and 2) the United Kingdom Comprehensive Studies Task Team maximum summer chl-a value of 10 µg/L as an estuarine eutrophic threshold (Painting et al. 2007. EPA maintained that frequently occurring, elevated chlorophyll a concentrations can be an expression of dominance by one or more phytoplankton species, potentially toxic or otherwise harmful or nuisance algae, citing cyanobacterial blooms in freshwater and brackish habitats (Chorus et al. 2000) and marine HABs (Anderson et al. 2008; Paerl et al. 2008; Glibert et al. 2010). They also utilized information on bloom frequencies typical of Florida estuaries and then identified concentrations typical of blooms of harmful or nuisance algae and indicative of imbalance of phytoplankton populations. One estimate for the range of observed monthly chl-a maxima was from 15 to 25 µg/L, depending on the type of estuary (coastal embayment, river-dominated, or lagoon) (Glibert et al. 2010). In a national survey, the average bloom chl-a concentrations were 20 µg/L or less for 7 of 10 large estuaries; concentrations were especially low for Florida Bay (8 µg/L) and Pensacola Bay (10 µg/L, Glibert et al. 2010) and higher for the St. Johns River Estuary (20 µg/L, Bricker et al. 2007). Based on this work, EPA selected a chl-a concentration target of 20 µg/L, with an allowable exceedance frequency of no more than 10 percent of monitoring data.

Chesapeake Bay

In the Chesapeake Bay, multiple lines of evidence were used to derived chlorophyll a criteria (EPA 2007), based on adverse effects associated with high chl-a in Chesapeake Bay include seasonal hypoxia or anoxia (Smith et al. 1992, Hagy et al. 2004, Bricker et al. 2008), decreased water clarity affecting submerged aquatic vegetation (SAV) (Dennison et al. 1993, Kemp et al. 2004), and blooms of potentially harmful algal taxa (HABs) (Cloern 2001, Marshall et al. 2005, 2009, Mulholland et al. 2009, Morse et al. 2011).These lines of evidence included (1) analysis of historical and recent data to establish baseline chl-a for the mainstem Bay; (2) detection of long-term trends of chl-a; (3) quantification of climatic forcing of chl-a; (4) identification of a relationship between DO and chl-a; (5) quantification of the effects of chl-a on water clarity and habitat suitability for SAV; (6) establishment of linkages between chl-a and cyanobacteria toxin concentrations.

Thresholds for the historical reference periods (1960-1980) ranged from 15 to 35 µg L^{-1} in spring, and from 7 to 54 µg L^{-1}, with the 1970s having higher thresholds than the 1960s (EPA 2007,). The oligohaline region had the highest surface chl-a thresholds, declining to the lowest thresholds for the polyhaline portion of Chesapeake Bay. The lowest thresholds were ~ 4-7 µg L^{-1}in the polyhaline region for the 1960s ranging up to the highest thresholds were ~ 40-55 µg L^{-1} in the oligohaline region for the 1970s historical reference period. The mesohaline and polyhaline regions had higher thresholds for surface chl-a in high-flow conditions than in mid- or low-flow conditions while the oligohaline region had higher thresholds for surface chl-a in low-flow than in high-flow conditions. The lowest thresholds were ~ 4-7 µg L^{-1}in the polyhaline region for the 1960s ranging up to the highest thresholds were ~ 40-55 µg L^{-1} in the oligohaline region for the 1970s historical reference period.

Low summer bottom-water DO occurred at high chl-a, with no observations of DO > 3 mg L-1 (the deep-water 30-d mean DO criterion) when May-Aug chl-a was > 16 µg L^{-1}, or of DO > 1.7 mg L-1 (the minimum DO criterion for fish; USEPA 2003) when May-Aug chl-a was > 22 µg L^{-1}.

Diatoms usually dominate the floral composition of Chesapeake Bay, with seasonally variable contributions by other algal taxa including dinoflagellates, cryptophytes, and cyanobacteria whose abundance varied seasonally. Exceptional occurrences of dinoflagellates blooms were not sufficient to support chl-a criteria on regional and seasonal bases. However, in tidal fresh and oligohaline regions, toxic blooms of the cyanobacteria, *Microcystis aeruginosa*, can reach high chl-a in summer. Simple linear regression showed significant relationships ($p < 0.05$) between surface chl-a and cell counts of *M. aeruginosa* for the upper Bay and four of seven tidal tributaries. Chl-a thresholds separating high-risk from middle- and low-risk for surface and above-pycnocline chl-a and were 29.2 and 29.0 µg L^{-1}, respectively. A threshold of 27.5 µg L^{-1} was established as protective against toxic Microcystis in the Bay (U.S. EPA 2007).

Based on these analyses, a set of reference criteria were developed for Chesapeake Bay (summarized in Table 3.2). These reference concentrations should only be applied to mainstem Chesapeake Bay surface, open-water habitats only during the spring (March 1 through May 31) and summer (July 1 through September 30) seasons, the most critical seasons for addressing algal-related impairments.

Although community composition was not directly incorporated into the EPA 2007 analysis, Buchanan et al. (2005) quantified the habitat conditions supporting phytoplankton reference communities in Chesapeake Bay. They reported maximum spring and summer chlorophyll a concentrations (in µg·liter-1), respectively, for tidal fresh (13.5, 15.9), oligohaline (24.6, 24.4), mesohaline (23.8, 13.5), and polyhaline (6.4, 9.2).

Table 3.2 Chesapeake Bay chlorophyll a reference concentrations (from EPA 2007).

Salinity Regime[2]/ Water Column Location	Season[3]	Water Clarity Criteria Application Depth[4] (m)	Chlorophyll *a* Reference Concentration ($\mu g \cdot liter^{-1}$)
Historical Chlorophyll *a* Reference Concentrations[5]			
Oligohaline	Spring	–[6]	18
Mesohaline	Spring	–	8
Polyhaline	Spring	–	4
Oligohaline	Summer	–	46
Mesohaline	Summer	–	23
Polyhaline	Summer	–	5
Dissolved Oxygen Impairment-Based Chlorophyll *a* Reference Concentrations			
Deeper Waters Which Stratify	Annual	–	10–15
Shallow Waters	Annual	–	30
Water Clarity Impairment-Based Chlorophyll *a* Reference Concentrations			
Tidal Fresh/Oligohaline	SAV	0.5	43
Tidal Fresh/Oligohaline	SAV	1.0	11
Mesohaline/Polyhaline	SAV	0.5	39
Mesohaline/Polyhaline	SAV	1	16
Mesohaline/Polyhaline	SAV	2	3

[1]All chlorophyll *a* reference concentrations apply as $\mu g \cdot liter^{-1}$ across the surface waters of open-water designated-use segments for the applicable salinity regime and season.

[2]Tidal Fresh = 0 – <0.5 ppt salinity; oligohaline = 0.5– <5 ppt salinity; mesohaline = 5–18 ppt salinity; polyhaline = >18 ppt salinity.

[3]Spring = March 1–May 31; Summer = June 1–September 30; SAV or SAV growing season: for tidal-fresh, oligohaline, and mesohaline habitats = April 1–October 31; for polyhaline habitats = March 1–November 30 (U.S. EPA 2003a).

[4]Water clarity criteria application depth for each Chesapeake Bay Program segment as published in U.S. EPA 2003b and as adopted into Delaware, Maryland, Virginia and the District of Columbia's water quality standards regulations.

[5]Reference concentrations only apply to mainstem Chesapeake Bay segments.

[6]Not applicable.

3.2 Non-Regulatory Assessment Frameworks

Over the past decade, much work has been done to establish standardized methodologies to assess ecological quality in estuaries, with several methods developed specifically for eutrophication [*Andersen et al.*, 2011; *Bricker et al.*, 2003; *Devlin et al.*, 2011; *Domingues et al.*, 2008; *Zaldivar et al.*, 2008] and conduct surveys to evaluate the magnitude and extent of eutrophication [*Andersen et al.*, 2011; *Borja et al.*, 2009; *Bricker et al.*, 1999; *Devlin et al.*, 2011; *Garmendia et al.*, 2012].

In Europe, there has been a vast expansion in methods, due to the adoption of the European Union Water Framework Directive (WFD). The aim of the WFD is to achieve good ecological status in all EU member state waterbodies, where good status represents a no more than 50% deviation from reference conditions. Assessments are carried out at a waterbody level, and reference conditions are defined for each waterbody type based on characteristics including tidal range, mixing, exposure and salinity [*Devlin et al.*, 2011]. Each EU member state is required to adopt the WFD process though the selection of waterbody types, reference conditions, specific indicator variables and assessment methods can vary among member states [*VIncent et al.*, 2002]. Birk et al. (2012) document over 300 methods developed for compliance with the WFD alone, as many countries preferred developing country-specific methods instead of a handful of methods applicable Europe-wide (e.g. Birk and Schmedtje, 2005; Borja et al., 2009).

Assessment Framework Utilizing Multiple Categories of Indicators

Several indicator-based assessment frameworks have been developed to assess eutrophic condition of estuaries with respect to eutrophication utilizing multiple indicators. The most representative assessment frameworks have been found to incorporate annual data with sampling throughout the year, to capture frequency of occurrence and spatial extent in indicator metrics, and use of a combination of indicators into an overall condition rating [*Devlin et al.*, 2011].

Tables 3.3-3.4. provides a brief summary of integrated assessment frameworks that utilize multiple groups of indicators (Ferreira et al. 2011). Studies comparing eutrophication status results generated for the same system using different assessment frameworks have indicated that results can vary slightly depending on which framework is applied (Table 3.5) [*Devlin et al.*, 2011; *Garmendia et al.*, 2012]. Different frameworks apply similar indicators, but differences in timeframes of data analysis (seasonal versus annual), characteristics included in the indicator metrics (concentration, spatial coverage, frequency of occurrence), and how to combine indicators into multiple lines of evidence, had an effect on the overall outcome of the assessment [*Devlin et al.*, 2011].

Table 3.3 Methods of eutrophication assessment and examples of biological and physico-chemical indicators used and integration capabilities (pressure-state and overall; modified from Borja et al. 2012). From Ferreira et al. 2012.

Method Name	Biological indicators	Physico-chemical indicators	Nutrient load related to impairments	Integrated final rating
TRIX[b]	Chl	DO, DIN, TP	no	yes
EPA NCA Water Quality Index[a]	Chl	Water clarity, DO, DIN, DIP	no	yes
ASSETS[e]	Chl, macroalgae, seagrass, HAB	DO	yes	yes
TWQI/LWQI[c]	Chl, macroalgae, seagrass	DO, DIN, DIP	no	yes
OSPAR COMPP[g]	Chl, macroalgae, seagrass, phytoplankton indicator species	DO, TP, TN, DIN, DIP	yes	yes
WFD[f]	phytoplankton, Chl, macroalgae, benthic invertebrates, seagrass,	DO, TP, TN, DIN, DIP, water clarity	no	yes
HEAT[d]	Chl, primary production, seagrass, benthic invertebrates, HAB, macroalgae	DIN, DIP, TN, TP, DO, water clarity	no	yes
IFREMER[h]	Chl, seagrass, macrobenthos, HAB	DO water clarity, SRP, TP, TN, DIN, sediment organic matter, sediment TN, TP	no	yes
STI[i]	Chl, Primary Production	DIN, DIP	no	no

[a] USEPA, 2005, 2008.
[b] Vollenweider *et al.*, 1998.
[c] Giordani *et al.*, 2009.
[d] HELCOM, 2009.
[e] Bricker *et al.*, 1999, 2003, 2007.
[f] Devlin, pers.Com.
[g] OSPAR, 2002, 2008.
[h] Souchu *et al.*, 2000.
[i] Ignatiades, 2005.

Table 3.4. Summary of approaches used for assessment of eutrophication applicable to shallow and deepwater unvegetated subtidal habitat. Adapted from Devlin et al. 2011.

		UK WFD	OSPAR	TRIX	ASSETS	EPA NCA	TWQI/LWQF	HEAT
Grouping of Variables	Causative Factors	Nutrient Load	DIN and DIP concentration, ratios, and loads	DIN and TP concentration	DIN and DIP loads	DIN, DIP conc	TN, TP, DIN and DIP conc.	DIN and DIP
Grouping of Variables	1^{ary} effects	Chl-a, PP indicator species, seasonal changes in cell abundance of diatoms/dinoflagellates, SAV, macroalgae	Chl-a, PP indicator species, macroalgae, microphytobenthos, SAV	Chl-A	Chl-a macroalgae	water clarity, chl-a	Chl a, SAV, macroalgae	Chl a, water clarity, SAV,
Grouping of Variables	2^{ary} effects	DO	DO, zoobenthos and/or fish kills, organic carbon	DO	Nuisance/toxic blooms	DO	DO	Benthic invertebrates
Grouping of Variables	Other effects		Algal toxins					
Temporal sampling framework		Annual chla and DO, winter DIN, monthly PP groups	Growing season chl-a (Mar-Sept), Winter DIN, summer DO	Annual	Annual	One sample per year (per station) within summer index period	Results can be derived based on one time period, multiple periods recommended	Growing season chl-a (Mar-Sept), Winter DIN, summer DO
Spatial sampling framework		Sampling in estuaries and nearshore defined by salinity, reported by waterbody	Sampling defined by salinity in estuaries, nearshore	Sampling mostly in larger offshore systems; results reported by region	Sampling in salinity zones, synthesized to waterbody, region, then national, with reporting at all levels	Sampling is regional, synthesized to national level, reported at regional and national level	For shallow, benthic PP dominated. Can be applied to single stations or groups of stations.	Sampling defined by salinity in Baltic Sea
Assessment of indicators		Deviation from reference conditions	Deviation from reference conditions	Placement on scale from 1-10 TRIX units	Deviation from reference conditions	Deviation from reference conditions	Deviation from reference condition	Deviation from reference condition
Comb-ination Method		Indicator scores are averaged within in indicator group. Final score gives classification status	One out, all out for individual categories and overall classification	Linear combo of logarithm of variables modified by scaling coeff.	Scores of ave. primary and secondary indicators combined in a matrix	Indicators assessed individually. WQI based on % of samples in 4 categories.	TWQI scores combined as the sum of weighted quality values for individual indicators.	One out, all out for individual categories and overall classification

Table 3.5 Summary of procedures used for evaluating the eutrophic status of estuarine and coastal waters and categories used for final classification. From Devlin et al. 2011.

Method	Summary of assessment procedure	Categories for final classification of status, and colour coding (Fig. 4)
WFD[a] used in the UK	Considers physico-chemical and biological QEs. For each element, one or more indicators are used. EQRs are calculated for each indicator using the ratio between measurements and reference condition. An EQR is calculated for each quality element by averaging the indicator scores for all components within each QE. The final BQE score, a number between 0 (worst) and 1 (best) relates to the five equidistant boundary classes. The worst classification status for the BQE's and the physico-chemical status are taken as the overall water body status. In this one-out-all-out approach, if any element has a state less than or equal to Moderate, then the water body is considered to be impacted	Classification status is calculated from the worst BQE BQE's are integrated by normalising outputs to 0–1 score (EQR score): High (best, blue) Good (green) Moderate (yellow) Poor (orange) Bad (worst, red)
OSPAR COMPP[b]	Thresholds are set for each indicator for estuarine, coastal and offshore waters. The full OSPAR COMPP procedure is applied only to areas indicated as Potential Problem Area or Problem Area by the screening procedure. Indicators are used in Categories I to IV to determine levels of nutrient enrichment and impacts, and combined within each category in a one-out-all-out process. A final assessment as Problem Area (+), Potential Problem Area or Non Problem Area (−) is made after evaluating scores for Categories I–IV with the one-out-all-out approach	Problem area (+) Potential problem area Non problem area (−)
TRIX[c]	The annual average of all parameters are combined without assigning individual parameter ratings, to give a ranking between 0 and 10, where: $TRIX = [\log10(\text{Chl } a) + \log10(\text{aD\%O}) + \log10(\text{DIN}) + \log10(\text{TP}) - k]/m$ The coefficients $k = -1.5$ and $m = 1.2$ are fixed to establish the lower and upper limits of the index	Index is scaled from 0 to 10, ranging from oligotrophy (scarcely productive—open sea) to eutrophy (highly productive)
ASSETS[d]	Three components are included: Influencing Factors (IF), Eutrophic condition (EC) and Future Outlook (FO) and the three components are combined into one ASSETS rating. Only the EC component is included in this study Five indicators are evaluated: primary symptoms (Chl-*a*, macroalgae) and secondary symptoms (DO, SAV, nuisance/toxic blooms). Area-weighted values are calculated from salinity zone assessments and area-weighted system wide ratings are made for each variable. Primary (average) and secondary (worst) symptom scores are combined in a matrix to determine an overall score or rating for eutrophic condition of the waterbody	The final scores for eutrophication status are: High (worst, red) Moderate high (orange) Moderate (yellow) Moderate low (green) Low (best, blue) Colour codes are consistent with WFD scales

Table. 3.5 continued

Method	Summary of assessment procedure	Categories for final classification of status, and colour coding (Fig. 4)
EPA NCA[e]	The Water Quality Index (WQI) uses a combination of DIN, DIP, DO, Chl-*a* and Water Clarity for assessment. Indicator scores are determined using the percentage of samples that are Good, Fair or Poor/No Data, e.g. DIN: Good: <10% of samples are Poor and >50% are Good Fair: 10–25% of samples are Poor and/or >50% are Poor or Fair Poor: >25% of samples are Poor All indicators are combined in a similar fashion to determine the rating for a site: Good = A maximum of one indicator is Fair and no indicators are Poor Fair = One of the indicators is rated Poor or two or more indicators are Fair Poor = Two or more of the five indicators are rated Poor To determine the WQI by region and nation, results from each area are used to determine a final assessment score where: Good: <10% of areas are in Poor condition and >50% are Poor or Fair Fair: 10–20% of areas are in Poor condition or >50% are Fair or Poor Poor: >20% of areas are in Poor condition	A final WQI assessment is determined by site and then by region and nation: Good (green) Fair (yellow) Poor (red)

Pressures (nutrient loading) and Future Outlook are not included here, as they are not used by all approaches

EQR ecological quality ratio, *DIN* dissolved inorganic nitrogen, *DIP* dissolved inorganic phosphorus, *TP* total phosphorus, *Chl-a* chlorophyll *a*, *DO* dissolved oxygen, *SAV* submerged aquatic vegetation, *spp.* species, *BQE* biological quality element, *PP* phytoplankton, *QE* quality element

[a] CEC (1991a, b)

[b] OSPAR (2002, 2005, 2008)

[c] Vollenweider et al. (1998)

[d] Bricker et al. (1999, 2003, 2007), Ferreira et al. (2007), www.eutro.org and www.eutro.org/register

[e] USEPA (2001a, b, 2005, 2008)

UK WFD Framework for Eutrophication

Here we review the United Kingdom (UK) assessment protocol for eutrophication. The WFD classifies waterbodies into one of five ecological condition categories: High, Good, Moderate, Poor or Bad. Initial risk of eutrophication is assessed based on nutrient load, turbidity, flushing time, and tidal range. The ecological condition category is assessed using three biological quality elements: phytoplankton, macroalgae, and angiosperms. The final assessment also includes a measure of physico-chemical status including dissolved inorganic nitrogen and dissolved oxygen.

Each biological quality element consists of one or more indicators that measure different aspects of the biological community (phytoplankton includes CHL-a and cell counts of abundance and composition, macroalgae includes biomass and areal coverage, angiosperms include biomass and area coverage) [*Devlin et al.*, 2011]. For each indicator, final measurements are converted into a normalized ecological quality ratio by first converting the data into a numerical scale between zero and one (where status class boundaries are not necessarily equidistant) and then averaging the scores for all indicators and related to one of the five assessment classes. Classification of overall ecological condition status is determined using a one-out-all-out approach: where the overall status reflects the worst category from results for any biological quality element or physico-chemical element [*Devlin et al.*, 2011]. In this review we focus specifically on the phytoplankton biological quality element and the nutrient physico-chemical element. Here we review the nutrient physico-chemical element and the phytoplankton biological quality element. The sampling period for all elements is a minimum of six years, with sampling frequency no less than 12 times per year, collected monthly [*Devlin et al.*, 2007b].

UK WFD Nutrients Water Quality Element. Nutrient thresholds for the UK WFD assessment framework are generated using a tool based on a cause and effect model that relates elevated nutrients indices of ecosystem response [*Devlin et al.*, 2007a]. The tool specifically looks at three indices: (1) Evidence of nutrient enrichment based on the calculation of an annual winter nitrogen concentration; (2) Modeling of potential primary production based on a waterbody characteristics and light availability; (3) Evidence of undesirable disturbance as measured by dissolved oxygen levels. A stepwise analysis scheme is employed to determine overall eutrophic condition. Initial classification of the water bodies is based on comparison of mean winter dissolved inorganic nitrogen concentration against predetermined nutrient thresholds. Winter is defined as the period when algal activity is lowest and when dissolved nutrients should show conservative behavior [*Devlin et al.*, 2007a]. Nutrient thresholds are also normalized to a salinity gradient, allowing for dilution of nutrients with increasing salinity. If estuaries exceed the initial thresholds for "Good" water quality, potential primary productivity is estimated from a simple screening model that uses equilibrium nutrient concentrations and light limited growth rates to calculate production [*Devlin et al.*, 2007a; *S Painting et al.*, 2006]. If the potential primary production is greater than 300 g C m^{-2} y^{-1}, a level defined by Nixon [1995] as representing eutrophic status, and winter dissolved inorganic nitrogen concentration is greater than 30 µM, than the estuary is considered to have moderate or worse eutrophic condition. The final metric, used to determine the severity of adverse impacts, is dissolved oxygen concentration. Dissolved oxygen concentration is reported as either a growing season mean (March to September). Thresholds for dissolved oxygen that mark the boundaries between Moderate and Poor and Poor and Bad are derived from criteria set for fish in transitional waters which supports conditions for juvenile fish in the freshwater reaches of estuaries [*Best et al.*, 2007]. Dissolved oxygen concentrations less than 5 mg L^{-1} negatively affect sensitive species of fish and invertebrates and is, thus, the boundary between moderate and poor. Dissolved oxygen levels below 2.5 mg L^{-1} negative impact most fish species and is thus the boundary between poor and bad condition. Overall condition is based on the combination of the three indices and is summarized in Table 3.6.

Table 3.6. UK WFD classification based on deviation from reference conditions. Classification is assessed via progression through the three indices [Devlin et al., 2007a]. Bold line indicators management action point.

		Index 1: Nutrient Concentration	Index 2: Production	Index 3: Undesirable Disturbance
Statistic for Index		**Mean Winter DIN (µM)**	**Growing Season Potential Primary Productivity**	**Growing Season Mean Dissolved Oxygen Concentration**
Units		□M	g C m^{-2} y^{-1}	mg L^{-1}
Index		I_{DIN}	I_{PP}	I_{DO}
Classification	**High**	$I_{DIN} \leq 12$	n/a	n/a
	Good	$I_{DIN} \leq 18$	n/a	n/a
	Good		$I_{DIN} \geq 30$ **µM** $I_{PP} < 300$	$I_{DO} > 5$
	Moderate		$I_{DIN} \geq 30$ **µM** $I_{PP} \geq 300$	$I_{DO} > 5$
	Poor		$I_{DIN} \geq 30$ **µM** $I_{PP} \geq 300$	$I_{DO} \leq 5$
			$I_{DIN} \geq 30$ **µM** $I_{PP} \geq 300$	$I_{DO} \leq 2$

UK WFD Phytoplankton Biological Quality Element . There are three indicators proposed for the phytoplankton biological quality element of the UK WFD for coastal waters: 1) phytoplankton biomass measure as CHL-a, 2) the frequency of elevated phytoplankton counts measuring individual species and total cell counts, and 3) seasonal progression of phytoplankton functional groups through the year [*Devlin et al.*, 2007b]. The first index, phytoplankton biomass as CHL-a (I_{CHL}), is defined as the 90th percentile of chlorophyll concentrations during the growing season (March to September). The boundary conditions are different by salinity strata. For marine waters, the reference value is proposed as 10 µg L^{-1} (implying 50% elevation of the background value of 6.7 µg L^{-1} and a reasonable C:Chl factor of 0.012). For low salinity waters, where the level of production may be expected to be higher, a reference value of 15 µg L^{-1} is proposed (implying a background value of 10 µg L^{-1} chlorophyll and a C:Chl factor of 0.02; Table. 3.X)[*Devlin et al.*, 2007b].

Table 3.7 Thresholds for concentrations of chl a, dissolved oxygen and dissolved inorganic nitrogen for the UK WFD assessment method. From Devlin et al. 2011.

<table>
<tr><th>Method</th><th>Chl-a reference thresholds (μg l^{-1})</th><th>DO reference thresholds (mg l^{-1})</th><th>DIN reference thresholds</th><th>Source and criteria</th></tr>
<tr>
<td>UK WFD</td>
<td>BIOMASS INDICATOR:
In the UK, five statistical measures of Chl-a biomass are made over two salinity bands using combined data for a six year reporting cycle. Compliance with the threshold is given a score of 1 for each statistical measurement, with an optimum score of 10. The final score (a value between 0 [bad] to 10 [best]) is normalized to an equidistant EQR score (0 – 1).

SALINITY RANGE – LOW (0 – 25ppt)
<u>No Chl-a measurement Threshold</u>
1: Average Chl-a conc, ≤15μg l^{-1}
2: Median Chl-<i>a</i> conc, ≤12μg l^{-1}
3: % Chl-<i>a</i> less than 10 μg l^{-1} >70%
4: % Chl-a less than 20 μg l^{-1} >80%
5: % Chl-a greater than 50 μg l^{-1} <5%

SALINITY RANGE – HIGH (>25 ppt)
<u>No Chl-a measurement Threshold</u>
6: Average chl-a conc, ≤10μg l^{-1}
7: Median chl-<i>a</i> conc, ≤8μg l^{-1}
8: % Chl-<i>a</i> less than 10 μg l^{-1} >75%
9: % Chl-a less than 20 μg l^{-1} >85%
10: % Chl-a greater than 50 μg l^{-1} <5%

TOTAL SCORE EQR STATUS CLASS
0 - 2 0.000 - 0.133 Bad
3 - 4 0.200 - 0.300 Poor
5 - 6 0.400 - 0.500 Moderate
7 - 8 0.600 - 0.700 Good
9 - 10 0.800 - 1.000 High

TAXA ABUNDANCE INDICATOR:
Cell counts for each sampling period are used to calculate the number of times the threshold is exceeded (as %) when:
1.Any Single taxon (species) >500,000 cell l^{-1}
2.Total Abundance > 10^{6} cells l^{-1}

<table>
<tr><th>% exceedances</th><th>Normalised score (= ref/value)</th><th>Final EQR</th></tr>
<tr><td>0-10</td><td>1.0--0.5</td><td>0.8 - 1.0</td></tr>
<tr><td>10–20</td><td>0.5--0.25</td><td>0.6 - 0.8</td></tr>
<tr><td>20–40</td><td>0.25--0.13</td><td>0.4 - 0.6</td></tr>
<tr><td>40–60</td><td>0.13--0.08</td><td>0.2-0.4</td></tr>
<tr><td>60–100</td><td>0.08--0.0</td><td>0 - 0.2</td></tr>
</table>

Normalised score calculated by reference condition (5%) divided by value. Final EQR normalised to equidistant boundaries (0 to 1)</td>
<td>Annual 5th percentiles:
>5.7 mg l^{-1} = High
4.0 <5.7 mg l $^{-1}$ = Good
2.4<4.0 mg l $^{-1}$ = Mod
1.6 <2.4 mg l $^{-1}$ = Poor
<1.6 mg l^{-1} = Bad

EQR
High: 0.8 – 1.0
Good: 0.6 – 0.8
Mod: 0.4 – 0.6
Poor: 0.2 – 0.4
Bad: 0.0 – 0.2

For DO, 5th percentile of all data (collected monthly). Reporting period is typically over 6 years.</td>
<td>Winter DIN thresholds for UK estuaries:
(clear estuaries)
as μM: as mg l^{-1}:
<20 = High <0.28 = High
<30 = Good <0.42 = Good
<45 = Mod <0.63 = Mod
<67 = Poor <0.94 = Poor
>67 = Bad >0.94 = Bad

These thresholds are for clear waters only, defined by mean annual SPM (SPM < 10 mg l^{-1})
In turbid waters (>10 mg l^{-1} SPM), a secondary threshold (70 μM, 0.98 mg l^{-1}) may be applied, where
<70 μM (0.98 mg l^{-1}) = Good
>70 μM (0.98 mg l^{-1}) = Moderate

If the secondary threshold is applied, then 99th percentiles are calculated from the data and compared to the threshold.</td>
<td>Chl-a
Devlin et al., 2007a, 2007b,
www.ukwfd.org

DO
Best et al., 2007

DIN
Devlin et al., 2007a
www.ukwfd.org.au</td>
</tr>
</table>

Table 3.7 continued

Method	Chl-*a* reference thresholds ($\mu g\ l^{-1}$)	DO reference thresholds ($mg\ l^{-1}$)	DIN reference thresholds	Source and criteria
ASSETS	*Annual 90th percentile*: 0-5 $\mu g\ l^{-1}$ = Low 5-20 $\mu g\ l^{-1}$ = Moderate >20 $\mu g\ l^{-1}$ = High >60 $\mu g\ l^{-1}$ = Hypereutrophic *Spatial coverage of worst case conditions:* 0-10% = Very Low 10 – 25% = Low 25% - 50% = Moderate >50% = High *Frequency of occurrence of worst case conditions*: Persistent Periodic Episodic	*Annual 10th percentile*: 0 $mg\ l^{-1}$ = Anoxia 0-2 $mg\ l^{-1}$ = Hypoxia 2-5 $mg\ l^{-1}$ = Biologically Stressful *Spatial coverage of worst case conditions:* 0-10% = Very Low 10 – 25% = Low 25% - 50% = Moderate >50% = High *Frequency of occurrence of worst case conditions*: Persistent Periodic Episodic	ASSETS does not use nutrient concentrations in the assessment formulation, only nutrient loads	Bricker et al. 1999, 2003, 2007 **Chl-*a*** 90th percentile of annual data is used. These thresholds are used for all systems except Florida Bay for which thresholds are lower (i.e. High [worst] for Florida waters = 2 – 5 $\mu g\ l^{-1}$) **DO** 10th percentile of annual data is used
EPA NCA	*Summer value*: 0-5 $\mu g\ l^{-1}$ = Good 5-20 $\mu g\ l^{-1}$ = Fair >20 $\mu g\ l^{-1}$ = Poor	*Summer value*: >5 $mg\ l^{-1}$ = Good 2-5 $mg\ l^{-1}$ = Fair <2 $mg\ l^{-1}$ = Poor	*Summer value:* as µM: / as $mg\ l^{-1}$: <7 = Good / <0.1 = Good 7-36 = Fair / 0.1-0.5 = Fair >36 = Poor / >0.5 = Poor	EPA 2001, 2005, 2008 Data from the summer index period are used for determination of Chl-*a* and DIN condition at individual sites of US East, Gulf and West coast systems. Reference conditions are different for sensitive waterbodies
OSPAR COMPP	*Growing season 90th percentile:* Threshold = 15 $\mu g\ l^{-1}$ >15 $\mu g\ l^{-1}$ = threshold exceeded indicating a Problem Area Maximum and mean concentrations may also be compared to this threshold.	*5th percentile of growing season data:* Threshold = 4 $mg\ l^{-1}$ <4 $mg\ l^{-1}$ = threshold exceeded indicating a Problem Area	*Winter DIN for UK estuaries*: Threshold = 30 µM (0.42 $mg\ l^{-1}$) >30 µM = threshold exceeded indicating a Problem Area N:P Ratio Threshold: 24:1 where >24:1 is indicative of a Problem Area	OSPAR 2005, 2008 A one-out-all-out procedure is used to determine the classification of each of the four Categories and for the Overall Assessment

Thresholds are not used in the TRIX approach. For DIN, the WFD (as applied within the UK) and OSPAR use µM units, while EPA NCA uses $mg\ l^{-1}$. Thresholds are given here in both units to enable comparisons among methods

Mod moderate, *SPM* suspended particulate matter

The second index, elevated phytoplankton abundance (I_E), assesses the presence, abundance and frequency of occurrence of elevated counts of algal species relative to undisturbed conditions. This index is based on three attributes, one which is a measure of the frequency that elevated biomass (CHL) exceeds a reference threshold and three of which focus on counts of algae that may result in the decline of ecosystem health in an undesirable disturbance (Table 3.8) [*Devlin et al.*, 2007b]. Each attribute is calculated from the number of times it exceeds the threshold as a proportion of the total number of sampling times per year, and is recorded as a six year mean. The proposed thresholds are for three groups of phytoplankton and for counts of chlorophyll exceeding a threshold. The first phytoplankton threshold identifies any species of phytoplankton, excluding *Phaeocystis* species, that exceed counts of 10^6 cells L^{-1} [S], the second phytoplankton threshold identifies *Phaeocystis sp*. that exceed counts of 10^6 cells L^{-1} [P], and the third threshold identifies where the total taxa counts exceeds counts of 10^7 cells L^{-1} [T]. The chlorophyll count within this index identifies any chlorophyll measurement that exceeds 10 µg L^{-1}. The final index is calculated as the sum of these attributes: $I_E = \Sigma\ (CHL + S + P + T)$.

Table 3.8 Proposed boundary conditions for phytoplankton abundance relating to occurrences of elevated taxa counts over a six year period. From Devlin et al. 2007b.

Normative definition	Index	Equation – {sum $[T]+[P]+[S]+[\text{chl}]\ldots/4\} * 100$	Classification boundaries	
Phytoplankton abundance	I_E	I_E: Sum of the occurrence of any single species ($>10^6$) plus *Phaeocystis* sp. ($>10^6$), plus total cell counts ($>10^7$) and counts of chlorophyll >10 µg l^{-1} over a six year period	High Good Moderate Poor Bad	<15% <30% <40% <50% >50%

This index is composed of counts of four attributes within the tool. Samples are taken in growing season between April and September.

The third index, seasonal succession of functional groups (I_F), represents the deviation of the natural progression of dominant functional groups throughout the seasonal cycle relative to undisturbed conditions. Counts of four major functional groups, including diatoms, dinoflagellates, microflagellates (excluding *Phaeocystis*) and *Phaeocystis sp*. are averaged for each month over a sampling year, and are normalized and reported as a monthly Z score. Monthly Z scores for each functional group are compared to a specific reference curve for different classes of waterbodies. A final score is based on the number of data points from the test waterbody which fell within the standard deviation range set for each monthly point of the reference growth curve [*Devlin et al.*, 2007b].

Trophic Index (TRIX)

TRIX integrates oxygen saturation, phytoplankton chlorophyll-a, nitrogen and phosphorus concentrations to assess the trophic state of coastal marine waters and lagoons [*Giovanardi and Vollenweider*, 2004; *Vollenweider et al.*, 1998]. TRIX is based on the assumption that eutrophication processes are mainly reflected by changes in the phytoplankton community, which is typically only true for coastal waters and estuaries dominated by deep subtidal habitat. It was developed for use in Italian coastal waters and lagoons. The index is given by equation 1:

Equation 1 $$\text{TRIX} = [\log_{10}(\text{CHLa} * \%\text{DO} * \text{N} * \text{P}) + 1.5] / 1.2$$

where CHLa is the chlorophyll-a concentration (µg L^{-1}), %DO is dissolved oxygen represented as the absolute percent deviation from saturation (%), N is the concentration of dissolved inorganic nitrogen (ammonia + nitrate + nitrite) in µg-at L^{-1}, P is the concentration of dissolved inorganic phosphorus as phosphate (µg-at L^{-1}). The TRIX score is scaled from 0 to 10, covering a range of four trophic states (0-4 high quality and low trophic level; 4-5 good quality and moderate trophic level; 5-6 moderate quality and high trophic level and 6-10 degraded and very high trophic level).

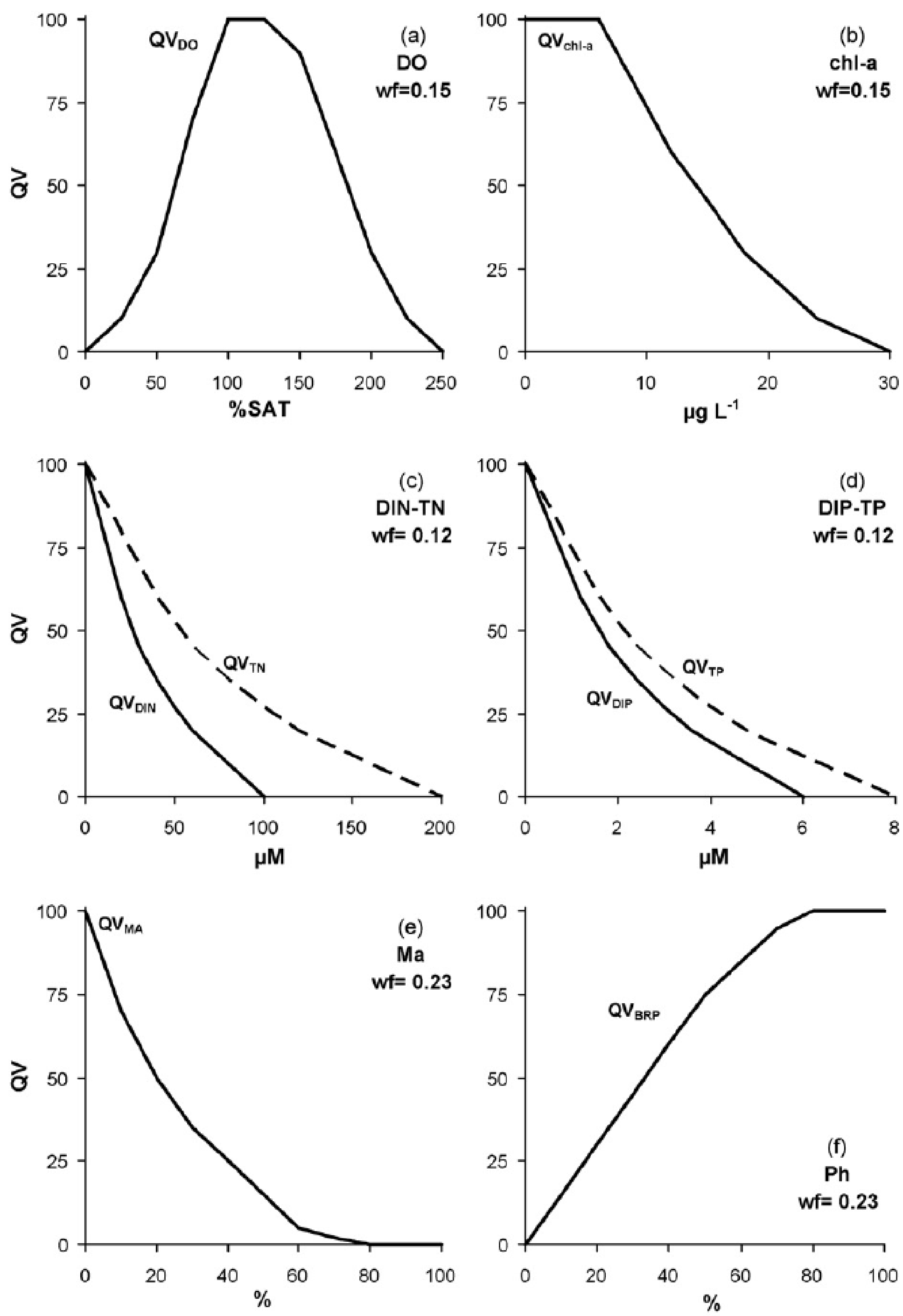


Figure 3.1. Relationships among analytical measurements of (a) dissolved oxygen saturation (DO), (b) chlorophyll-a (Chl-a), (c) dissolved inorganic and total nitrogen (DIN-TN), (d) dissolved inorganic and total phosphorus (DIP-TP), (e) macroalgal coverage (Ma), (f) phanerogam coverage (Ph) and respective Q values (QV). wf: weighting factors used in TWQI calculation[Giordani et al., 2009].

Assessment of Estuarine Trophic Status (ASSETS)

ASSETS is an integrated methodology used to comparatively rank the eutrophication status of estuaries and coastal areas. It was developed for use in the U.S. National Estuarine Eutrophication Assessment (NEEA), but has been extended and refined for use in other estuarine systems around the world. The methodology is described in detail elsewhere [*Bricker et al.*, 2003; *Bricker et al.*, 1999].

The ASSETS assessment includes three diagnostic tools: an assessment of pressure (influencing factors [IF]), an evaluation of state (eutrophic condition [EC]), and the expected response (future outlook [FO])[*Bricker et al.*, 2003; *Bricker et al.*, 1999; *Devlin et al.*, 2011; *Garmendia et al.*, 2012]. The IF assessment is based on two factors: the nutrient loading (input) from the watershed and/ or ocean and the susceptibility of the system (capability of the system to dilute or flush the nutrient inputs). The overall IF falls into one of five categories (low, moderate-low, moderate, moderate-high, and high) that are determined by a matrix that combines susceptibility and load factors. The EC is evaluated based on a combination of primary and secondary symptoms of eutrophication sampled monthly. The two primary symptoms are phytoplankton (evaluated as CHL-a concentration, frequency, and spatial coverage) and macroalgae (magnitude and frequency of “problem status,” where “problem” indicates a detrimental impact on any biological resource). The three secondary symptoms are bottom water dissolved oxygen (concentration, spatial coverage, and frequency of low events), nuisance and toxic blooms (duration and frequency of “problem status”), and submerged aquatic vegetation (SAV) (“problem status or change in spatial coverage” and the magnitude of the change)[*Bricker et al.*, 1999; *Garmendia et al.*, 2012]. The EC rating is determined by a matrix that combines the average score of the primary symptoms (chlorophyll “a” and macroalgae) and the highest score (worst impact) of the secondary symptoms (dissolved oxygen, nuisance and toxic blooms and SAV) and categorizes estuaries into one of five categories (low, moderate-low, moderate, moderate-high, and high). The FO rating, is determined by a matrix that combines the susceptibility and expected change in loading factors and classifies estuaries into one of the five categories (worsen-high, worsen-low, no change, improve-low, and improve-high). The assessment then combines results of the three components into a single overall rating of bad, poor, moderate, good, and high trophic status using a matrix approach [*Bricker et al.*, 2003; *Bricker et al.*, 1999; *Devlin et al.*, 2011; *Garmendia et al.*, 2012]. Thresholds for each indicator are given in Table 3.3.

Table 3.9. Indicators and thresholds applied in the ASSETS framework [Bricker et al., 2003].

	Index	Indicator	Statistic for Index	Thresholds and Ranges
Primary Symptoms	**Phytoplankton**	CHL-a	90^{th} percentile of monthly data	Hypereutrophic: $> 60\ \mu g\ L^{-1}$ High: $> 20\ \mu g\ L^{-1}$ but $\leq 60\ \mu g\ L^{-1}$ Medium: $> 5\ \mu g\ L^{-1}$ but $\leq 20\ \mu g\ L^{-1}$ Low: $\leq 5\ \mu g\ L^{-1}$
		Spatial Coverage	Heuristic of Monthly Data	High, Moderate, Low, or Very Low
		Frequency		Periodic, Episodic, or Persistent
	Macroalgae or Epiphytes	Biomass and Cover	Heuristic of Monthly Data	Problem: detrimental impact to biological resources No Problem: no apparent impact on biological resources
		Spatial Coverage		High, Moderate, Low, or Very Low
		Frequency		Periodic, Episodic, or Persistent
Secondary Symptoms	**Dissolved Oxygen**	Bottom water Concentration	10^{th} percentile of monthly data	Anoxia: $0\ mg\ L^{-1}$ Hypoxia: $> 0\ mg\ L^{-1}$ but $\leq 2\ mg\ L^{-1}$ Biologically Stressful: $> 2\ mg\ L^{-1}$ but $\leq 5\ mg\ L^{-1}$
		Spatial Coverage	Heuristic of Monthly Data	High, Moderate, Low, or Very Low
		Frequency		Periodic, Episodic, or Persistent
	SAV Loss	Magnitude of Loss	Analysis of Monthly Data	High Loss: ≥ 50 but ≤ 100 % of estuarine surface water area Medium Loss: ≥ 25 but > 50% of estuarine surface water area Low: ≥ 10 but > 25% of estuarine surface water area Very Low: ≥ 0 but > 10% of estuarine surface water area
	Nuisance and Toxic Blooms	Observed Occurrence	Cell Counts of Dominant Species	Problem: detrimental impact to biological resources No Problem: no apparent impact on biological resources
		Duration	Monthly Data	Hours, Days, Weeks, Seasonal, Other
		Frequency	Heuristic of Monthly Data	Periodic, Episodic, or Persistent

OSPAR

OSPAR is the mechanism by which fifteen Governments of the western coasts and catchments of Europe, together with the European Community, cooperate to protect the marine environment of the North-East Atlantic. The OSPAR Eutrophication Strategy sets the objective to combat eutrophication in the OSPAR maritime area. The OSPAR Common Procedure is used to identify the eutrophication status and assess compliance with the Ecological Quality Objectives (EcoQO) for eutrophication for the North Sea (www.OSPAR.org).

The specific Ecological Quality Objectives for eutrophication agreed at the 5th North Sea Conference (Bergen Declaration 2002) are (OSPAR 2005):

- Winter DIN and/or DIP should remain below elevated levels, defined as concentration > 50% above salinity related and/or region-specific natural background concentrations;
- Maximum and mean region-specific chlorophyll a concentrations during the growing season should remain below region-specific elevated levels, defined as concentrations > 50% above the spatial (offshore) and/or historical background concentration;
- Region/area-specific phytoplankton eutrophication indicator species should remain below respective nuisance and/or toxic elevated levels (and increased duration);
- Oxygen concentration, decreased as an indirect effect of nutrient enrichment, should remain above region specific oxygen deficiency levels, ranging from 4-6 mg oxygen per litre;
- There should be no kills in benthic animal species as a result of oxygen deficiency and/or nuisance/toxic phytoplankton indicator species for eutrophication.

Under OSPAR (2005), nutrient concentrations are assessed by plotting the winter nutrient concentrations of each year in relation to the respective measured salinity values ("mixing diagrams"). In winter, defined as period when algal activity is lowest, DIN and DIP show a conservative behavior and, therefore, a good linear relationship with salinity (decreasing concentration with increasing salinity from coast to offshore). The salinity normalized nutrient concentration (with 95% confidence interval) is plotted in relation to the respective year in order to establish trends in the winter nutrient concentrations and the level of elevation (compared with background concentration).

In determining the maximum and mean chlorophyll a levels in estuaries, chlorophyll a concentrations are averaged over the salinity range during the growing season. Table 3.10 gives the area-specific natural background and elevated concentrations of chl-a.

Table 3.10 Area specific background concentrations and elevated nutrient concentrations of chlorophyll a during growing season in relation to salinity. From OSPAR 2005.

	Region	Salinity	Background concentration	Background concentration	Elevated levels
			Chlorophyll *a* µg/l, means	Chlorophyll *a* µg/l, maxima	Chlorophyll *a* µg/l, means
North Sea	Belgium		10	15	
	Coast		10		>15
	Denmark	>34.5	2-4		>4.5
	Coast	<34.5	2-10		
	Germany	>34.5	2	10-13	3
	Coast	<34.5	2-4	13-18	3-6
	Netherlands	>34.5	2-4		>4.5
	Coast	<34.5	10	10	>15
	Norway		2-4		>4.5
	Coast		2-10		
	UK	>34.5	5-10	10	>10
	Coast	<34.5	8-12	15	>20
Channel	France	>34	2	10	> 4
Wadden Sea	Denmark	<30			>22-24 (needs verification)
	Germany	29-32	2-4	12-20	3-6
	Netherlands	<30		16	>22-24 (needs verification)
Skagerrak	Denmark	32-34	<1.25		
	Norway	33			
	Sweden		1.5		>2
Kattegat	Denmark		1.5		>2
	Sweden		1.5	1.5	>2
Atlantic	France		2	10	>4
	Ireland coast	>34.5	<7		>10
	Norway				
	Portugal				
	Spain	>34.5			>12
	coast			8	
	UK/Scotland	>34.5	5	10	>10
	coast	<34.5	10	15	>15
Southern Irish Sea and Eastern Celtic Sea	Ireland Offshore	>34.8			
Atlantic to Irish Sea	Coast	>34.5	<7		>10
Estuaries	Belgium				

Table. 3.10 Continued

	Region	Salinity	Background concentration	Background concentration	Elevated levels
			Chlorophyll *a* µg/l, means	Chlorophyll *a* µg/l, maxima	Chlorophyll *a* µg/l, means
	Denmark				
	France			13 (variable at sal <30)	>18-20(variable at sal < 30)
	Germany	0-30	5-8	12-40	7.5-12
	Ireland				
	Netherlands	<30		2-6	
	Western Scheldt				>9-10
	Ems Dollar				>18-20
	Norway				
	Portugal:				
	Sado		10?		>9
	Tagus				>14
	Mondego		10?		>9
	Spain				
	Sweden				
	UK				

OSPAR distinguishes two types of phytoplankton indicator species: nuisance species (forming dense "blooms") and toxic species (already toxic at low cell concentrations). Examples of levels considered as elevated levels and their effects are provided in Table 3.11. Use of nuisance and toxic blooms has not seen wide-spread use because of uncertainty in linkage to anthropogenic nutrients.

Table 3.11 Elevated levels of area-specific nuisance and toxic phytoplankton indicator species and the types of their effects. From OSPAR 2005.

Phytoplankton indicator species	Elevated levels	Effects
Nuisance species		
Phaeocystis spp. (colony form)	> 10^6 cells/l (and >30 days duration)	nuisance, foam, oxygen deficiency
Noctiluca scintillans	> 10^4 cells/l (area coverage > 5 km^2)	nuisance, oxygen deficiency
Toxic (toxin producing) species		
Chrysochromulina polylepis	> 10^6 cells/l	toxic; fish and benthos kills
Gymnodinium mikimotoi	> 10^5 cells/l	toxic; fish kills, PSP mussel infection
Alexandrium spp.	> 10^2 cells/l	toxic; PSP mussel infection
Dinophysis spp.	> 10^2 cells/l	toxic; DSP mussel infection
Prorocentrum spp.	> 10^4 cells/l	toxic; DSP mussel infection

HELCOM Eutrophication Assessment Tool (HEAT)

HEAT is a multi-metric indicator-based tool for assessment of eutrophication status [*HELCOM*, 2009]. HEAT has been developed specifically for the HELCOM Integrated Thematic Assessment of Eutrophication in the Baltic Sea. Ecological objectives related to eutrophication were adopted in the HELCOM Baltic Sea Action Plan. They are: concentrations of nutrients close to natural levels, clear water, natural level of algal blooms, natural distribution and occurrence of plants and animals, and natural oxygen levels [*HELCOM*, 2009]. HEAT is an indicator based assessment framework which groups indicators as follows: (1) physical- chemical features (PC), (2) phytoplankton (PP), (3) submerged aquatic vegetation (SAV), and (4) benthic invertebrate communities (BIC). Groups 1 and 2 (PC and PP) are considered 'primary signals' of eutrophication, while groups 3 and 4 (SAV and BIC) are considered 'secondary signals' [*HELCOM*, 2009]. For each indicator a eutrophication quality objective (EutroQO) or target is calculated from the reference condition (RefCon) and the acceptable deviation (AcDev) from reference condition. When the actual status (AcStat) exceed the EutroQO, the area in question is regarded as 'affected by eutrophication'' or falling below the "good-moderate" threshold [*Andersen et al.*, 2011].

Reference Conditions (RefCon), are the biological quality elements that exist, or would exist, with no or very minor disturbance from human activities. They should represent the continuum that is naturally present and must reflect variability. The HEAT tool uses three principles for setting RefCons: (1) reference sites, (2) historical data, and (3) modeling. Expert judgment can also be used as a supplement. RefCons as applied in the Baltic sea were typically basin specific and varied by an order of magnitude over the salinity gradient of the sea.

The acceptable deviation (AcDev) values are basin specific. Two different principles were used for setting the AcDev, according to whether indicators show a positive response (increasing in value) to increases in nutrient inputs or a negative response (decreasing in value). For an indicator showing positive response (e.g. nutrient concentrations and chlorophyll-a), AcDev has an upper limit of +50% deviation from RefCon [*HELCOM*, 2009]. Setting AcDev to 50% implies that low levels of disturbance (defined as less than +50% deviation) resulting from human activity are considered acceptable while moderate (greater than +50%) deviations are unacceptable (boundary between good and moderate in the WFD) [*Andersen et al.*, 2011]. For indicators responding negatively to increases in nutrient input (e.g. Secchi depth and depth limit of SAV) the AcDev's have in principle a limit of -25% [*HELCOM*, 2009], although AcDev's used for benthic invertebrates are slightly greater in magnitude, ranging from -27 to -40% [*HELCOM*, 2009]. Whereas an indicator with positive response can theoretically show unlimited deviation, indicators showing negative response have a maximum deviation of -100% and a deviation of -25% is, in most cases, interpreted as the boundary between good and moderate in the WFD [*Andersen et al.*, 2011].

Each site is assigned an ecological condition category as set up by the WFD: high (best condition), good, moderate, poor, and bad (worst condition) [*HELCOM*, 2009]. To assign a category, an Ecological Quality Ratio (EQR) is calculated for each site based on the RefCon and

AcStat. The boundary between good and moderate status is where the deviation from RefCon is equal to the AcDev. All other categories are assigned based on a defined deviation of the AcStat from RefCon [*Andersen et al.*, 2011]. An EQR value and a set of class boundaries are calculated for each indicator, but the overall status classification depends on a combination of indicators. First, indicator EQR values are combined to give an EQR value for a specific Quality Element (QE), and similarly the indicator class boundaries are combined to give the class boundaries for the QE. In the simplest case, where all indicators within a QE have equal weights, the EQR for the QE is the average of the indicators' EQRs within the QE and each QE class boundary (e.g. Moderate/Good boundary) is found as the average of the class boundary values for all indicators representing that specific QE. Within a QE, it is also possible to assign weighting factors to indicators according to expert judgment. The classification of the QE is then given by comparison of the weighted averages of the EQRs with the weighted averages of the individual class boundaries. Thus, the same weighting is applied both in calculation of the EQR for the specific QE as well as QE class boundary values. The lowest rated of the QEs will because of the 'One out—all out' principle determine to final status classification [*Andersen et al.*, 2011].

Transitional Water Quality Index (TWQI)

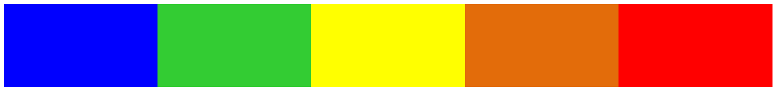
The TWQI was developed to assess trophic status and water quality in transitional (i.e. estuarine) aquatic ecosystems of Southern Europe [*Giordani et al.*, 2009]. It was developed specifically for shallower estuarine systems, where benthic vegetation controls primary productivity, making phytoplankton only indices unsuitable. The index was based on the water quality index of the U.S. National Sanitation Foundation and integrates the main causal factors (inorganic nutrients), key biological elements (primary producers) and indicator effects (dissolved oxygen). The TWQI utilizes six main variables: relative coverage of seagrass and opportunistic macroalgae species, concentration of dissolved oxygen, phytoplankton chlorophyll-a, dissolved inorganic nitrogen and phosphorus. Non-linear functions are used to transform each measured variable into a Quality Value (QV) (Figure 3.1.) [*Giordani et al.*, 2009]. Each quantity is then multiplied by a weighting factor to account for the relative contribution of each variable to the overall water quality (adding up to a total percentage of 100): dissolved oxygen = 15%, CHL-a = 15%, DIN-TN = 12%, DIP-TP = 12%, macroalgal coverage = 23%, seagrass coverage = 23%. The QV_{DO} for dissolved oxygen follows a bell shaped curve where the QV increases from 0 to 100 from dissolved oxygen levels of 0 percent saturation to 125 % saturation and decreases again from 100 to 0 as DO saturation increases from 125% to 250% (saturations over 125% are often associated with blooms in primary producer groups). The QV_{CHLa} is zero (worst condition) when concentrations of CHL-a are greater than 30 mg m^{-3} and 100 (best condition) when CHL-a concentrations are less than 6 mg m^{-3}. The QV_{DIN} is inversely related to DIN concentrations where QV_{DIN} is 100 when DIN is 0 µM and QV_{DIN} is 0 when DIN is greater than 100 µM. The most significant decrease in QV_{DIN} is imposed at the 0-20 µM range because the main transformation in primary production was found to occur in this range [*Viaroli et al.*, 2008], and it was found to be a critical threshold for other lagoons (see Souchu et al. 2000). The QV_{DIP} was set up similar to QV_{DIN} where QV_{DIP} is 100 when DIP is 0 µM and QV_{DIP} is 0 when DIP is greater than 6 µM. The QV_{Ph} and QV_{Ma} are based on the percent of estuarine surface area colonized. The QV_{Ma} is zero (worst condition) when macroalgae percent cover exceeded 80% of estuarine surface area and 100 (best condition) when macroalgae percent cover was less than 10%. The utility function for seagrass was opposite to macroalgae such that QV_{Ph} is zero (worst condition) when seagrass percent cover was less than 10% of estuarine surface area and 100 (best condition) when seagrass percent cover was greater than 80%. An index value is calculated as the sum of the weighted quality values, ranging from 0 (poorest) to 100 (best condition). The index has been tested and validated in several estuarine systems that differ in anthropogenic pressures and eutrophication levels.

The French Research Institute for the Exploration of the Sea (IFREMER) Classification for Mediterranean Lagoons

The IFREMER developed a classification scheme for benthically-dominated French Mediterranean lagoons [*Souchu et al.*, 2000; *Zaldivar et al.*, 2008], which is based on several physical, chemical and biological potential indicators of eutrophication in the various components of the lagoon ecosystem: benthic, phytoplankton, macrophytes, macrofauna, sediments and water. It allows for the classification of a lagoon into five eutrophication levels formalized by five different colors from blue (no eutrophication), green, yellow, orange, and red (high eutrophication), similar to the color scheme used by the Water Framework Directive (WFD). Overall classification is based on the worst partial value of the elements listed above. Each component of the ecosystem is assessed independently allowing for identification of which component is experience degradation. Indicators are scored against thresholds based on an annual average of the data. Elements and thresholds used to assess the water column are presented in Table 3.12. Thresholds are based on an annual average of data collected.

Table 3.12. Water quality elements and thresholds measured in the IFREMER assessment framework for French Mediterranean lagoons. Eutrophication is scored from blue (no eutrophication) to red (high eutrophication) [Souchu et al., 2000; Zaldivar et al., 2008].

Condition Category



U.S. EPA's National Coastal Assessment

The US EPA's National Coastal Assessment (NCA) is implemented through a federal—state partnership, and is designed to answer questions on environmental conditions in coastal waterbodies at a regional – national scale. The results supplement the US Clean Water Act (CWA) where waterbodies identified as not meeting state water quality criteria for designated uses require actions to correct pollution caused impairments [*USEPA*, 2001; 2005; 2008]. Of the five EPA NCA indices of condition in coastal waterbodies, the Water Quality Index (WQI) is the indicator describing nutrient related conditions and will be the only one reviewed here. This method uses five indicators: dissolved inorganic nitrogen (DIN), dissolved inorganic phosphorus (DIP), Chl-a, water clarity (by Secchi depth and by comparison of light reaching the water surface and at 1 m depth) and dissolved oxygen. The WQI uses the EPA Environmental Monitoring and Assessment Program's (EMAP) probabilistic randomly selected sampling framework where samples are taken once per year (per station) by region during a summer index period (June through September; [*USEPA*, 2001]). An evaluation is made for each of the five indicators at each site by comparison with regionally defined reference conditions and a combined water quality index rating is calculated for each site, then for the region and the nation based on the ratio of individual indicators that are rated as Good, Fair or Poor [*Devlin et al.*, 2011]. Thresholds for each indicator are based on assumed reference conditions, are given in Table 3.13.

An indicator is considered Good if less than 10% of samples are Poor and 50% are Good; condition is fair if 10–25% of samples are Poor and/or 50% are Poor or Fair; and condition is Poor if more than 25% of samples are Poor. All indicators are combined in a similar fashion to determine the rating for a site: where Good is a maximum of one indicator is Fair and no indicators are Poor; Fair is one of the indicators is rated Poor or two or more indicators are Fair; and Poor is two or more of the five indicators are rated Poor. To determine the WQI by region and nation, results from each area are used to determine a final assessment score where: Good is less than 10% of areas are in Poor condition and more than 50% are Poor or Fair; Fair is 10–20% of areas are in Poor condition or greater than 50% are Fair or Poor; and Poor if greater than 20% of areas are in Poor condition.

Table 3.13. Thresholds for each indicator used in the US EPA NCA [Devlin et al., 2011].

Classification

Indicator Specific Assessment Frameworks-Phytoplankton Index of Biotic Integrity

One use of phytoplankton community structure data is to combine it into an index of biological integrity (IBI). IBIs are becoming more common for assessment of estuarine ecological condition and management focus in the face of physical and chemical transformation, habitat destruction, and changes in biodiversity (Borja et al. 2008). An IBI describes the biological condition of an assemblage of plants or animals, typically based on the diversity and relative abundance of species or the presence or absence of pollution tolerant species. A key element of developing an IBI is the ability to describe the community response of the assemblage (e.g., benthic invertebrates, phytoplankton, etc.) along gradient of physical or chemical stress from minimally disturbed or “reference state” to highly disturbed.

IBIs developed and used in Chesapeake Bay present an example of how phytoplankton community structure data can be synthesized to provide information about the ecological health of the Estuary and about the ability to support specific beneficial uses. A Phytoplankton Index of Biotic Integrity (P-IBI) was developed in Chesapeake Bay using an 18 year data set (Lacouture et al. 2006). The P-IBI combined the scores of pollution-sensitive, biologically important metrics of the phytoplankton community into a single index. Like other multi-metric indexes, the P-IBI is more sensitive to habitat conditions than its component metrics, which include chlorophyll-a, the abundances of several potentially harmful species, and various indicators of cell function and species composition (Lacouture et al. 2006).

Thirty-eight phytoplankton metrics were used to quantify the status of phytoplankton communities relative to water quality conditions (Table 3.12). Least-impaired (reference) habitat conditions have low dissolved inorganic nitrogen (DIN) and orthophosphate (P04) concentrations and large Secchi depths. Impaired (degraded) habitat conditions have high DIN and P04 concentrations and small Secchi depths. The phytoplankton communities of these contrasting habitat conditions showed many significant differences (Table. 3.14, Buchanan et al. 2005). Twelve discriminatory metrics were chosen, and different combinations of these twelve metrics were scored and used to create phytoplankton community indexes for spring and summer in the four salinity regimes in Chesapeake Bay.

Table 3.14 Phytoplankton metrics examined in the development of the Chesapeake Bay Index of Biotic Integrity. From Lacouture et al. 2006.

TABLE 2. Phytoplankton metrics examined for discriminatory ability and their Kruskal-Wallis (χ^2) test results for significant difference between reference and degraded communities (* $p = 0.05$–0.1, ** $p = 0.01$–0.05, *** $p < 0.01$, ns = not significant). All metric values are for the above-pycnocline layer, except surface chlorophyll *a*. Carbon:chlorophyll *a* is the ratio of total nano-micro phytoplankton (2–200 μm) biomass to chlorophyll *a* in the above-pycnocline layer. Since picophytoplankton biomass is not included in the numerator, carbon:chlorophyll *a* values are somewhat underestimated in summer when picophytoplankton are most abundant. Average cell size is total biomass divided by total abundance of the nano-micro phytoplankton size fractions. See Table 1 for season and salinity zone definitions.

Metric	Spring				Summer			
	F	O	M	P	F	O	M	P
Chlorophyll *a*	**	***	**	***	***	***	***	***
Chlorophyll *a* surface	**	***	***	***	***	***	***	***
Pheophytin	***	***	***	***	***	***	***	***
Total biomass nano-micro phytoplankton	ns	**	***	ns	***	***	ns	***
Total abundance nano-micro phytoplankton	ns	***	ns	*	***	***	**	ns
Carbon:chlorophyll *a*	*	ns	***	***	ns	ns	***	***
Average cell size nano-micro phytoplankton	ns	*	***	**	ns	*	ns	***
Chlorophyte abundance	ns	ns	***	ns	***	***	***	ns
Chlorophyte biomass	ns	ns	***	ns	***	***	***	ns
Chrysophyte abundance	ns	**	ns	ns	**	**	ns	***
Chrysophyte biomass	ns	**	ns	ns	**	**	***	***
Cryptophyte abundance	ns	***	ns	***	ns	ns	*	*
Cryptophyte biomass	ns	**	ns	***	ns	ns	ns	*
% total biomass composed of cryptophytes	ns	ns	***	**	***	***	ns	***
Cyanophyte abundance	ns	*	**	ns	***	***	ns	***
Cyanophyte biomass	*	ns	**	ns	***	***	ns	***
% total biomass composed of cyanophytes	ns	ns	***	ns	ns	ns	ns	***
Diatom abundance	*	ns	ns	*	***	***	***	ns
Diatom biomass	ns	ns	***	ns	***	***	ns	***
% total biomass composed of diatoms	ns	***	***	ns	ns	***	ns	ns
Dinoflagellate abundance	ns	***	ns	*	ns	***	***	ns
Dinoflagellate biomass	ns	***	*	ns	ns	***	ns	***
% total biomass composed of dinoflagellates	ns	***	ns	ns	ns	***	ns	ns
Prasinophyte abundance	ns	***	ns	ns	ns	**	ns	**
Prasinophyte biomass	ns	***	ns	ns	ns	**	ns	**
Picoplankton abundance (Virginia only)	**	ns	*	ns	**	*	***	ns
Picoplankton biomass (Virginia only)	**	ns	*	ns	**	*	***	ns
Cochlodinium heterolobatum abundance			ns				ns	ns
Cochlodinium heterolobatum biomass			ns				ns	ns
Microcystis aeruginosa abundance	ns	ns	ns	ns	***	**	ns	ns
Microcystis aeruginosa biomass	ns	ns	ns	ns	***	**	ns	ns
Prorocentrum minimum abundance	**	***	***	***	ns	*	ns	ns
Prorocentrum minimum biomass	**	***	***	***	ns	*	ns	ns
Dissolved oxygen	ns	***	***	***	ns	**	***	***
Dissolved organic carbon	***	**	***	ns	***	***	***	ns
Particulate carbon	**	ns	***	***	***	***	ns	***
Total organic carbon	***	*	***	**	***	***	***	ns
Total suspended solids	***	***	***	***	***	***	***	***

4 References

Andersen, J. H., et al. (2011), Getting the measure of eutrophication in the Baltic Sea: towards improved assessment principles and methods, *Biogeochemistry*, *106*, 137-156.

Anderson, L. A., and J. L. Sarmiento (1994), Redfield ratios and remineralization determined by nutrient data analysis, *Global Biogeochemical Cycles*, *8*(1), 65-80.

Anderson, R. A. (Ed.) (2005), *Algal Culturing Techniques*, Elsevier.

Anderson, D.M., J.M. Burkholder, W.P. Cochlan, P.M. Glibert, C.J. Gobler, C.A. Heil, R. Kudela, M.L. Parsons, J.E. Rensel, D.W. Townsend, V.L. Trainer, and G.A. Vargo. 2008. Harmful algal blooms and eutrophication: Examining linkages from selected coastal regions of the United States. *Harmful Algae* 8:39-53.

Best, M. A., A. W. Wither, and S. Coates (2007), Dissolved oxygen as a physico-chemical supporting element in the Water Framework Directive, *Marine Pollution Bulletin*, *55*, 53-64.

Birk, S. and U. Schmedtje (2005). "Towards harmonization of water quality classification in the Danube River Basin: overview of biological assessment methods for running waters." Archiv für Hydrobiologie Supplement Large Rivers 16: 171–196.

Birk, S., W. Bonne, et al. (2012). "Three hundred ways to assess Europe's surface waters: An almost complete overview of biological methods to implement the Water Framework Directive." Ecological Indicators **18**: 31-41.

Borja, A., A. Basset, S. Bricker, J.-C. Dauvin, M. Elliot, T. Harrison, J.-C. Marques, S. B. Weisberg, and R. West (2011), Classifying ecological quality and integrity of estuaries, in *Treatise on Estuarine and Coastal Science*, edited by E. Wolanski and D. S. McLusky, pp. 125-162, Waltham: Academic Press.

Borja, A., S.B. Bricker, D.M. Dauer, N.T. Demetriades, J.G. Ferreira, A.T. Forbes, P. Hutchings, X.P. Jia, R. Kenchington, J.C. Marques and C.B. Zhu. 2008. Overview of integrative tools and methods in assessing ecological integrity in estuarine and coastal systems worldwide. *Marine Pollution Bulletin* 56:1519–1537.

Borja, A., J. Bald, J. Franco, J. Larreta, I. Muxika, M. Revilla, J. G. Rodriguez, O. Solaun, A. Uriarte, and V. Valencia (2009), Using multiple ecosystem components, in assessing ecological status in Spanish (Basque Country) Atlantic marine waters, *Marine Pollution Bulletin*, *59*(1-3), 54-64.

Borja, A., D. M. Dauer, et al. (2012). "The importance of setting targets and reference conditions in assessing marine ecosystem quality." Ecological Indicators **12**(1): 1-7.

Boynton, W. R., and W. M. Kemp (2008), Estuaries, in *Nitrogen in the Marine Environment, 2nd Edition*, edited by D. G. Capone, D. A. Bronk, M. R. Mulholland and E. J. Carpenter, pp. 809-856, Elsevier Inc., Burlington, Massachusetts.

Bricker, S. B., J. G. Ferreira, and T. Simas (2003), An integrated methodology for assessment of estuarine trophic status, *Ecological Modelling*, *169*(1), 39-60.

Bricker, S. B., C. G. Clement, D. E. Pirhalla, S. P. Orlando, and D. R. G. Farrow (1999), National Estuarine Eutrophication Assessment: Effects of Nutrient Enrichment in the Nation's Estuaries*Rep.*, NOAA, National Ocean Service, Speical Projects Office and National Centers for Coastal Ocean Science, Silver Springs, MD.

Bricker, S. B. (2007). Effects of Nutrient Enrichment in the Nation's Estuaries: A Decade of Change: National Estuarine Eutrophication Assessment Update, US Department of Commerce, National Oceanic and Atmospheric Administration, National Ocean Service, National Centers for Coastal Ocean Science.

Bricker, S.B., B. Longstaff, W. Dennison, A. Jones, K. Boicourt, C. Wicks and J. Woerner. 2008. Effects of nutrient enrichment in the nation's estuaries: A decade of change. *Harmful Algae* 8:21-32.

Buchanan C, Lacouture RV, Marshall HG, Olson M, Johnson JM. Phytoplankton reference communities for Chesapeake Bay and its tidal tributaries. Estuaries. 2005;28(1):138-59.

Carstensen, J., M. Sanchez-Camacho, C. M. Duarte, D. Krause-Jensen, and N. Marba (2011), Connecting the Dots: Responses of Coastal Ecosystems to Changing Nutrient Concentrations, *Environmental Science & Technology*, *45*(21), 9122-9132.

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

Cloern, J. E. (2001), Our evolving conceptual model of the coastal eutrophication problem, *Marine Ecology-Progress Series*, *210*, 223-253.

Cloern J.E. and Dugdale R. 2010. San Francisco Bay. In Nutrient in Estuaries: A Summary Report of the National Estuarine Experts Workgroup 2005–2007. P.M. Glibert, C.J.. Madden, W. Boynton, D. Flemer,C. Heil and J. Sharp(eds). US Environmental Protection Agency. Available online http://water.epa.gov/scitech/swguidance/standards/criteria/nutrients/upload/Nutrients-in-Estuaries-November-2010.pdf

Cloern, J.E. and R. Dufford. 2005. Phytoplankton community ecology: principles applied in San FranciscoBay. Marine Ecology Progress Series. 285: 11-28.

Conley, D. J., H. Kaas, F. Mohlenberg, B. Rasmussen, and J. Windolf (2000), Characteristics of Danish estuaries, *Estuaries*, *23*(6), 820-837.

Dennison WC, R.J. Orth, K.A. Moore, J.C. Stevenson, V. Carter, S. Kollar, P.W. Bergstrom, and R.A. Batiuk, R. Assessing water quality with submersed aquatic vegetation. Bioscience. 1993;43(2):86-94.

Devlin, M., S. Painting, and M. Best (2007a), Setting nutrient thresholds to support an ecological assessment based on nutrient enrichment, potential primary production and undesirable disturbance, *Marine Pollution Bulletin*, *55*(1-6), 65-73.

Devlin, M., S. Bricker, and S. Painting (2011), Comparison of five methods for assessing impacts of nutrient enrichment using estuarine case studies, *Biogeochemistry*, *106*(2), 177-205.

Devlin, M., M. Best, D. Coates, E. Bresnan, S. O'Boyle, R. Park, J. Silke, C. Cusack, and J. Skeats (2007b), Establishing boundary classes for the classification of UK marine waters using phytoplankton communities, *Marine Pollution Bulletin*, *55*, 91-103.

Domingues, R. B., A. Barbosa, and H. Galvao (2008), Constraints on the use of phytoplankton as a biological quality element within the Water Framework Directive in Portuguese waters, *Marine Pollution Bulletin*, *56*(8), 1389-1395.

Ferreira JG, Andersen JH, Borja A, Bricker SB, Camp J, da Silva MC, et al. Overview of eutrophication indicators to assess environmental status within the European Marine Strategy Framework Directive. Estuarine Coastal and Shelf Science. 2011;93(2):117-31.

Garmendia, M., S. Bricker, M. Revilla, A. Borja, J. Franco, J. Bald, and V. Valencia (2012), Eutrophication Assessment in Basque Estuaries: Comparing a North American and a European Method, *Estuaries and Coasts*, *35*(4), 991-1006.

Giordani, G., J. Zaldivar, and P. Viaroli (2009), Simple tools for assessing water quality and trophic status in transitional water ecosystems, *Ecological Indicators*, *9*(5), 982-991.

Giovanardi, F., and R. A. Vollenweider (2004), Trophic conditions of marine coastal waters: experience in applying the Trophic Index (TRIX) in two areas of the Adriatic and Tyrrhenian seas, *Journal of Limnology*, *63*, 199-218.

Hagy, J.D., W.R. Boynton, C.W. Keefe and K.V. Wood. 2004. Hypoxia in Chesapeake Bay, 1950-2001: Long-term change in relation to nutrient loading and river flow. *Estuaries* 27:634-658.

Hejzlar, J., et al. (2009), Nitrogen and phosphorus retention in surface waters: an inter-comparison of predictions by catchment models of different complexity, *Journal of Environmental Monitoring*, *11*(3), 584-593.

HELCOM (2009), Eutrophication in the Baltic Sea -- An integrated thematic assessment of the effects of nutrient enrichment and eutrophication in the Baltic Sea region, *Baltic Sea Environment Proceedings*(No. 115B), 148.

Huntington, B. E., and K. E. Boyer (2008), Effects of red macroalgal (Gracilariopsis sp.) abundance on eelgrass Zostera marina in Tomales Bay, California, USA, *Marine Ecology-Progress Series*, *367*, 133-142.

Karr JR, Chu EW. Biological monitoring: Essential foundation for ecological risk assessment. Hum Ecol Risk Assess. 1997;3(6):993-1004.

Kemp, P. E., B. F. Sheer, E. B. Sheer, and J. J. Cole (Eds.) (1993), *Handbook of Methods in Aquatic Microbial Ecology*, Lewis Publishers.

Kemp WM, R. Batiuk, R. Bartleson, P. Bergstrom, V. Carter, C.L. Gallegos, and W. Hunley. Habitat requirements for submerged aquatic vegetation in Chesapeake Bay: Water quality, light regime, and physical-chemical factors. Kennedy VS, editor. New York: Academic Press; 2004.

Kemp, W. M., et al. (2005), Eutrophication of Chesapeake Bay: historical trends and ecological interactions, *Marine Ecology-Progress Series*, *303*, 1-29.

Klausmeler, C. A., E. Litchman, T. Daufreshna, and S. A. Levin (2004), Optimal nitrogen-to-phosphorus stoichiometry of phytoplankton, *Nature*, *429*(6988), 171-174.

Lacouture RV, Johnson JM, Buchanan C, Marshall HG. 2006. Phytoplankton index of biotic integrity for Chesapeake Bay and its tidal tributaries. Estuaries and Coasts. 29(4):598-616.

McKee LJ, Sutula, Gilbreath, A.N., Gillett D., Beagle, J., Gluchowski, D., and Hunt, J. 2011. Numeric nutrient endpoint development for San Francisco Bay- Literature review and Data Gaps Analysis. Southern California Coastal Water Research Project Technical Report No. 644. www.sccwrp.org

McLaughlin K. and Sutula M. Developing nutrient numeric endpoints and TMDL tools for California estuaries: An implementation plan. K McLaughlin, M Sutula. Technical Report 540. Southern California Coastal Water Research Project. Costa Mesa, CA.

Morse, R.E., J. Shen, J.L. Blanco-Garcia, W.S. Hunley, S. Fentress, M. Wiggins, and M.R. Mulholland. 2011. Environmental and physical controls on the formation and transport of blooms of the dinoflagellate Cochlodinium polykrikoides Margalef in lower Chesapeake Bay and its tributaries. Estuaries and Coasts 34: 1006-1025.

Mulholland , M.R., R.E. Morse, G.E. Boneillo, P.W. Bernhardt, K.C. Filippino, L.A. Procise, J.L. Blanco-Garcia, H.G. Marshall, T.A. Egerton, W.S. Hunley, K.A. Moore, D.L. Berry and C.J. Gobler. 2009. Understanding causes and impacts of the dinoflagellate, *Cochlodinium polykrikoides*, blooms in the Chesapeake Bay. *Estuaries and Coasts* 32: 734-747.

Nixon, S. W. (1995), Coastal marine eutrophication: a definition, social causes, and future concerns, *Ophelia*, *41*, 199-219.

Paerl, H.W., J.J. Joyner, A.R. Joyner, K. Arthur, V.J. Paul, J. M. O'Neil, and C. A. Heil. 2008. Co-occurrence of dinoflagellate and cyanobacterial harmful algal blooms in southwest Florida coastal waters: A case for dual nutrient (N and P) input controls. *Marine Ecology Progress Series* 371:143-153.

Painting, S., et al. (2006), Assessing the impact of nutrient enrichment in estuaries: susceptibility to eutrophication, *Marine Pollution Bulletin*, *55*, 74-90.

Painting, S. J., et al. (2007), Assessing the impact of nutrient enrichment in estuaries: Susceptibility to eutrophication, *Marine Pollution Bulletin*, *55*(1-6), 74-90.

Redfield, A. C., B. H. Ketchum, and F. A. Richards (1963), The influence of organisms on the composition of seawater, in *The Sea*, edited by M. N. Hill, pp. 26-77, Wiley Intersciece.

Senn D. and Novick E. 2013. A Conceptual Model of Nutrient Cycling in San Francisco Bay. San Francisco Estuary Institute. www.sfei.org

SFRWQCB 2012. San Francisco Bay Nutrient Management Strategy. http://www.waterboards.ca.gov/sanfranciscobay/water_issues/programs/planningtmdls/amendments/estuarineNNE/Nutrient_Strategy%20November%202012.pdf

Smith, D.E., M. Leffler and G. Mackiernan (Eds). 1992. *Oxygen dynamics in Chesapeake Bay: A synthesis of research.* University of Maryland Sea Grant, College Park, Maryland, USA, 234 p.

Smith, S. V., D. P. Swaney, R. W. Buddemeier, M. R. Scarsbrook, M. A. Weatherhead, C. Humborg, H. Eriksson, and F. Hannerz (2005), River nutrient loads and catchment size, *Biogeochemistry*, *75*(1), 83-107.

Souchu, P., M. C. Ximenes, M. Lauret, A. Vaquer, and E. Dutrieux (2000), Mise a jour d'indicators du niveau d'eutrophisation des milieux lagunaaires mediterraneens*Rep.*, 412 pp, Ifremer-Creocean-Universite Montpellier II.

Sutula M. (2013)2013 Developing Nutrient Numeric Endpoints For California Estuaries: Conceptual Framework and Technical Work plan Technical Report 540. Southern California Coastal Water Research Project. Costa Mesa, CA.

Sutula M, Creager C, Wortham G. (2007) Technical approach to develop nutrient numeric endpoints for California estuaries. Costa Mesa: Southern California Coastal Water Research Project, Tetra Tech, 2007

Steeman-Nielsen, E. (1952), The use of radioactive carbon (14C) for measuring organic production in the sea, *Journal du Conseil*, *18*, 117-140.

USEPA (2001), National Coastal Condition Report I *Rep.*, United States Environmental Protection Agency, Office of Research and Development, Office of Water, Washington, DC.

USEPA (2001). Nutrient Criteria Technical Guidance Manual: Estuarine and Coastal Marine Waters. U.S. Environmental Protection Agency, Washington D.C.; 2001.

USEPA (2003). *Ambient water quality criteria for dissolved oxygen, water clarity and chlorophyll a for Chesapeake Bay and its tidal tributaries.* EPA 903-35 U.S. Environmental Protection Agency, Region 3, Chesapeake Bay Program Office, Annapolis, Maryland, USA.

USEPA (2005), National Coastal Condition Report II*Rep.*, United States Environmental Protection Agency, Office of Research and Development, Office of Water, Washington DC.

USEPA (2006). Technical Approach to Develop Nutrient Numeric Endpoints for California. IU.S. Environmental Protection Agency WDC, Lafayette, CA: Tetra Tech, INC.; 2005. p. 115.

USEPA (2007). Ambient Water Quality Criteria for Dissolved Oxygen, Water Clarity and Chlorophyll a for the Chesapeake Bay and Its Tidal Tributaries 2007 Chlorophyll Criteria Addendum. Region III Chesapeake Bay Program Office, October 2007. Report No.: Contract No.: EPA 903-R-07-005 CBP/TRS 288/07.

USEPA (2008), National Coastal Conditions Report III*Rep.*, United States Environmental Protection Agency, Office of Research and Development, Office of Water, Washington DC.

USEPA. 2010a. *Methods and Approaches for Deriving Numeric Criteria for Nitrogen/Phosphorus Pollution in Florida's Estuaries, Coastal Waters, and Southern Inland Flowing Waters*. http://yosemite.epa.gov/sab/sabproduct.nsf/fedrgstr_activites/C439B7C63EB9141F85257 73B004E53CA/$File/FL+EC+Final+Methods+-+Chapters1-6.pdf. Last updated November 17, 2010.

Viaroli, P., M. Bartoli, G. Giordani, M. Naldi, S. Orfanidis, and J.-M. Zaldivar (2008), Community shifts, alternative stable states, biogeochemical controls and feedbacks in eutrophic coastal lagoons: a brief overview, *Aquatic Conservation: Marine and Freshwater Ecosystems*, *18*, S105-S117.

VIncent, C. H., A. Heinrich, K. Edwards, K. Nygaard, and K. Haythornthwaite (2002), Guidance on typology, reference conditions and classification systems for transitional and coastal waters*Rep.*, 119 pp, CIS Working Group 2.4 (COAST), Common Implementation Strategy of the Water Framework Directive, European Commission.

Vollenweider, R. A., F. Giovanardi, G. Montanari, and A. Rinaldi (1998), Characterization of the trophic conditions of marine coastal waters with specia reference to the NW Adriatic Sea: proposal for a trophic scale, turbidity and generalized water quality index, *Environmetrics*, *9*, 329-357.

Walker WW. Emprical methods for predicting eutrophication in impoundments; Report 3, Phase III: Model Refinements. Vicksburg, MS: US Army Corps of Engineers Waterways Experiment Station, 1985.

Zaldivar, J.-M., et al. (2008), Eutrophication in transitional waters: an overview, *Transitional Waters Monographs*, *1*, 1-78.

APPENDIX I – CATALOGUE OF SF BAY DATA AVAILABLE FOR ANALYSIS OF EXISTING DATA

The existing data available to test out assessment approaches generally falls into two categories: 1) USGS water quality sampling and 2) IEP monitoring data.

The parameters sampled and the time periods for which these data are available are summarized in this appendix.

USGS

USGS consists of a long term data set collected from 1975-2011, with the exact coverage varying by station (Figure A1.1, Table A1.1). Nutrients were sampled regularly beginning in 2004 at a subset of all stations. Parameters consist of Chl-a, DO, SPM, salinity, temp, depth, and nutrients (NO2, NO3, NH3, PO4, Si). During the period of 1992-2001, USGS also collected phytoplankton composition data. These data were analyzed by Cloern and Dulford (2005).

Figure A1.1 USGS water quality sampling stations in SF Bay.

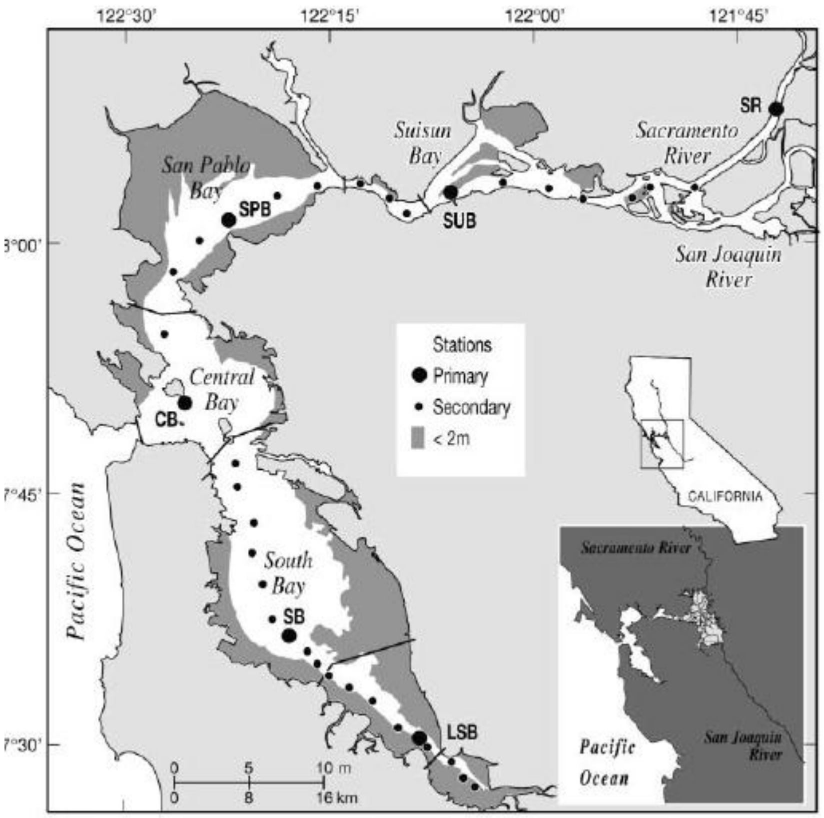


Figure A1.2 Station at which phytoplankton taxonomic composition data were collected (primary stations) during 1992-2001.

'

DWR-IEP

The Department of Water Resources (DWR) and the Interagency Ecological Program (IEP) have been collecting data from 1975-2011, with exact coverage varying by station (Figure A1.3, Table A1.2). Parameters collected include Chl-a, BOD, SPM, TDS, VSS, salinity, depth, pH, DO, turbidity, temp, pheophytin-a, DOC, TOC, nutrients (NH_3, TKN, NO_3, NO_2, DON, TON, PO_4, TP, Si), and taxonomic assemblage. For the latter, 16 phytoplankton species were enumerated prior to 2008 while 21 species were enumerated from 2008-2010.

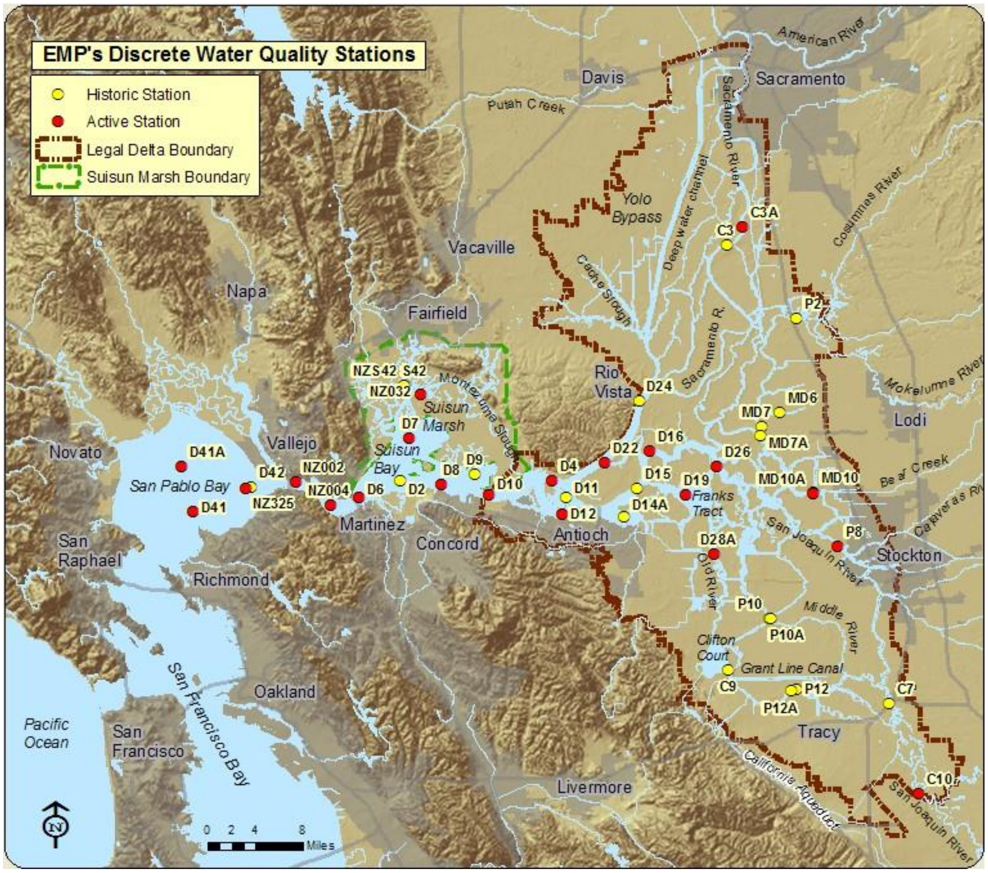


Figure A1.3 Stations sampled under the DWR-IEP monitoring program.

Scientific Bases for Establishing Chlorophyll-*a* Thresholds for San Francisco Bay

Martha Sutula[1], Raphael Kudela[2], James D. Hagy III[3], Lawrence W. Harding, Jr.[8], David Senn[9], James E. Cloern[6], Gry Mine Berg[4], Suzanne Bricker[5], Richard Dugdale[7], and Marcus Beck[3]

[1] Southern California Coastal Water Research Project, Costa Mesa, California 92626 USA

[2] Ocean Sciences Department, University of California Santa Cruz, California 95064 USA

[3] U.S. Environmental Protection Agency, Office of Research and Development, Gulf Breeze, Florida 32561 USA

[4] Applied Marine Sciences, Santa Cruz, California USA 95060

[5] NOAA National Centers for Coastal Ocean Science, Silver Spring, Maryland 20910 USA

[6] U.S. Geological Survey, Menlo Park, California 94025 USA

[7] Romberg Tiburon Center, San Francisco State University, Tiburon, California 94920 USA

[8] Department of Atmospheric and Oceanic Sciences, University of California, Los Angeles, California 90095 USA

[9] San Francisco Estuary Institute, Richmond, California 94804 USA

For Submittal to: *Estuarine, Coastal and Shelf Science*

Author for correspondence: Martha Sutula email – marthas@sccwrp.org

Keywords: eutrophication; water quality criteria; dissolved oxygen, HAB, *chl-a*; SPATT

Draft: 03 November 2015

Abstract

San Francisco Bay (SFB) receives high nutrient loads from agricultural runoff, storm water, and treated wastewater effluent from 37 Publically Owned Treatment Works (POTWs), although to date the estuary appears resistant to classic symptoms of eutrophication. Recent trends of increasing chlorophyll-*a* (*chl-a*), harmful algal blooms (HAB), and dissolved oxygen concentrations (DO) suggest this resistance may be weakening. These findings motivated development of water-quality criteria (WQC) for SFB protective from adverse effects of nutrient over-enrichment. WQC consisting of thresholds of phytoplankton biomass as *chl-a* are based on strong relationships between nutrients, *chl-a*, and water-quality impairments in several estuaries. Although plankton ecology is well chronicled for SFB, data from several decades of monitoring have not been used heretofore to support WQC. Here, we analyze long-term data on *chl-a* (1993-2014), phytoplankton species composition (1993-2014), algal toxins (2012-2014), and DO (1993-2014) to derive: (1) quantitative relationships of HAB abundances, toxin levels, and DO to *chl-a;* and (2) *chl-a* thresholds and related uncertainties corresponding to "protected" and "at risk" categories based on WQC for DO and HAB alerts. Although *chl-a* is lower and DO higher in SFB than comparable estuaries experiencing nutrient over-enrichment, we report trends of increasing *chl-a*, declining DO, ubiquitous presence of HAB species, and toxin concentrations exceeding alert levels in ~35% of samples over the last 20 years. Quantile regressions of *chl-a* with HAB abundance and DO were significant, indicating SFB is poised for increased risk of impairments by HAB and low DO with increasing phytoplankton biomass. Coordinated statistical analyses showed *chl-a* thresholds associated with HAB and DO impairments converged on comparable values. We identified monthly mean *chl-a* < 13 mg m^{-3} as an inflection point, below which probabilities for exceeding alert levels for HAB abundances and HAB toxins

were reduced. This HAB-based *chl-a* threshold was similar to a *chl-a* threshold of 13 - 16 mg m^{-3} for meeting the WQC for DO of 7 mg L^{-1}. At the high-end of risk, *chl-a* thresholds from 25 - 40 mg m^{-3} corresponded to a 0.5 probability of exceeding alert levels for HAB abundance, and with consistent excursions of DO in lower South Bay (LSB) and South Bay (SB) below the WQC of 5.0 mg L^{-1} for DO. We suggest that if available nutrients in SFB were assimilated into phytoplankton biomass, mean *chl-a* in all sub-embayments of SFB could reach “high risk” thresholds. These findings justify the establishment of *chl-a* thresholds to support nutrient management of SFB, given uncertainty about the future trajectory of water quality in this important estuarine ecosystem.

Introduction

Nutrient over-enrichment of the world's estuaries has led to multiple ecosystem impairments that express cultural eutrophication (Nixon, 1995; Paerl 1997; Cloern 2001; Diaz and Rosenberg, 2008; Bricker et al. 2008). Identifying specific water-quality goals for nutrients has proven difficult, however, because ecological responses to nutrients are complex. San Francisco Bay (SFB) is a well-documented example of a nutrient-enriched estuary that exhibits this complexity (Cloern and Jassby, 2012). Extensive long-term data suggest that, to date, SFB has been resistant to classic symptoms of nutrient over-enrichment such as high phytoplankton biomass, harmful algal blooms (HAB), and low dissolved oxygen (DO). A number of factors have precluded widespread development of these symptoms in SFB, including high turbidity and concomitant light-limitation of primary productivity, intense tidal mixing that reduces biomass accumulation and DO depletion, and grazing by large populations of filter-feeding clams that regulates phytoplankton biomass in some areas of the bay (cf. Cloern and Jassby, 2012; Cloern et al., 2007; Kimmerer and Thompson, 2014).

Recent evidence suggests resistance to nutrient over-enrichment may be weakening in SFB, such as: (1) a three-fold increase of chlorophyll-*a* (*chl*-*a*) in South Bay (SB) during summer-fall since 1999 (Cloern et al., 2007); (2) regular occurrences of HAB species (Lehman et al., 2005; Cloern et al., 2005; Cloern and Dufford, 2005); and (3) diurnal depressions of DO to hypoxic conditions with DO < 2.8 mg L^{-1} in restored salt ponds (Thebault et al., 2008; Topping et al., 2009). These observations call for a water-quality framework to inform management actions, consisting of thresholds for key properties that would be "protective" from adverse effects of nutrient over-enrichment. Phytoplankton biomass as *chl-a* is an integrative indicator of nutrient loadings with established links to water-quality impairments, commonly used to assess eutrophication and support regulatory goals (Bricker et al., 2003; Zaldivar et al., 2008; Harding et al., 2014). Quantitative thresholds leading to management endpoints can be based on deviations from "reference" conditions when data prior to degradation are available (Andersen et al. 2010, 2015), or on ecosystem impairments such as low DO, HAB, or water clarity (e.g., Harding et al., 2014). We lack *chl-a* records for SFB prior to human disturbance, limiting the use of reference conditions, but long-term data on *chl-a* support quantitative analyses of relationships between *chl-a* and potential impairments.

Two pathways of nutrient over-enrichment that culminate in adverse effects on humans, marine mammals, and other aquatic life include: (1) low DO associated with excess organic matter; and (2) increased HAB occurrences (Rosenberg et al., 1991; Diaz and Rosenberg, 1995; Kirkpatrick et al., 2004; Glibert et al., 2005; Baustin and Rabalais, 2009). Recognizing that factors other than nutrients affect low DO and HABs, causal links are established for nutrient loadings, *chl-a*, hypoxia, and HABs (Tett et al., 2007; Heisler et al., 2008; Anderson et al., 2012). Such links have been used to support water quality criteria (WQC) for Chesapeake Bay, relating risk of impairments to increased *chl-a* (e.g., Harding et al., 2014). As in many estuaries, *chl-a* has increased significantly over the past 15-20 years in SFB, amounting to a three-fold increase from the mid-1990s to mid-2000s in South Bay (SB) and Lower South Bay (LSB) (Cloern et al., 2007). Of particular concern are regular occurrences of fall blooms of phytoplankton in SB and LSB since the late 1990s, areas that rarely experienced such outbreaks in the past (Cloern and Jassby, 2012), and significant increases of *chl-a* in other sub-embayments during the same period. Despite these upward trends of *chl-a* and reports of HAB occurrences (see Cloern et al., 1994), routine monitoring for algal toxins has not been conducted (Cloern and Dufford, 2005). Moreover, long-term data on *chl-a*, phytoplankton species composition, and DO to quantify risk of low DO or HAB occurrences with increasing *chl-a* have yet to be assembled for various sub-embayments of SFB.

Here, we present relationships between DO, HAB, and *chl-a* in SFB to derive quantitative thresholds based on water-quality impairments. We then apply these thresholds as endpoints to support assessments of status and trends of water quality required by both scientists and managers (Sutula et al., 2015). Our goals were to: (1) determine relationships DO, HAB occurrences, and algal toxins to *chl-a*; and (2) quantify *chl-a* thresholds and associated uncertainties using statistical approaches that identify "protected" and "at risk" categories in the context of WQC for DO and HAB alerts.

Materials and Methods

Study Area

SFB is the largest estuary in California, consisting of several major sub-embayments (Nichols et al., 1986). The estuary receives nutrient loads from 37 publicly owned wastewater treatment works (POTW) serving the area's population of 7.2 million (Fig. 1). Most POTW perform only secondary treatment without additional nitrogen (N) or phosphorus (P) removal. Freshwater flow into SFB comes from two major sources, the Sacramento and San Joaquin Rivers, large rivers that drain 40% of California's landscape. Intense agriculture in the heavily farmed Central Valley combined with urban sources such as Sacramento ~100 km upstream of Suisun Bay (SUB) contribute to high nutrient loads entering the northern estuary from the Sacramento/San Joaquin Delta. Storm-water runoff from densely populated urban areas surrounding SFB also contributes significant nutrients.

Conceptual Approach

Long-term data from the SFB Research Program (1993-2014) of the US Geological Survey (USGS) and concurrent measurements of algal toxins (2012-2014) supported analyses of trends for DO and HAB. Relationships of DO and HAB to *chl-a* were used to identify *chl-a* thresholds that correspond to risks of low DO or HAB alerts. Increased *chl-a* does not uniformly correspond to increased HAB occurrences, particularly for a single phytoplankton species or toxin, and both high-biomass and high-toxicity events are well described (Anderson et al., 2012). For the former, significant relationships between HAB and *chl-a* have been used to support WQC (Shutler et al., 2012; Schaeffer et al., 2012; Harding et al., 2014). Predominance of a particular taxonomic group, i.e., diatoms, expressed as cell counts or fraction of bio-volume is often accompanied by increased *chl-a*. Conditions that support increased *chl-a*, however, are known to increase abundance of the entire phytoplankton community, not just HAB species (Barber and Hiscock, 2006). For our analyses, we assumed increased *chl-a* reflected increased abundance of all phytoplankton, including potentially toxic HAB based on previous studies (Bricker et al., 2008; Glibert et al., 2005).

Decomposition of excess phytoplankton biomass supports DO consumption, leading to hypoxia (DO < 2.8 mg L^{-1}) in stratified conditions. Spatial and temporal displacement of high *chl-a* and DO depletion commonly occurs in estuaries (Rabalais et al., 2014), reflecting strong seasonality of production and consumption (e.g., Wheeler et al., 2003). Empirical relationships between DO and *chl-a* exhibit time lags, with analyses requiring consideration of relevant time and space scales for individual ecosystems. Accordingly, we aggregated DO and *chl-a* data for a range of time scales for the six sub-embayments to evaluate the strength of these relationships.

Data Sources

USGS SFB Research Program. Our analyses drew on time-series data collected on regular cruises by the USGS along a 145-km transect from 1993 – 2014. These observations provided a complete record of *chl-a*, DO, conductivity, temperature, turbidity, and photosynthetically available radiation (PAR) (http://sfbay.wr.usgs.gov/access/wqdata/query/index.html). Vertical profiles were conducted with a Seabird Electronics SBE9+ CTD and rosette sampler equipped with a Turner Designs C3 fluorometer, Li-Cor LI 192 transmissometer, and Seabird SBE 43 DO electrode. Concurrent grab samples were collected for identification and enumeration of phytoplankton species. Discrete measurements of DO and *chl-a* were used to calibrate instruments and correct for turbidity.

Data were aggregated by sub-embayment (Fig. 1) as geomorphology and nutrient loadings affect ecological responses to nutrient inputs in SFB (Jassby et al., 1997). Sub-embayments consist of: (1) Lower South Bay (LSB), the area south of Dumbarton Bridge; (2) South Bay (SB), from Dumbarton Bridge to San Bruno Shoal; (3) Central Bay (CB), from San Bruno Shoal to Angel Island; (4) North Central Bay (NCB), from Angel Island to Pt. San Pablo; (5) San Pablo Bay (SPB), from Pt. San Pablo to Martinez; and (6) Suisun Bay (SUB), east of Martinez. USGS stations corresponding to these sub-embayments are 34-36, 24-32, 23-20, 18-16, 15-10, and 4-8, respectively. For some analyses, data from several sub-embayments were combined based on the statistical similarities to obtain bay-wide metrics, and to increase sample sizes for uncommon, but potentially deleterious HAB species.

HAB Species and Toxins. HAB species identified by the USGS were used for this analysis (Table 1). Seasonal and inter-annual patterns were identified for the three most common HAB species in SFB, *Pseudo-nitzschia sp.*, *Alexandrium sp.*, *Dinophysis sp.*, for several dinoflagellate species, including *Heterosigma akashiwo, Karenia mikimotoi, Karlodinium veneficum*, and for cyanobacteria including the genera *Microcystis, Oscillatoria, Planktothrix, Anabaenopsis*, and *Anabaena*. Some rare species with low frequencies-of-occurrence were excluded from the analyses. SFB does not currently have established guidance for potentially deleterious HABs, so we used alert levels from the literature, monitoring programs, and analyses of available data. These included: 10^6 cells L^{-1} for cyanobacteria (WHO 2003), presence/absence for *Alexandrium* (http://www.scotland.gov.uk/Publications/2011/03/ 16182005/37), 10^2 - 10^3 cells L^{-1} for *Dinophysis spp.* (http://www.scotland.gov.uk /Publications/2011/03/16182005/37; Vlamis et al., 2014), and 10^5 – 5 x10^5 cells L^{-1} for *Pseudo-nitzschia.* The dinoflagellates *H. akashiwo, K. mikimotoi*, and *K. veneficum* lack guidance on alert levels, so we used 5 x 10^5 cells L^{-1} based on expert opinion. No defined alert levels exist for toxin concentrations estimated using Solid Phase Adsorption Toxin-Tracking (SPATT - MacKenzie et al., 2004), thus alert levels were defined as 1 ng g^{-1} for microcystins (MCY), and 75 ng g^{-1} for domoic acid (DA) based on laboratory calibrations and studies at the Santa Cruz Municipal Wharf and Pinto Lake, California (Lane et al., 2010; Kudela, 2011; Gibble and Kudela, 2014).

We deployed SPATT samplers in the flow-through system of the *R/V Polaris* (~1 m intake) from October 2011 to November 2014 to assess the presence of DA and MCY. Individual SPATT deployments encompassed South SFB (stations 36-18, representing LSB+SB+CB), and typically stations 36-24 for full-bay cruises (representing LSB+SB), NCB (stations 21-16), SPB (stations 15-9), and SUB and the Delta (stations 8-657). Data were binned by sub-embayment, with SB and South-CB defined as stations 36-24 and 36-18, respectively. SPATT were operated as described previously (Lane et al., 2010; Kudela, 2011; Gibble and Kudela, 2014) for MCY, and reported as the total of LR, RR, YR, and LA congeners and domoic acid (DA). SPATT toxin concentrations were reported in units of ng toxin g^{-1} resin, and represented a weighted-average for the length of deployment (sub-embayment).

Statistical Analyses

Three statistical approaches were used: (1) seasonal Mann-Kendall Test to quantify temporal trends of DO, HAB and *chl-a*; (2) ordinary least squares regressioin (OLS), robust (e.g., Least Absolute Deviation - LAD), and quantile regression to determine relationships of DO, HAB, and *chl-a*; and (3) quantile regression and conditional probability analyses (CPA) to derive quantitative thresholds for *chl-a* based on a failure to achieve DO benchmarks, or an increased risk of reaching HAB abundances to trigger HAB alerts. Quantile regression was used to determine the 10^{th} and 50^{th} (median) quantiles of DO or HAB cell densities conditional on *chl-a*. Quantile regression is statistically analogous to rank-based correlation; it is based on ordering the observations, is robust to extreme values, and does not require assumptions about distributions of residuals (Cade and Noon, 2003). LAD regression is similar to quantile regression when the median quantile is used, although LAD uses ranks. CPA was used to analyze risk of DO below a WQC, or HAB abundance above a set alert level based on *chl-a* $\geq$ a specified concentration (R package *CProb;* Hollister et al., 2008). The baseline probability is the overall probability of exceedance among all observations, without regard to *chl-a* (i.e., *chl-a* > minimum value). Inflection points in the relationship are interpreted as *chl-a* above which probability of an adverse DO or HAB event increases at a faster rate relative to increases of *chl-a*. A probability of 0.5 is nominally defined as a benchmark of “elevated risk” because above this level, an adverse event is more likely to occur than not.

Statistical Analyses of HAB – *Chl-a*. HAB cell densities and toxins were analyzed in the context of seasonal and inter-annual patterns of *chl-a*. Near-surface samples ($\leq$ 2 m) collected from April - November were used for these analyses. "Calculated *chl-a*" consisting of fluorescence calibrated by discrete samples was also used for these analyses. All cell counts for known HAB species were used, regardless of depth or location, to increase sample sizes. More than 95% of HAB were from near-surface samples. USGS enumerates phytoplankton for samples with *chl-a* ≥ 5 mg m^{-3}, introducing a possible sampling bias by neglecting HAB at low *chl-a*. Another potential bias in the data is a lack of records for *Microcystis* spp., suggesting these cells were not identified by microscopy although they are regularly observed in northern SFB.

OLS and LAD regressions of HAB cell counts and SPATT toxin concentrations on *chl-a* were based on log-transformed data to improve normality. For toxin analysis, $\log_{10}$-transformed SPATT were compared to mean or maximum *chl-a* from corresponding SFB sub-embayments. Cell counts and *chl-a* were transformed by natural logarithm. HAB alert levels (see above) were used to derive probabilities that HAB or toxins would reach problematic levels with increased *chl-a*. Selection of alert level influenced the probability derived from CPA (see below).

Quantile regression and CPA were used to identify *chl-a* thresholds based on the risk of exceedances of HAB cell densities or toxin alert levels. First, CPA was conducted on HAB cell densities and SPATT data aggregated for all sub-embayments. A "HAB event of concern" was classified as a site with at least one HAB species exceeding cell-density alert levels. Second, quantile and OLS regressions were used to quantify relationships between cell densities of *Alexandrium, Dinophysis, Heterosigma, Karlodinium*, and *Pseudo-nitzschia* and *chl-a*. Corresponding analyses were performed for SPATT data aggregated among years and sub-embayments. Additional analyses were conducted by sub-embayment, but the results were similar (albeit with reduced statistical power) and were omitted for brevity. Analyses were also performed with and without *Alexandrium*, potentially biasing the analysis because this HAB has a low alert level. Concern that patterns may be strongly influenced by exchange with the open coast led to CPA on the complete time-series stratified by sub-embayment. Finally, data were divided into pre- and post-2002 to identify potential decadal differences.

Statistical Analyses of DO - *Chl-a*. Relationships between DO and *chl-a* were derived using the USGS data (1993-2014). Mean *chl-a* from depths ≤ 2 m was calculated for each station for periods-of-interest identified in previous analyses as showing *chl-a* changes (Cloern et al., 2007). These periods include: 1) spring bloom (February-May), 2) summer baseline (June-September), and 3) combination of these two periods (February-September). Mean February-September *chl-a* proved integrative of changing phytoplankton productivity in SFB and was chosen as the time period to derive thresholds of risk of low DO.

The evaluation period for DO was based on periods of non-compliance using existing WQC for DO in SFB: 1) an instantaneous WQC > 7 mg L^{-1} upstream and > 5 mg L^{-1} downstream of the Carquinez Bridge, not to fall below these values more than 10% of the time (SFRWQCB, 2011); and 2) > 80% saturation in running three-month medians in any sub-embayment of SFB. Medians for percent saturation and concentration (mg L^{-1}) of DO were computed from vertical profiles at each station. The number of stations below the WQC was tabulated by sub-embayment over the 20-year period. The three-month intervals with the most DO exceedances were used for further statistical analyses as these periods are sensitive to low DO and would correspond to protective thresholds for *chl-a*.

Quantile regression was used to investigate relationships between DO and *chl-a* by sub-embayment for several time lags. DO percent saturation was preferred to DO concentration as it removed variability associated with effects of temperature and salinity on solubility. Median (i.e., τ =0.5) quantiles of percent DO saturation were used to test the significance and slope of the relationship between DO and *chl-a* for three periods of integration. Sub-embayments for which a there was a positive slope between median percent DO saturation and *chl-a* were omitted from further analyses.

For sub-embayments with significant ($p < 0.05$) negative relationships between DO and *chl-a*, thresholds of increased risk of falling below DO benchmarks were quantified using two approaches. First, quantile regression using $\tau = 0.1$ was used to predict the mean and 95% confidence intervals of *chl-a* at which a gradient of DO percent saturation of 80%, 72%, 57%, and 46% would be attained 90% of the time. The remaining 10% non-attainment corresponds to California State Water Resource Control Board guidance for listing of impaired waters (SWRCB, 2005). These percent DO saturation are equivalent to DO concentrations of 7.0, 6.3, 5.0 and 4.0 mg L^{-1} at mean summer temperature 15^{o} C and salinity 24. Benchmark concentrations of 6.3 and 5.0 mg L^{-1} are the lowest DO concentrations to which salmonid and non-salmonid fish, respectively, can be exposed indefinitely without resulting in > 5% impact to estuarine populations (Bailey et al., 2014). 7.0 and 5.0 mg L^{-1} benchmarks are the established WQC for DO for SFB sub-embayments. In addition, *chl-a* at which DO percent saturation was expected to meet the median three-month percent saturation WQC of 80% for SFB was estimated as the 50^{th} quantile regression line. Finally, CPA was used to identify change points in the probability of DO falling below established WQC for DO with increasing *chl-a*.

Results

HAB Cell Densities and Algal Toxins

HAB species were detected in ~50% of samples and exceeded alert levels in ~35% of samples. Of samples exceeding alert levels, 53% were associated with *Alexandrium*, 11% with *Dinophysis*, and 7% with *Pseudo-nitzschia* (Fig. 2). Few toxic HAB events have been reported for SFB, but SPATT data confirm common occurrences of toxins DA and MCY (Fig. 3). Of 158 SPATT samplers we deployed, 72% showed detectable MCY and 97% showed detectable DA. Mean concentrations were 0.75 ng g^{-1} for MCY and 57 ng g^{-1} for DA, with ranges of detectable toxin for SPATT from 0.01 - 25.5 ng g^{-1} for MCY and from 1.69 - 1650 ng g^{-1} for DA (Fig. 3a-b).

DO, HAB Cell Densities, Algal Toxins, and Chl-a

Significant decreases of DO and increases of *chl-a* occurred in all sub-embayments from 1993 – 2013 ($p < 0.05$) based on a seasonal Mann-Kendall test (Fig. 4). Particularly notable were increases of summer baseline *chl-a* throughout SFB, with the largest increases in central and southern sub-embayments (Fig. 4). Sen slopes ranged from -0.9 to -1 percent saturation yr^{-1} for DO and from 0.041 to 0.096 mg m^{-3} yr^{-1} for *chl-a*. Cell counts were aggregated for all sub-embayments to increase sample sizes, except for *Pseudo-nitzschia* (n = 166) and *Alexandrium* (n = 261). HAB organisms *Alexandrium,* several cyanobacteria, *Dinophysis, Heterosigma, Karlodinium,* and *Pseudo-nitzschia* showed no significant increases based on Kendall's Tau tests ($p > 0.1$), while *Karenia* showed a significant, positive trend ($p < 0.05$). Cell counts for *Pseudo-nitzschia* or *Alexandrium* analyzed by sub-embayment showed no significant trends (ANCOVA, $p > 0.05$).

The 10th percentile of summer DO ranged from 5.7 – 7.8 mg L^{-1} on south to north transects (Supplemental Material, Table S1 and Fig. S1). DO was > 5 mg L^{-1} from 97.1 to 100% of the time along these transects, with DO > 7 mg L^{-1} in SUB 100% of the time. For most sub-embayments, evaluation periods that most frequently fell below the WQC for DO consisting of a three-month running median of 80% saturation were May-July and June-August (Supplemental Material, Table S2), and these periods were used in quantile regressions.

Relationships of HAB and Chl-a

Relationships of HAB abundance and SPATT toxins to *chl-a* showed considerable scatter. Cell counts of *Alexandrium, Dinophysis, Karlodinium,* and *Pseudo-nitzschia* increased with increasing *chl-a* and slopes of LAD regressions were significant ($p < 0.05$). Slopes and corresponding [R^2] for these HAB were 0.48 [0.25], 0.56 [0.33], 1.4 [0.44], and 0.43 [0.45] respectively. Cyanobacteria, *Heterosigma*, and *Karenia* showed no significant trends. SPATT data analyzed by sub-embayment showed a significant increase of MCY and DA with increasing mean *chl-a*, and a significant increase of DA with maximum *chl-a*.

Relationships of DO and Chl-a

Median DO in May-July and June-August showed similar patterns with consistently negative slopes for SUB, SPB, SB, and LSB, regardless of the evaluation period for *chl-a* (Table 1; Fig. 5). Slopes were generally steepest and most significant for mean February-September *chl-a*. Unlike other sub-embayments, the DO - *chl-a* relationship for SB was relatively insensitive to the evaluation period for *chl-a*, with significant relationships for most combinations. DO - *chl-a* relationships were significant in LSB for several evaluation periods. The June-September mean was only negative and significant in SPB, SB, and LSB, while the February-May mean was only negative and significant when correlated with SB and LSB (Table 1). In contrast, slopes were often positive for CB and NB. Quantile regressions of SUB and SPB, while significant, contained relatively few observations at high *chl-a* (Fig. 5). In addition, NB and CB had insufficient exceedances of WQC for DO to run CPA. For this reason, all sub-embayments except SB and LSB were omitted from further analyses to derive DO-related *chl-a* thresholds.

Thresholds Based on Chl-a

<u>HAB Relationships to *Chl-a*.</u> The baseline probability of HAB occurrences for the full range of *chl-a* was 0.35 - 0.40 (Fig. 6a). Interpretation of this baseline probability is that 35-40% of all samples from 1993-2014 exceeded HAB alert levels based on abundance (cells L^{-1}). A mean probability of 0.5 to exceed HAB alert levels corresponded to *chl-a* > 37.5 mg m^{-3} with an upper 95% confidence interval of 13.5 mg m^{-3}. An inflection point for probability corresponding to increased risk occurred at *chl-a* > 25 mg m^{-3}We repeated this analysis after removing *Alexandrium* to determine if this species with low alert level affected CPA outputs (Fig. 6b). The relationship between HAB abundance and *chl-a* was weaker at higher *chl-a*, but presence/absence of *Alexandrium* did not affect the baseline probability. Setting an *Alexandrium* alert level other than "present" had little effect on CPA outputs as mean abundance was ~8,000 cells L^{-1} (range: 100-290,000 cells L^{-1}), and an alert level of > 1,000 cells L^{-1} gave similar patterns.

The *chl-a* thresholds derived using CPA were consistent with relationships of HAB species to *chl-a* using quantile regressions or LAD, with a 0.50 probability of HAB corresponding to a broad range of *chl-a* from 3.5 - 40 mg m^{-3}. Low-biomass, highly toxic genera such as *Alexandrium* and *Dinophysis* occupied the low end of the *chl-a* range, while high-biomass genera such as *Heterosigma* and *Pseudo-nitzschia* occurred at the other high end. CPA for individual sub-embayments were affected by sample size with limited observations at high *chl-a*, but comparable thresholds were derived using spatially aggregated data. Exceptions included NCB and CB that showed flat relationships with *chl-a* (e.g., Fig. 7). Other sub-embayments showed increased probabilities of HAB occurrences with increasing *chl-a*, exceeding 0.80 at highest *chl-a* in SPB and SB. More than 90% of *chl-a* observations in NB and CB were < 13 mg m^{-3} for the 20-year record, and only in SB and LSB were *chl-a* commonly > 13 mg m^{-3} (18% and 26%, respectively; Supplemental Materials, Fig. S2).

OLS regressions of SPATT toxin on *chl-a* were not statistically significant, but CPA on toxins and *chl-a* gave similar inflection points as we derived for HAB organisms. The baseline probability for DA began at ~0.35 (i.e., across all *chl-a* levels) and increased to ~0.6 for observations with *chl-a* >13 mg m^{-3} (Fig. 8a). MCY showed a similar pattern, with a baseline probability of ~0.3 (Fig. 8b). Very few SPATT observations exceeded *chl-a* thresholds for HAB alert levels (>13 mg m^{-3}), but an increased probability of exceeding toxin thresholds at *chl-a* > 10 mg m^{-3} was consistent with the probability of exceeding alert levels for HAB abundance based on CPA (Fig. 6a-b).

Thresholds Relating DO to *Chl-a*. Quantile regression of mean *chl-a* from February-September and DO from May-July in SB and LSB showed consistently significant ($p < 0.1$), negative slopes for $\tau = 0.1$ and 0.5 using all three *chl-a* evaluation periods (Table 1). Slopes were slightly steeper and more significant for May-July then for June-August. Based on quantile regressions for SB using DO from May-July, a mean *chl-a* from February - September of 14 mg m^{-3} was associated with a low frequency of DO falling below the WQC for DO, while the likelihood was higher at *chl-a* of 17 mg m^{-3} (Table 2). Comparison of predicted *chl-a* values for a gradient of DO is instructive. At *chl-a* of 14 mg m^{-3}, 90% of DO observations were predicted to exceed 7 mg L^{-1}, while at *chl-a* of 42 mg m^{-3}, 90% of DO observations were predicted to exceed 5.0 mg L^{-1} (Table 2). For context, the February-September *chl-a* measured at SB sites was below 14 mg m^{-3} 95% of the time over the 20-year record (supplemental materials, Fig. S2).

DO in LSB was predicted to fall below DO benchmarks at lower *chl-a* than in SB, althought confidence intervals were larger. At a mean *chl-a* from February-September of 16 mg m^{-3}, there was an elevated risk of falling below WQC for DO based on the three-month median for percent saturation (Table 2). A 10% probability of exceeding the WQC for DO was associated with *chl-a* of 4 mg m^{-3}, with a negative lower 95th CI. This suggests advection of DO-depleted water into the study area such that even at extremely low values of *chl-a*, the probability of falling below the WQC for DO is high. Similarly, the CPA showed a baseline probability of 0.2 for falling below the WQC for DO (Fig. 9). This baseline was moderately high considering a mean probability of 0.5 based on the WQC for DO to *chl-a* > 14 mg m^{-3} with an upper 95% confidence interval of > 10 mg m^{-3}. We interpret this result to mean *chl-a* at or above these thresholds entails increased risk of DO below the WQC for DO of 80% saturation with increased *chl-a* (Fig. 9). Applying the CPA and comparing results to DO and *chl-a* distributions in SFB, we observed that 90% of DO values would exceed 6.3 mg L^{-1} and 5.0 mg L^{-1}, respectively, at *chl-a* of 15 mg m^{-3} and 36 mg m^{-3} (Table 2). Long-term data for 20 years showed 95% of *chl-a* measured in LSB was < 25 mg m^{-3} (Supplemental Materials, Fig. S2), and hypoxia associated with high *chl-a* remains uncommon in the open channel habitat of LSB.

Discussion

Current Status and Potential for Eutrophication in SFB

Humans have enriched the world's bays and estuaries with nitrogen and phosphorus, but the responses to enrichment vary widely across ecosystems (Cloern, 2001). Nutrient supply sets the potential for environmental degradation through excess production of algal biomass, but the realization of that potential – i.e., the efficiency with which exogenous nutrients are converted into biomass – depends on factors that regulate phytoplankton population growth, including light availability, toxins, grazing, pathogens, and transport processes. Nutrient concentrations in SFB exceed those that have led to degradation of water quality in other estuaries, but its phytoplankton biomass (mean *chl-a* concentration) is lower and DO concentrations higher than for other enriched estuaries such Chesapeake Bay, Neuse Estuary, Seine Bay, and the Westerschelde (Bricker et al., 2007; Cloern and Jassby 2008, Fig. 1). However, estuaries are highly dynamic ecosystems that exhibit complex responses to human disturbances, climate variability, and climate change (Cloern et al., 2015, Harding et al., 2015). Changes in SFB during the past two decades include significant increases of *chl-a*, ubiquitous presence of HAB species known to be toxic in other nutrient-enriched estuaries, and significant decreasing trends of DO.

HAB cell densities exceeded alert levels in ~35% of samples from SFB, indicating the potential for adverse effects on ecosystem health. HAB species are expected to occur at some baseline level, based on the cosmopolitan distributions of many species (Lundholm and Moestrop, 2006). However, the probability of a HAB event is high, once seeded, due to nutrient over-enrichment that characterizes SFB. The high baseline of occurrence documented in this study reflects strong connectivity with at least two documented sources of HAB seed populations. The first is the coastal ocean adjacent to SFB, a source of toxic phytoplankton species that lead to closures of shellfish harvesting half the year because of potential exceedances of alert levels based on HAB abundance (Lewitus et al., 2012). The second source is the South Bay salt ponds where the presence of dinoflagellates, *Alexandrium* spp. and *Karenia mikimotoi*, the raphidophyte, *Chattonella marina*, and the cyanophytes, *Anabaenopsis* spp. and *Anabaena* spp. has been confirmed (Thébault et al., 2008). Samples from SB contained other HAB species that were rare in SFB prior to the opening of the Salt Ponds, including *Karlodinium veneficum*, *Chattonella marina*, and *Heterosigma akashiwo*, while abundances of *K. mikimotoi* and *K. veneficum* in LSB and SB increased after breaching of the Salt Ponds. Distributions of these species show a spatial pattern reveal expansion into the rest of the SFB, suggesting that they pose an emerging threat.

While presence of HAB species above a defined alert level indicates a potential threat, the presence of toxins elevates that threat considerably as it demonstrates that environmental conditions within SFB or connected habitats are conducive to toxin production. SFB is not routinely monitored for algal toxins and no acute wildlife mortalities or human illnesses have been directly attributed to HAB from 1993-2014. MCY in SFB, however, has been linked to negative impacts on aquatic food webs (Lehman et al., 2010), and there is increasing evidence that chronic, sub-lethal exposure to DA constitutes a significant impairment (Goldstein et al., 2008; Montie et al., 2012). *Pseudo-nitzschia* exceeded alert levels in only 11% of samples, and cyanobacteria cells were not recorded, nonetheless, 72% and 96% of SPATT showed measurable quantities of the toxins MCY and DA, respectively. SPATT detects low concentrations of toxins compared to traditional methods (Lane et al., 2010; Kudela, 2011), and removal of SPATT data with the lowest toxin levels still left ~35% of samples with toxin levels of concern. These findings suggest that dissolved toxins are widely distributed in SFB. SPATT were not analyzed for other toxins that may occur in SFB, and the threat of HAB toxins remains requires further study.

Significant relationships of HAB abundance to *chl-a* were detected in SFB, while low DO and high *chl*-a were rarely observed and relationships differed by sub-embayment. In Northern SFB, DO was high and relatively low *chl-a* accompany depressed primary production with several possible causes, including inhibition and grazing (Dugdale et al., 2007; Cloern et al., 2014). The lack of consistent, significant relationships between DO and *chl-a* in SUB, SPB and CB sub-embayments suggests that physical processes, such as strong tidal mixing and a lack of persistent stratification, partially alleviate the development of low DO, despite high phytoplankton biomass (Smith and Hollibaugh, 2006). These modulating factors appear important in both CB and SUB, sub-embayments that are adjacent to the coastal ocean and the Sacramento/San Joaquin Delta, respectively. In contrast to CB and SUB sub-embayments, DO was lowest and *chl-a* was highest in LSB, a lagoonal sub-embayment with a long residence time that is near productive intertidal habitats that experience hypoxia, such as the restored salt ponds in SB (Thebault et al., 2008) and tidal sloughs (Senn et al., 2014).

SFB is responsive to both climate forcing and climate change (Cloern et al., 2015), and these factors can lead to shifts in the efficiency of nutrient assimilation into phytoplankton biomass, as reported for the Baltic Sea (Riemann et al., 2015). The high ambient nutrient concentrations that characterize SFB suggest a potential for accumulation of phytoplankton biomass sufficient to impair water quality. To evaluate this potential, we computed median concentrations of dissolved inorganic nitrogen (DIN) and *chl-a* for four sub-embayments (Table 3). We then estimated potential *chl-a* as the sum of measured *chl-a* plus the amount of *chl-a* that would be produced if all remaining DIN was assimilated into phytoplankton biomass, assuming a conversion factor of 1 g *chl-a* per mol N (Eppley et al., 1971). We found median *chl-a* in all sub-embayments of SFB would increase an order of magnitude if this potential was realized. Given uncertainty about the future trajectory of water quality in SFB, a potential for increased phytoplankton biomass justifies establishment of *chl-a* thresholds to support nutrient management directed at reducing risk of impairments.

Chl-a as the Basis to Assess Water-Quality Impairments

Chl-a is an integrative measure of water quality that has been used to assess eutrophication in estuaries around the world (Bricker et al., 2003; Zaldivar et al., 2008). Our analyses have related specific water-quality impairments in SFB to *chl-a*, consistent with published work that applies *chl-a* as a pivotal indicator of nutrient over-enrichment. We present several key findings that support this approach. First, we documented significant relationships between HAB abundance, DO and *chl-a* using quantile regressions. Our results are consistent with a conceptual model of increased risk for HAB abundance, toxins, and low DO at increased phytoplankton biomass (Cloern, 2001). Second, several statistical approaches yield consistent ranges for chl-a threshold based on HAB and DO. An inflection point at mean monthly *chl-a* < 13 mg m^{-3} was a threshold below which the probability of potentially deleterious conditions quantified by HAB abundance and SPATT-derived toxins decreased. This *chl-a* threshold was similar to mean seasonal *chl-a* of 13 - 16 mg m^{-3} associated with attainment of the WQC for DO in SFB, based on the three-month median percent saturation of 7 mg L^{-1}. At the opposite end of the risk continuum, inflection points of heightened risk of HAB cell density (*chl-a* from 25 - 40 mg m^{-3}) corresponded well to mean seasonal *chl-a* thresholds of 35-40 mg m^{-3} required for LSB and SB to fall more consistently below the 5.0 mg L^{-1} DO WQC. Third, *chl-a* thresholds we derived for SFB were in agreement with published water-quality criteria using a variety of assessment methods. Several examples are consistent with *chl-a* thresholds that we derived for SFB based on relationships with HAB and DO. Harding et al. (2014) reported that mean summer *chl-a* from 7.2 – 11 mg m^{-3} precluded low DO in the deep waters of Chesapeake Bay, and that mean annual *chl-a* of 15 mg m^{-3} was associated with decreased risk of *Microcystis* spp. toxins. Bricker et al. (2003) designated > 20 mg m^{-3} as a threshold of "high" risk for eutrophication, a value agreed upon by expert judgment. Similarly, *chl-a* thresholds of 10, 20 and 50 mg m^{-3} are used to define categories of low, high, and very high risk of eutrophication in the Phytoplankton Biological Quality Element for the European Union (EU) Water Framework Directive (WFD) proposed in the United Kingdom (Devlin et al., 2011).

"Risk Assessment" and Uncertainty in Chl-a Thresholds

Environmental management and regulation are firmly grounded in a paradigm of "risk assessment" (US Environmental Protection Agency, 1998). For this reason, risk represents a useful context to express *chl-a* thresholds and uncertainties with respect to WQC for SFB. Here, we used CPA and quantile regression to derive *chl-a* thresholds corresponding to low and high risk of exceeding HAB alert levels. A similar approach is commonly used to derive WQC for freshwater ecosystems, but few applications exist for the marine environment (Paul et al., 2005). Both CPA and quantile regression provide quantitative measures of uncertainty, a key element to support environmental decision-making (National Research Council, 2009).

CPA and quantile regression provided estimates of statistical uncertainty for *chl-a* thresholds based on HAB and DO, but other sources of uncertainty should be considered when applying these thresholds to nutrient management. First, the ecological significance of HAB species in SFB is not well known. Data needs include bio-accumulation of particulate and dissolved toxins in the biota, and acute and chronic impairments of ecosystem health. Such efforts should be coupled to an improved understanding of relationships between HAB toxins and *chl-a* specific to each sub-embayment. Second, spatial and temporal dynamics of low summer DO and seasonal maxima of *chl-a* that support DO consumption require additional study. Conceptually, it is possible that the mechanism behind this relationship is that high primary production on seasonal to annual time scales is expected to promote increased abundance of detritus, which, during the summer, leads to an increased probability of net ecosystem heterotrophy (Caffrey, 2003). Large spring blooms and subsequent fall blooms that were prominent features of the annual phytoplankton cycle in 2000 (Cloern et al., 2007) have not occurred in the past five years; in contrast, the summertime baseline that has seen the largest magnitude increase from 1993 to 2014 (Fig. 4). In-depth investigations into phytoplankton contribution to the SFB carbon budget and its relative influence on the coupling of pelagic and benthic metabolism are needed to better understand the relationships behind these empirical relationships between DO and *chl-a* in SB and LSB (e.g., Murrell et al., 2013).

Finally, there is a need to review the relevance and adequacy of scientific data supporting WQC for DO in SFB, specifically in LSB. Over the last 20 years, LSB has met 5.7 mg L^{-1}, the benchmark proposed by Best et al. (2007) that corresponds to the highest ecological condition category in EU estuaries. However, it has frequently not met the WQC based on three-month median percent DO saturation of > 80%, a value that at mean summer salinities and temperatures is equivalent to 7 mg L^{-1}. The question is whether 7 mg L^{-1} is a reasonable expectation for DO in LSB, given that this sub-embayment is strongly influenced by highly productive, intertidal habitats (Thebault et al., 2008; Shellenbarger et al., 2008).

Such investigations should be nested within an improved monitoring program, as the complexity of these patterns remind us that SFB is in a continuing state of change, one that is likely to continue over the next century (Cloern et al., 2011). Although it is attractive to consider relationships of impairments such as HAB abundance and low DO to *chl-a* as constant, we recognize that *chl-a* thresholds are responsive to changes in fundamental drivers of phytoplankton dynamics, such as oceanic exchange, top-down grazing, and light limitation. Changes in the relationships of impairments to *chl-a* will almost certainly respond to climate variability and climate change, as reported for this and other ecosystems (Cloern et al., 2014; Riemann et al., 2015).

Summary

This study demonstrated that, while DO is higher and *chl-a* lower in SFB than in other estuaries subject to nutrient over-enrichment, this important ecosystem is poised to express symptoms of cultural eutrophication. We found that evidence of ubiquitous HAB abundance, HAB toxins, declining DO, and increasing *chl-a*, supporting generalized conceptual models that describe increased risk of HAB cell densities and toxin concentrations and declining DO with increasing phytoplankton biomass. The majority of SFB subembayments are currently below *chl-a* < 13 mg m^{-3}, representative of baseline probabilities of HAB occurrence and attainment of SFB's 3-month median percent saturation DO WQC. However, SFB has sufficient dissolved inorganic nutrients to reach chl-a levels defined by "high risk" thresholds in the range of 25-40 mg m-3 *chl-a*, suggesting a potential for increased biomass accumulation that could lead to cultural eutrophication. Given the uncertainty in SFB's trajectory amidst global change, it is this potential for high biomass production that motivates establishment of *chl-a* water quality goals to support nutrient management of SFB, and underlines the need for continued monitoring of SFB to understand how these fundamental relationships may change in the future.

Acknowledgements

Funding for this study was provided through a contract with the San Francisco Water Quality Control Board (11-151-120). This work benefited from many discussions with the SFB Nutrient Technical Workgroup and Steering Committee.

References

Andersen, J. H., Murray, C., Kaartokallio, H., Axe, P., Molvær, J., 2010. Confidence rating of eutrophication status classification. Mar. Pollut. Bull. 60, 919–924.

Andersen, J. H., Carstensen, J., Conley, D. J., Dromph, K., Fleming-Lehtinen, V., Gustafsson, B.G., Josefson, A.B., Norkko, A., Villnas, A., Murray, C., 2015. Long-term temporal and spatial trends in eutrophication status of the Baltic Sea. Biol. Rev. doi: 10.1111/brv.12221

Anderson, D. M., Cembella, A. D., Hallegraeff, G. M., 2012. Progress in understanding harmful algal blooms: paradigm shifts and new technologies for research, monitoring, and management. Ann. Rev. Mar. Sci., 4, 143-176.

Bailey, H., Curran, C., Poucher, S., Sutula, M., 2014. Science supporting dissolved oxygen objectives for Suisun Marsh. Southern California Coastal Water Research Project Authority Technical Report 830. www.sccwrp.org. 33 p.

Barber, R. T., Hiscock, M. R., 2006. A rising tide lifts all phytoplankton: Growth response of other phytoplankton taxa in diatom-dominated blooms. Global Biogeochem. Cycles 20, GB4S03, doi:10.1029/2006GB002726.

Baustian, M. M., Rabalais, N. N., 2009. Seasonal composition of benthic macroinfauna exposed to hypoxia in the Northern Gulf of Mexico. Estuar. Coasts 32, 975-983.

Best, M. A., Wither, A. W., Coates, S., 2007. Dissolved oxygen as a physico-chemical supporting element in the Water Framework Directive, Mar. Pollut. Bull. 55, 53-64.

Bricker, S. B., Ferreira, J. G., Simas, T., 2003. An integrated methodology for assessment of estuarine trophic status. Ecol. Modell. 169, 39-60.

Bricker, S., Longstaff, B., Dennison, W., Jones, A., Boicourt, C., Wicks, C., Woerner, J., 2007. Effects of nutrient enrichment in the Nation's estuaries: a decade of change. NOAA Coastal Ocean Program Decision Analysis Series No. 26. National Centers for Coastal Ocean Science, Silver Spring, Maryland, USA. 328 p.

Cade, B. S., Noon, B. R., 2003. A gentle introduction to quantile regression for ecologists. Front. Ecol. Environ. 1, 412-420.

Caffrey, J. M., 2003. Production, respiration and net ecosystem metabolism in US estuaries. Environ. Mon. Assess. 81, 207-219.

Cloern, J. E., 2001. Our evolving conceptual model of the coastal eutrophication problem. Mar. Ecol. Prog. Ser. 210, 223-253.

Cloern, J. E., Cole, B. E., Hager, S. W., 1994. Notes on *Mesodinium rubrum* red tides in San Francisco Bay (California, USA). J. Plankton Res. 16, 1269–1276.

Cloern, J. E., Dufford, R., 2005. Phytoplankton community ecology: principles applied in San Francisco Bay. Mar. Ecol. Prog. Ser. 285, 11-28.

Cloern, J. E., Schraga, T. S., Lopez, C. B., Knowles, N., Labiosa, R. G., Dugdale, R., 2005. Climate anomalies generate an exceptional dinoflagellate bloom in San Francisco Bay. Geophys. Res. Lett. 32, L14608, doi:10.1029/2005GL023321.

Cloern, J., Jassby, A., Thompson, J., Hieb, K., 2007. A cold phase of the East Pacific triggers new phytoplankton blooms in San Francisco Bay. Proc. Natl. Acad. Sci. U.S.A. 104, 18,561-18,565.

Cloern, J. E., Jassby, A., 2008. Complex seasonal patterns of primary producers at the land-sea interface. Ecol. Lett. 11, 1294-303. doi: 10.1111/j.1461-0248.2008.01244.

Cloern, J. E., Knowles, N., Brown, L. R., Cayan, D., Dettinger, M. D., Morgan T. L., Schoellhamer, D. H., Stacey, M. T., van der Wegen, M., Wagner, R. W., Jassby, A., 2011. Projected Evolution of California's San Francisco Bay-Delta-River System in a Century of Climate Change. PLoS ONE 6, e24465. doi:10.1371/journal.pone.0024465

Cloern, J. E., Jassby, A. D., 2012. Drivers of change in estuarine-coastal ecosystems: discoveries from four decades of study in San Francisco Bay. Rev. Geophys. 50, RG4001, doi:10.1029/2012RG000397.

Cloern, J. E., Foster, S. Q., Kleckner, A. E., 2014. Phytoplankton primary production in the world’s estuarine-coastal ecosystems. Biogeosciences 11, 2477–2501. doi:10.5194/bg-11-2477-2014

Cloern, J. E., Abreu, P. C., Carstensen, J., Chauvaud, L., Elmgren, R., Grall, J., Greening, H., Johansson, J. O. R., Kahru, M., Sherwood, E. T., Xu, J., Yin, K., 2015. Human activities and climate variability drive fast-paced change across the world's estuarine-coastal ecosystems. Glob. Chang. Biol. doi:10.1111/gcb.13059

Devlin, M., Bricker, S., Painting, S., 2011. Comparison of five methods for assessing impacts of nutrient enrichment using estuarine case studies, Biogeochemistry 106, 177-205.

Diaz, R. J., Rosenberg, R., 1995. Marine benthic hypoxia: a review of its ecological effects and the behavioral responses of benthic macrofauna. Oceanogr. Mar. Biol. Ann. Rev. 33, 245-303.

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

Dugdale, R. C., Wilkerson, F. P., Hogue, V. E., Marchi, A., 2007. The role of ammonium and nitrate in spring bloom development in San Francisco Bay. Estuar. Coast. Shelf Sci. 73, 17-29.

Eppley, R. W., Carlucci, A. F., Holm-Hansen, O., Kiefer, D., McCarthy, J. J., Venrick, E., Williams, P. M., 1971. Phytoplankton growth and composition in shipboard cultures supplied with nitrate, ammonium, or urea as the nitrogen source. Limnol. Oceanogr. 16, 741-751.

Gibble, C., Kudela, R., 2014. Detection of persistent microcystins toxins at the land-sea interface in Monterey Bay, California. Harmful Algae 39, 146-153. doi:10.1016/j.hal.2014.07.004

Glibert, P. M., Anderson, D. M., Gentien, P., Graneli, E., Sellner, K. G., 2005. The global, complex phenomena of harmful algal blooms. Oceanography 18, 137–147.

Goldstein, T., Mazet, J. A. K., Zabka, T. S., Langlois, G., Colegrove, K. M., Silver, M., Bargu, S., Van Dolah, F., Leighfield, T., Conrad, P. A., Barakos, J., Williams, D. C., Dennison, S., Haulena, M., Gulland, F. M. D., 2008. Novel symptomatology and changing epidemiology of domoic acid toxicosis in California sea lions *Zalophus californianus*: an increasing risk to marine mammal health. Proc. Royal Soc. B: Biol. Sci. 275, 267-276.

Harding, L. W., Jr., Batiuk, R. A., Fisher, T. R., Gallegos, C. L., Malone, T. C., Miller, W. D., Mulholland, M. R., Paerl, H. W., Perry, E. S., Tango, P., 2014. Scientific bases for numerical chl-a criteria in Chesapeake Bay. Estuar. Coasts 37, 134-148.

Harding, L. W., Jr., Gallegos, C. L., Perry, E. S., Miller, W. D., Adolf, J. E., Mallonee, M. E., Paerl, H. W., 2015. Long-term trends of nutrients and phytoplankton in Chesapeake Bay. Estuar. Coasts. doi:10.1007/sl2237-015-0023-7.

Heisler, J., Glibert, P. M., Burkholder, J. M., Anderson, D. M., Cochlan, W., Dennison, W. C., Dortch, Q., Gobler, C. H., Heil, C. A., Humphries, E., Lewitus, A., Magnien, R., Marshall, H. G., Sellner, K. G., Stockwell, D. A., Suddleson, M., 2008. Eutrophication and harmful algal blooms: a scientific consensus. Harmful Algae 8, 3-13.

Hollister J., Walker, H., Paul, J. F., 2008. CProb: A computational tool for conducting conditional probability analysis. J. Environ. Qual. 37, 2392-2396.

Jassby, A.D., Cole, B. E., Cloern, J. E., 1997. The design of sampling transects for characterizing water quality in estuaries. Estuar. Coast. Shelf Sci. 45, 285–302.

Kimmerer, W. J., Thompson J. K., 2014. Phytoplankton growth balanced by clam and zooplankton grazing and net transport into the low-salinity zone of the San Francisco Estuary. Estuar. Coasts 37, 1202-1218.

Kirkpatrick, B., Fleming, L. E., Squicciarini, D., Backer, L. C., Clark, R., Abraham, W., Benson, J., Cheng, Y. S., Johnson, D., Pierce, R., Zaias, J., Bossart, G., Baden, D. G., 2004. Literature review of Florida red tide: Implications for human health effects. Harmful Algae 3, 99–115.

Kudela, R. M., 2011. Characterization and deployment of Solid Phase Adsorption Toxin Tracking (SPATT) resin for monitoring of microcystins in fresh and saltwater. Harmful Algae 11, 117-125.

Lane, J. Q., Roddam, C. M., Langlois, G. W., Kudela, R. M., 2010. Application of Solid Phase Adsorption Toxin Tracking (SPATT) for field detection of domoic acid and saxitoxin in coastal California. Limnol. Oceanogr. Methods 8, 645–660.

Lehman, P. W., Boyer, G., Hall, C., Waller, S., Gehrts, K., 2005. Distribution and toxicity of a new colonial *Microcystis aeruginosa* bloom in the San Francisco Bay Estuary, California. Hydrobiol. 541, 87-99.

Lehman, P. W., Teh, S. J., Boyer, G. L., Nobriga, M., Bass, E., Hogle, C., 2010. Initial impacts of *Microcystis* on the aquatic food web in the San Francisco Estuary. Hydrobiol. 637, 229–248.

Lewitus, A. J., Horner, R. A., Caron, D. A., Garcia-Mendoza, E., Hickey, B. M., Hunter, M., Huppert, D. D., Kudela, R. M., Langlois, G. W., Largier, J. L., 2012. Harmful algal blooms along the North American west coast region: History, trends, causes, and impacts. Harmful Algae 19, 133-159.

Lundholm, N, Moestrop, O., 2006. The biogeography of harmful algae. In: Ecology of Harmful Algae (Granelli, E., Turner, J. T. Eds.), Ecological Studies 189, Springer.

Mackenzie, L., Beuzenberg, V., Holland, P., Mcnabb, P., Selwood, A., 2004. Solid phase adsorption toxin tracking (SPATT): a new monitoring tool that simulates the biotoxin contamination of filter feeding bivalves. Toxicon 44, 901-918. doi:10.1016/j.toxicon.2004.08.020.

Montie, E. W., Wheeler, E., Pussini, N., Battey, T. W. K., Van Bonn, W., Gulland, F., 2012 Magnetic resonance imaging reveals that brain atrophy is more severe in older California sea lions with domoic acid toxicosis. Harmful Algae 20, 19-29.

Murrell, M. C., Stanley, R., Lehrter, J. C., Hagy, J. D., 2013. Plankton community respiration, net ecosystem metabolism, and oxygen dynamics on the Louisiana continental shelf: Implications for hypoxia, Cont. Shelf Res. 52, 27-37.

Nichols, F. H., Cloern, J. E., Luoma, S. N., Peterson, D. H., 1986. The modification of an estuary. Science 231, 567–573.

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

National Research Council, 2009. Science and decisions: advancing risk assessment. Committee on Improving Risk Analysis Approaches Used by U.S. EPA, Board on Environmental Studies and Toxicology. doi: 10.17226/12209.

Paerl, H. W., 1995. Coastal eutrophication in relation to atmospheric nitrogen deposition: Current perspectives. Ophelia 41, 237-259.

Paul, J. F., McDonald, M. E., 1997. Development of empirical, geographically-specific water quality criteria: a conditional probability analysis approach. J. Am. Water Resour. Assoc. 41, 1211-1223.

Rabalais, N. N., Cai, W. -J., Carstensen, J., Conley, D. J., Fry, B., Hu, X., Quiñones-Rivera, z., Rosenberg, R., Slomp, C. P., Turner, R. E., Voss, M., Wissel, B., Zhang, J., 2014. Eutrophication-driven deoxygenation in the coastal ocean. Oceanography 27, 172–183.

Riemann, B., Carstensen, J., Dahl, K., Fossing, H., Hansen, J. W., Jakobsen, H. H., Josefson, A. A., Krause-Jensen, D., Markager, S., Stæhr, P.A., Timmermann, K., Windolf, J., Andersen, J. H., 2015. Recovery of Danish coastal ecosystems after reductions in nutrient loading: A holistic ecosystem approach. Estuar. Coasts doi: 10.1007/s12237-015-9980-0.

Rosenberg, R., Hellman, B. Johansson, B., 1991. Hypoxic tolerance of marine benthic fauna Mar. Ecol. Prog. Ser. 79, 127-131.

Schaeffer, B. A., Hagy, J. D., Conmy, R. N., Lehrter, J. C., Stumpf, R. P., 2012. An approach to developing numeric water quality criteria for coastal waters using the SeaWiFS satellite data record. Environ. Sci. Technol. 46, 916-922.

Shellenbarger, G. G., Schoellhamer, D. H., Morgan, T. L., Takekawa, J. Y., Athearn, N. D., Henderson, K. D., 2008. Dissolved oxygen in Guadalupe Slough and Pond A3W, South San Francisco Bay, California, August and September 2007. U.S. Geological Survey Open-File Report 2008–1097, 26 p.

Shutler, J. D., Davidson, K., Miller, P. I., Swan, S. C., Grant, M. G., Bresnan, E., 2012. An adaptive approach to detect high-biomass algal blooms from EO chl-a-*a* data in support of harmful algal bloom monitoring. Rem. Sens. Lett. 3, 101-110.

Smith S.V., Hollibaugh, J. T., 2006. Water, salt, and nutrient exchanges in San Francisco Bay. Limnol. Oceanogr. 51, 504–517.

State Water Resource Water Quality Control Board (SFRWQCB), 2011. San Francisco Bay Basin (Region 2) Water Quality Control Plan. As amended December 2011, available at www.waterboards.ca.gov/rwqcb2/basin_planning.shtml.

Sutula, M., Senn, D., 2016. Scientific bases for assessment of nutrient impacts on San Francisco Bay. Southern California Coastal Water Research Project Authority Technical Report 864. www.sccwrp.org. 56 pp.

Tett, P., Gowen, R., Mills, D., Fernandes, T., Gilpin, L., Huxham, M., Kennington, K., Read, P., Service, M., Wilkinson, M., Malcolm, S., 2007. Defining and detecting undesirable disturbance in the context of marine eutrophication. Mar. Poll. Bull. 55, 282-297.

Thébault, J., Schraga, T. S., Cloern, J. E., Dunlavey, E. G., 2008. Primary production and carrying capacity of former salt ponds after reconnection to San Francisco Bay. Wetlands 28, 841-851.

Topping, B. R., Kuwabara, J.S., Athearn, N.D., Takekawa, J.Y., Parchaso, F., Henderson, K.D., Piotter, S., 2009. Benthic oxygen demand in three former salt ponds adjacent to south San Francisco Bay, California. U.S. Geological Survey Open-File Report 2009-1180, 21 p.

U.S. Environmental Protection Agency, 1998. Guidelines for Ecological Risk Assessment. EPA/630/R-95/002F.

Vlamis A. and P. Katikou. 2014. Climate influence on Dinophysis spp. spatial and temporal distributions in Greek coastal water. Plankton Benthos Res 9(1): 15–31.

Wheeler, P.A., Huyer, A., Fleischbein, J., 2003. Cold halocline, increased nutrients and higher chl-a off Oregon in 2002. Geophys. Res. Lett. 30, doi: 10.1029/2003GL017395. issn: 0094-8276.

Zaldivar, J. M., Viaroli, P., Newton, A., De Wit., R., Ibanez, C., Reizopoulou, S., Somma, F., Razinkovas, A., Basset, A., Holmer, M., Murray, N., 2008. Eutrophication in transitional waters: an overview, Trans. Waters Mono. 1, 1-78.

Figure Legends

Fig. 1. SFB showing distribution of habitat types and locations of sub-embayments Suisun Bay (SUB), San Pablo Bay (SBP), North Central Bay (NCB), Central Bay (CB), South Bay (SB), and Lower South Bay (LSB), defined by Jassby et al. (1997), relative to the locations of major cities in region.

Fig. 2. Time series of major HAB in SFB from 1993-2014. Symbols indicate cell densities (cells mL^{-1}) by cruise. The station with the highest cell density is indicated for cruises with HAB enumerated at multiple locations. Inset values give cell densities at stations > 200 cells mL^{-1}.

Fig. 3. Concentration of (a) DA (ng g^{-1}) and (b) MCY (LR, RR, YR, and LA in ng g^{-1}) from SPATT deployed in the R/V *Polaris* surface mapping system for regions representing the following sub-embayments: SUB+ Delta station, SPB, NCB, and SB+CB during full Bay cruises, and LSB+SB during South SFB only cruises sub-embayments. Circles indicate DA (top) or MCY concentrations (bottom); for DA > 400 ng g^{-1} and MCY > 10 ng g^{-1}, and numeric values indicate the concentrations.

Fig. 4. Monthly geomean and 95% CI of *chl-a* over the periods from 1993-1999 and 2000-2014, by subembayment, from north to south, (a) SUB, (b) SPB, (c) NCB, (d) CB, (e) SB, and (f) LSB. Comparison of *chl-a* before and after 1999 is important temporal benchmark as Cloern et al. (2007) identified a *chl-a* step change coincident with the shifting of the NE Pacific to its cool phase.

Fig. 5. Comparison of quantile regressions relating May-July DO percent saturation to *chl-a* in selected sub-embayments from north to south: (a) SUB, (b) SPB, (c) SB and (d) LSB. Lines for the 10^{th} (τ=0.1, red) and median quantiles (τ=0.5, blue) are shown for the quantile regressions. Results of regression analyses are given in Table 1.

Fig. 6. Probability of HAB cell densities higher than alert levels as specified value of *chl-a* is exceeded for data in which (a) all HAB species are included and (b) excluding *Alexandrium*. The black line represents mean probability. Grey dashed lines are lower and upper 95% confidence intervals of bootstrap values (100 iterations).

Fig. 7. Mean probability of observing any HAB species at concentrations higher than defined alert levels if specified value of *chl-a* is exceeded, by sub-embayment for CB (open squares), SB (black triangle), and LSB (grey circle).

Fig. 8. Probabilities of DA (top panel) or MCY (bottom panel) > 75 ng g^{-1} and 1 ng g^{-1}, respectively, indicating risk when specified values of *chl-a* are exceeded. The black line represents mean probability. Dashed lines are lower and upper 95% confidence intervals from bootstrap (100 iterations).

Fig. 9. Probability of DO percent saturation < 80% during the months of June-August in LSB as specified value of February – September mean *chl-a* is exceeded. The black line represents mean probability. Grey dashed lines are lower and upper 95% confidence intervals of bootstrap values (100 iterations).

Figure Legends (Supplemental)

Fig. S1. Cumulative frequency distribution of minimum monthly DO by sub-embayment stations.

Fig. S2. Cumulative frequency distribution of (left panel) annual calendar mean and February-September *chl-a* for South and Lower South Bays as a proportion of site -years of 1993-2013 and (right panel) monthly *chl-a* by all sub-embayments as a proportion of site-years.

Table 1. Slopes of quantile regressions at Tau= 0.1 and 0.5 by DO integrating period (May-July and June-August) and chlorophyll-a averaging period (February-May, June-September, and February-September mean chlorophyll-a). * designates p< 0.1, ** designates p<0.05 and * designates P< 0.01.**

Subembayment and DO Integrating Period	**Slope of Quantile Regressions and Significance Level**					
	February-May Mean		**June-September Mean**		**February-September Mean**	
	0.1 Tau	**0.5 Tau**	**0.1 Tau**	**0.5 Tau**	**0.1 Tau**	**0.5 Tau**
May-July						
Lower South	0.06	-0.04	-0.22	-0.62**	-0.73**	-0.61**
South	-0.38***	-0.28***	-0.17	-0.58***	-0.78***	-0.73***
Central	-0.43	0.01	2.15***	0.74**	-0.73	0.15
North Central	-0.20	0.14	1.18	0.87	-0.84	0.85
San Pablo	-0.36	-0.44	-0.93	-0.58***	-0.77	-0.37
Suisun	-0.85	-0.57	-0.86	-0.45	-1.99***	-0.16
June-August						
Lower South	-0.14	-0.23***	0.62	0.39	-0.14	-0.20
South	-0.29***	-0.17***	0.16	-0.02	-0.60***	-0.39***
Central	-0.47	-0.10	0.74*	0.70**	0.27	0.27
North Central	-0.25	-0.13	0.98	0.60	0.39	0.39

San Pablo	-0.20	-0.11	-0.11	-0.36*	-0.33	-0.33
Suisun	0.02	0.05	-1.08	-0.82	0.49	-0.49

4

Table 2. Comparison of mean and 95% CI (in parentheses) of predicted *chl-a* (mg m^{-3}) from quantile regressions of February –September mean *chl-a* and May-July DO for specified DO benchmarks. 80% saturation at a τ= 0.5 is equivalent to SFB's percent saturation WQC. Predicted *chl-a* at τ= 0.1 represent a 10% frequency of falling below a gradient of DO benchmarks from the literature (i.e. 80%, 72%, 66% and 57% saturation, with corresponding to DO concentrations at mean summertime temperature of 15ºC and salinity of 24 ppt in SB and LSB). All regressions were significant for $p< 0.05$ (Table 1).

DO Percent saturation, with Equivalent DO Concentration	Predicted Mean *Chl-a* (95% CI)	
	LSB (N=48)	SB (N=161)
τ= 0.5		
80% (~7 mg L^{-1})	15.6 (9.2 – 21.8)	17.3 (15.1 – 19.5)
τ = 0.1		
80% (~ 7.0 mg L^{-1})	4.3 (-4.1 – 12.1)	14.3 (12.6 – 15.5)
72% (~6.3 mg L^{-1})	15.3 (5.3 – 29.3)	24.6 (21.9 – 24.7)
66% (~5.7 mg L^{-1})	23.5 (15.3 – 39.3)	32.3 (29.5 – 32.3)
57% (~ 5.0 mg L^{-1})	35.8 (30.3 – 54.3)	43.8 (40.5 – 45.9)
46% (~4.0 mg L^{-1})	50.9 (41.4 – 60.4)	57.9 (56.2 – 59.2)

Table 3. Median values of dissolved inorganic nitrogen (DIN), measured *chl-a* concentration, and potential *chl-a* concentration if all DIN was assimilated into additional phytoplankton biomass. Data from the USGS SFB water-quality measurement program for years 2000-2014.

Sub-embayment	DIN (μM)	Measured *Chl-a* (mg m^{-3})	Potential *Chl-a* (mg m^{-3})
SUB	36.9	2.5	39.7
SPB	29.0	3.8	33.6
SB	31.4	5.5	39.2
LSB	57.5	7.5	67.0

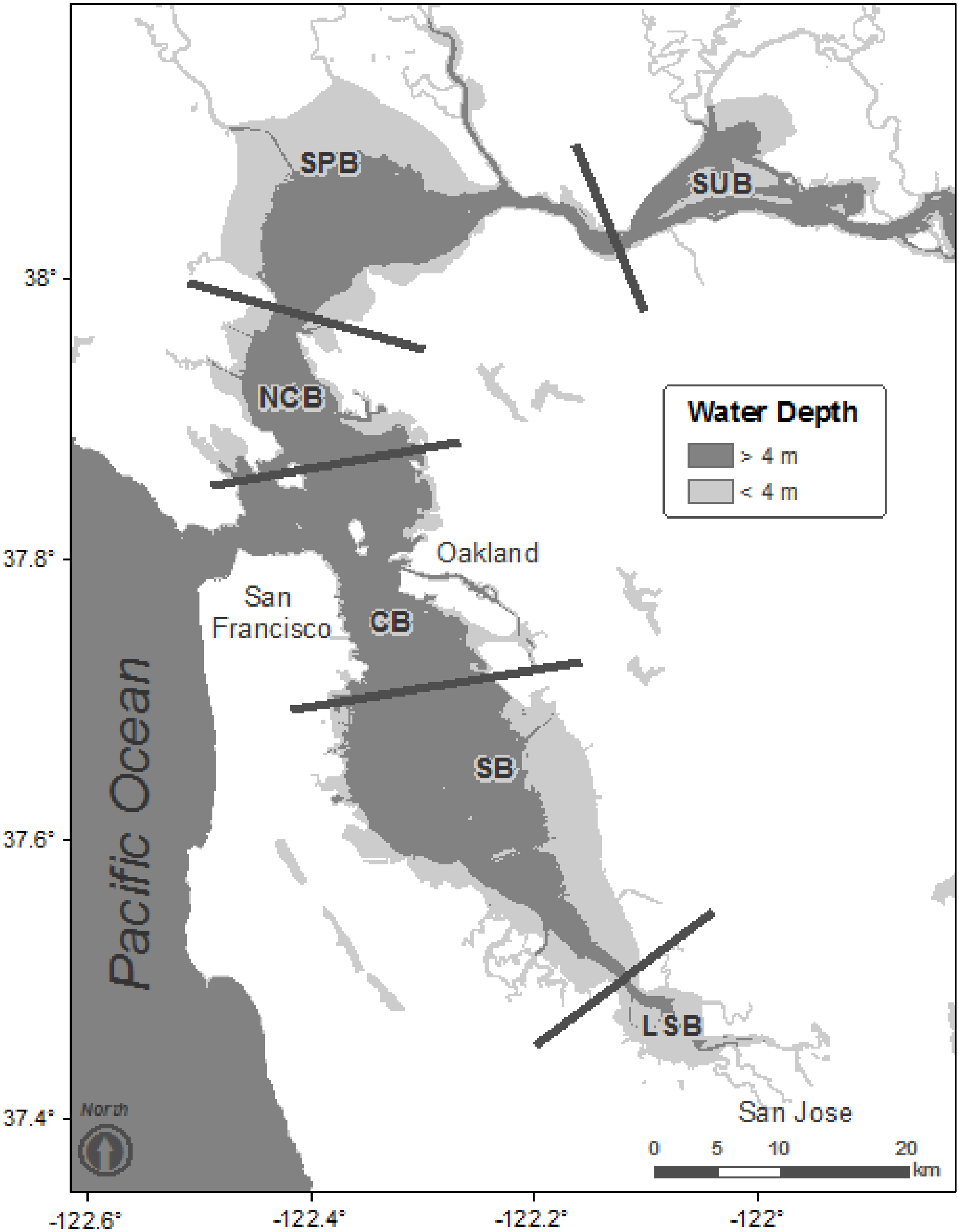
SUPPLEMENTAL MATERIALS

Table A1. **10^{th} Percentile of the vertical median and minimum summer (May-August) DO concentration over the period of 1993-2014 and percentage of time over that period that DO concentration was less than 5 mg L^{-1}.**

Sub-embayment	**10th Percentile of Summer Vertical Median DO (mg L^{-1})**	**10th Percentile of DO Summer Vertical Minimum (mg L^{-1})**	**% of Time Summer DO< 5 mg L^{-1}**
Lower South Bay	5.7	5.6	2.9
South Bay	5.9	5.8	0.5
Central Bay	6.5	6.5	0.2
North Central Bay	6.8	6.4	1.9
San Pablo Bay	7.1	7	0
Suisun Bay	7.8	7.7	0

Table A4. Percent of site-events that fell below DO objectives of 3 month median< 80% Saturation. Julian Day designates 3-month DO median aggregating period (e.g. Days 120-210 are May-July and 150-240 are June-August). N= Total number of site events.

Julian Day	Lower South		South		Central		No. Central		San Pablo		Suisun	
	N	%	N	%	N	%	N	%	N	%	N	%
30-120	**63**	0%	210	0%	82	0%	60	2%	126	1%	96	0%
60-150	**63**	0%	210	0%	82	0%	60	8%	126	4%	96	0%
90-180	**63**	6%	210	0%	82	2%	60	13%	126	3%	96	0%
120-210	**61**	18%	201	0%	80	5%	59	15%	126	1%	100	1%
150-240	**54**	28%	189	1%	77	0%	58	0%	126	0%	97	0%
180-270	**53**	13%	176	0%	73	0%	60	0%	126	0%	94	0%
210-300	**50**	6%	166	0%	68	0%	55	0%	117	0%	89	0%



SPB
SUB
NCB
Water Depth
> 4 m
< 4 m
Oakland
San
Francisco
CB
Pacific Ocean
SB
LSB
San Jose
North
0
5
10
20
km
38°
37.8°
37.6°
37.4°
-122.6°
-122.4°
-122.2°
-122°

Fig. 1

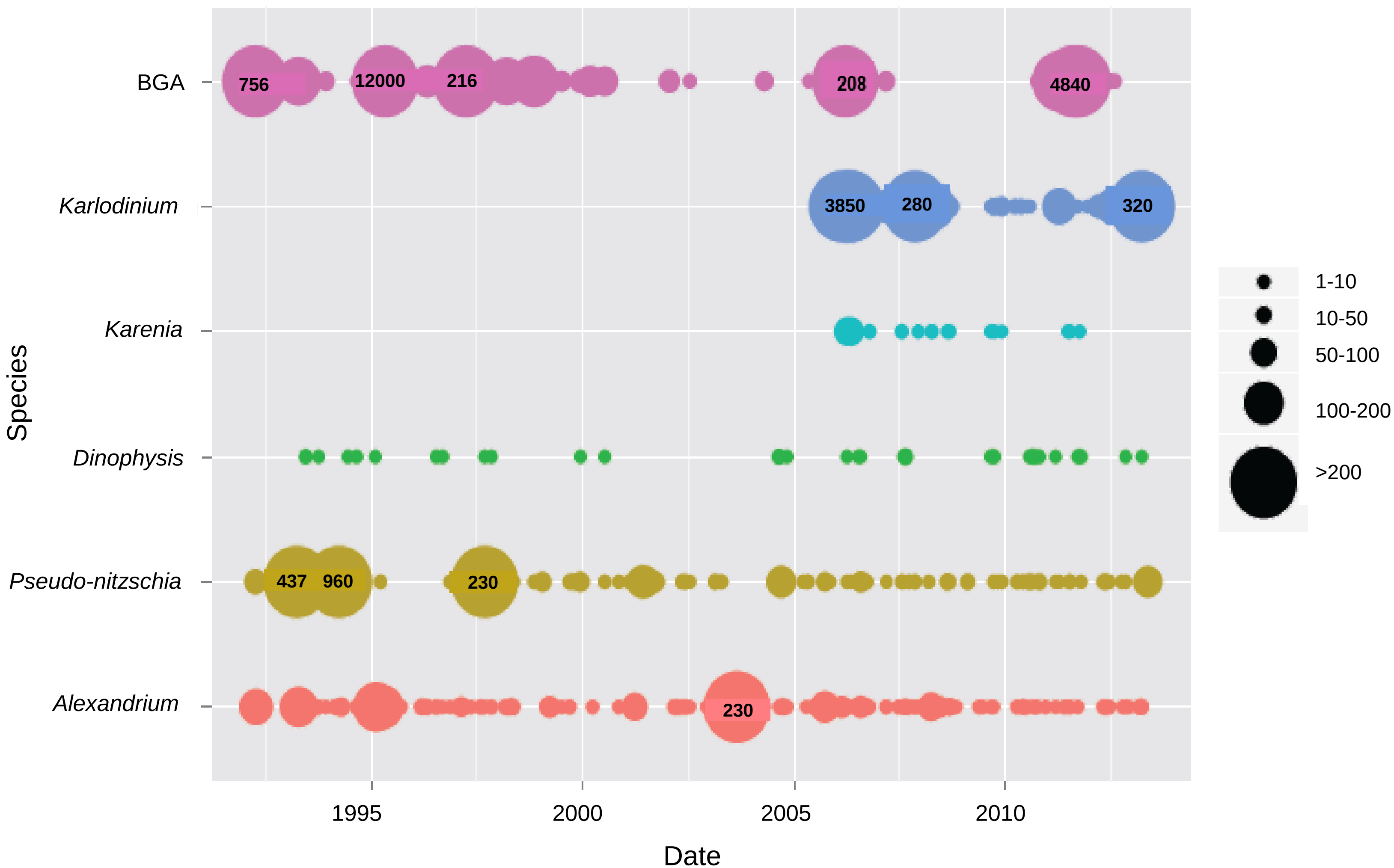

BGA
Karlodinium
Karenia
Dinophysis
Pseudo-nitzschia
Alexandrium
Species
756
12000
216
208
4840
3850
280
320
437
960
230
230
1995
2000
2005
2010
Date
1-10
10-50
50-100
100-200
>200

Fig. 2

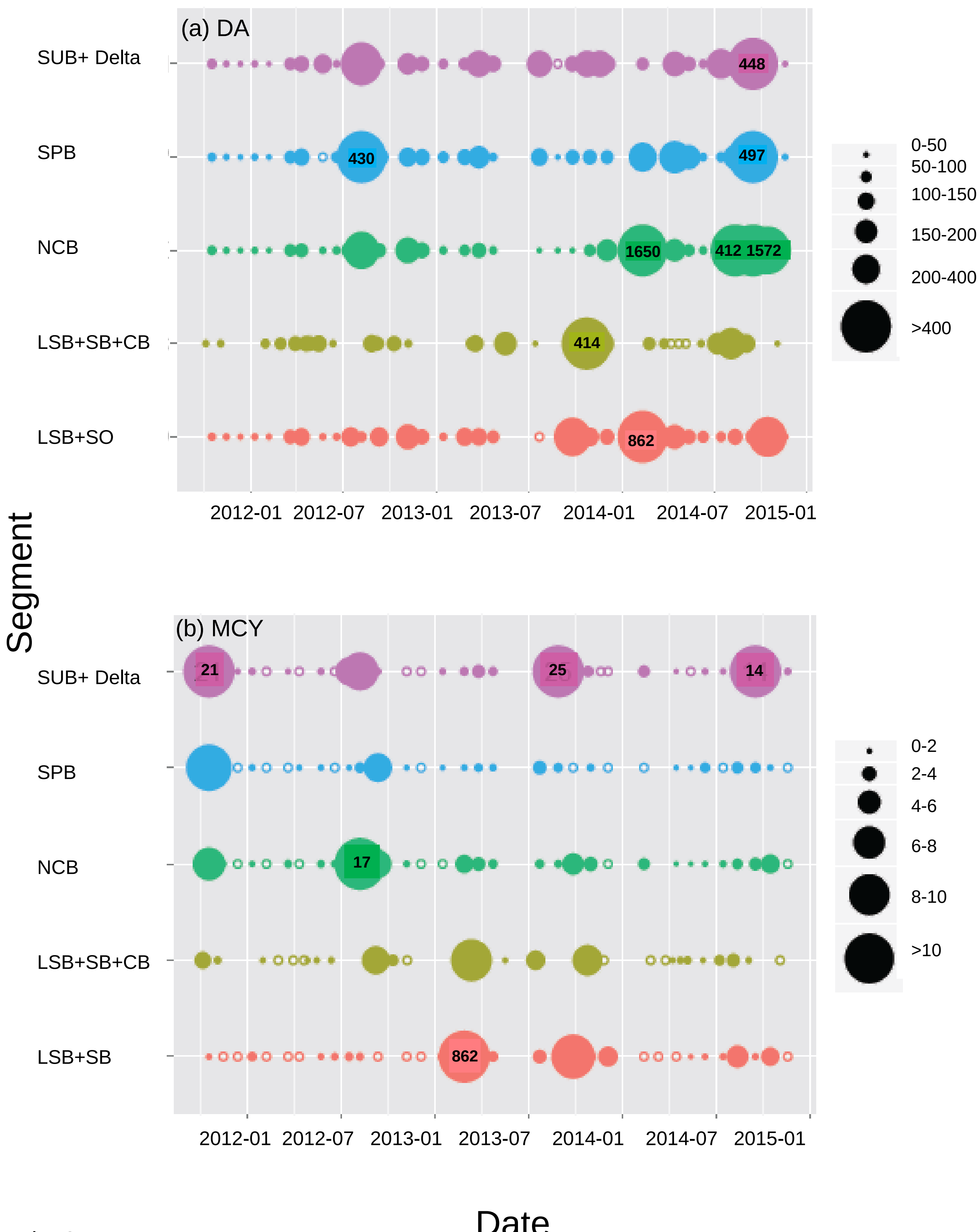

(a) DA
SUB+ Delta
SPB
NCB
LSB+SB+CB
LSB+SO
448
430
497
1650
412 1572
414
862
2012-01
2012-07
2013-01
2013-07
2014-01
2014-07
2015-01
0-50
50-100
100-150
150-200
200-400
>400
Segment
(b) MCY
SUB+ Delta
SPB
NCB
LSB+SB+CB
LSB+SB
21
25
14
17
862
2012-01
2012-07
2013-01
2013-07
2014-01
2014-07
2015-01
0-2
2-4
4-6
6-8
8-10
>10
Date

Fig. 3

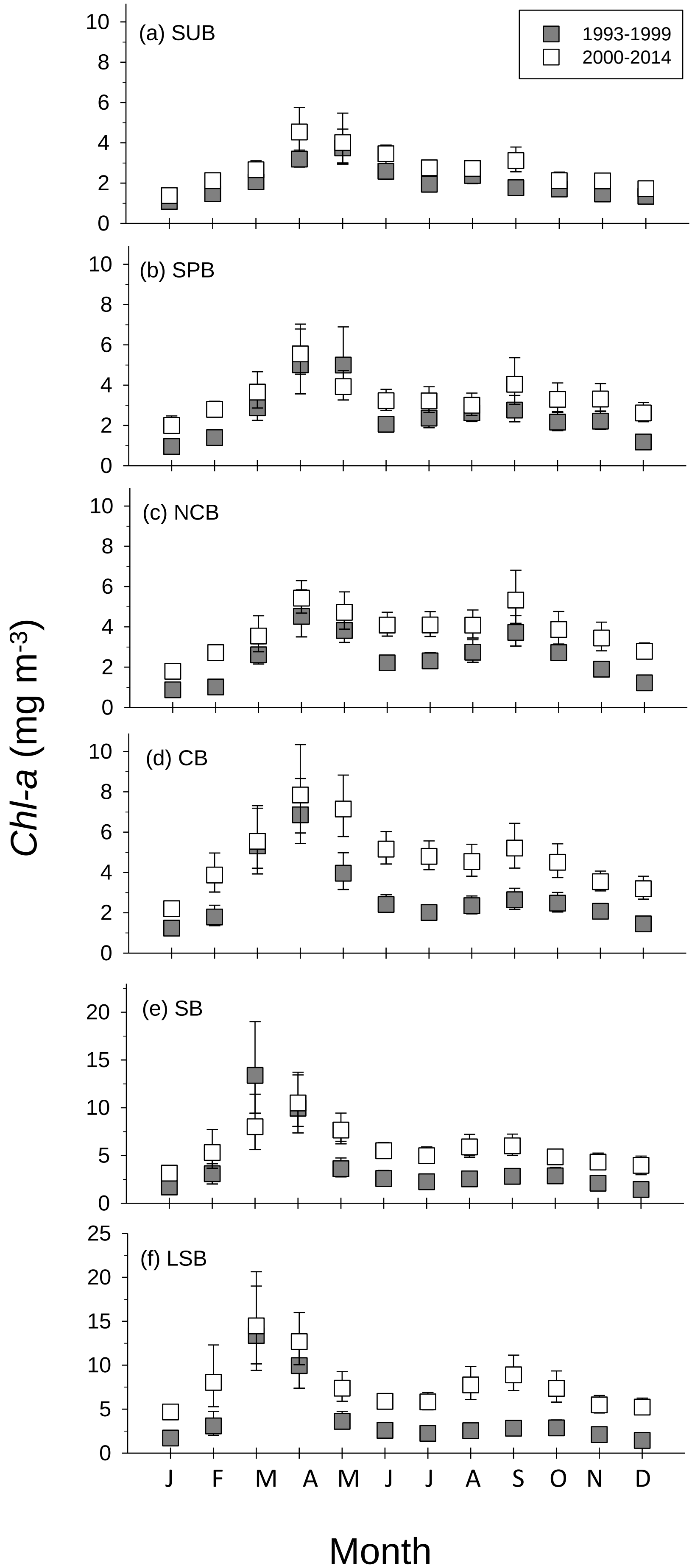

(a) SUB
1993-1999
2000-2014
(b) SPB
(c) NCB
(d) CB
(e) SB
(f) LSB
Chl-a (mg m-3)
J F M A M J J A S O N D
Month

Fig. 4

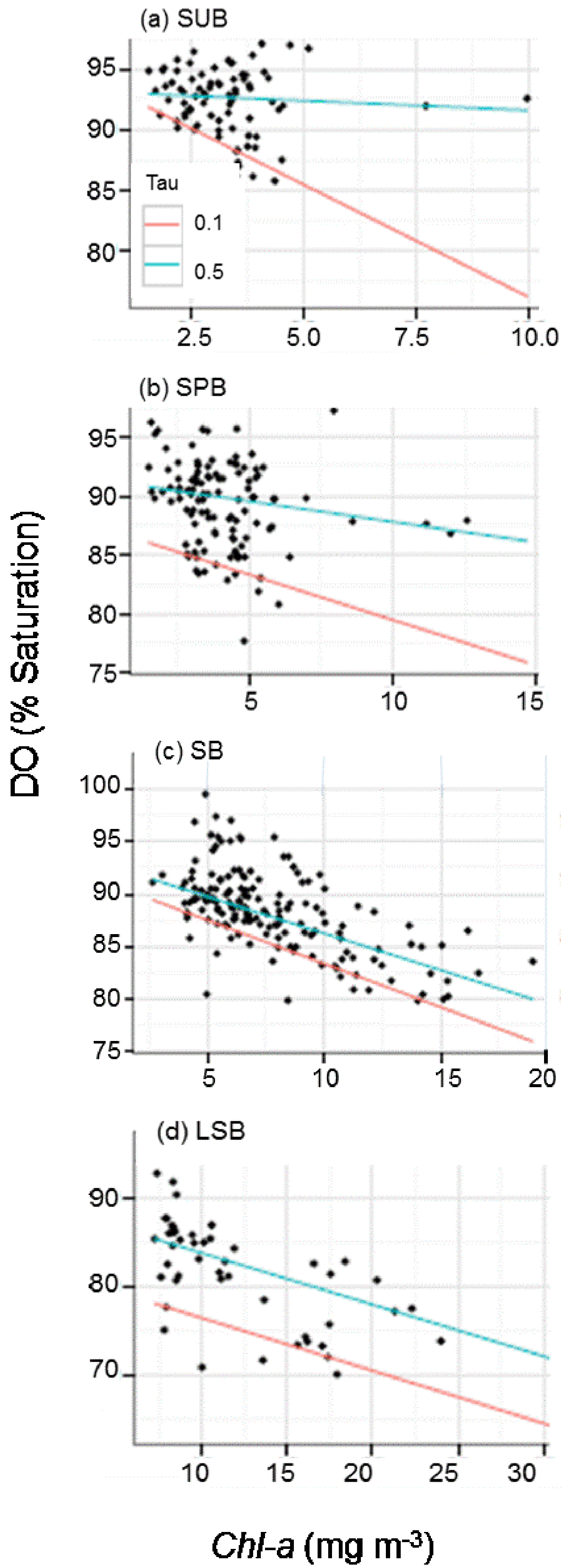

(a) SUB
95
90
85
80
Tau
0.1
0.5
2.5
5.0
7.5
10.0
(b) SPB
95
90
85
80
75
5
10
15
(c) SB
100
95
90
85
80
75
5
10
15
20
(d) LSB
90
80
70
10
15
20
25
30
DO (% Saturation)
Chl-a (mg m-3)

Fig. 5

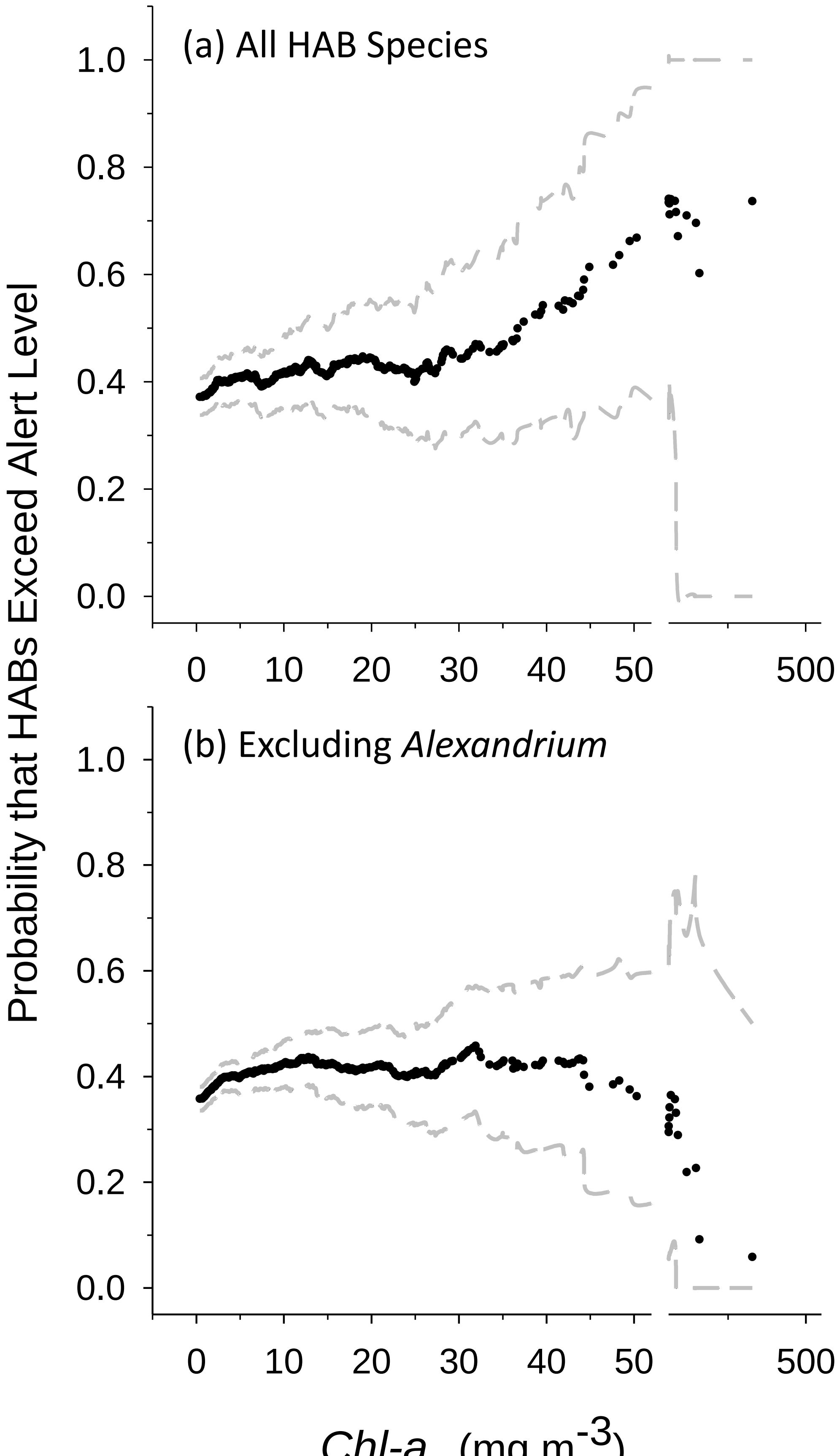
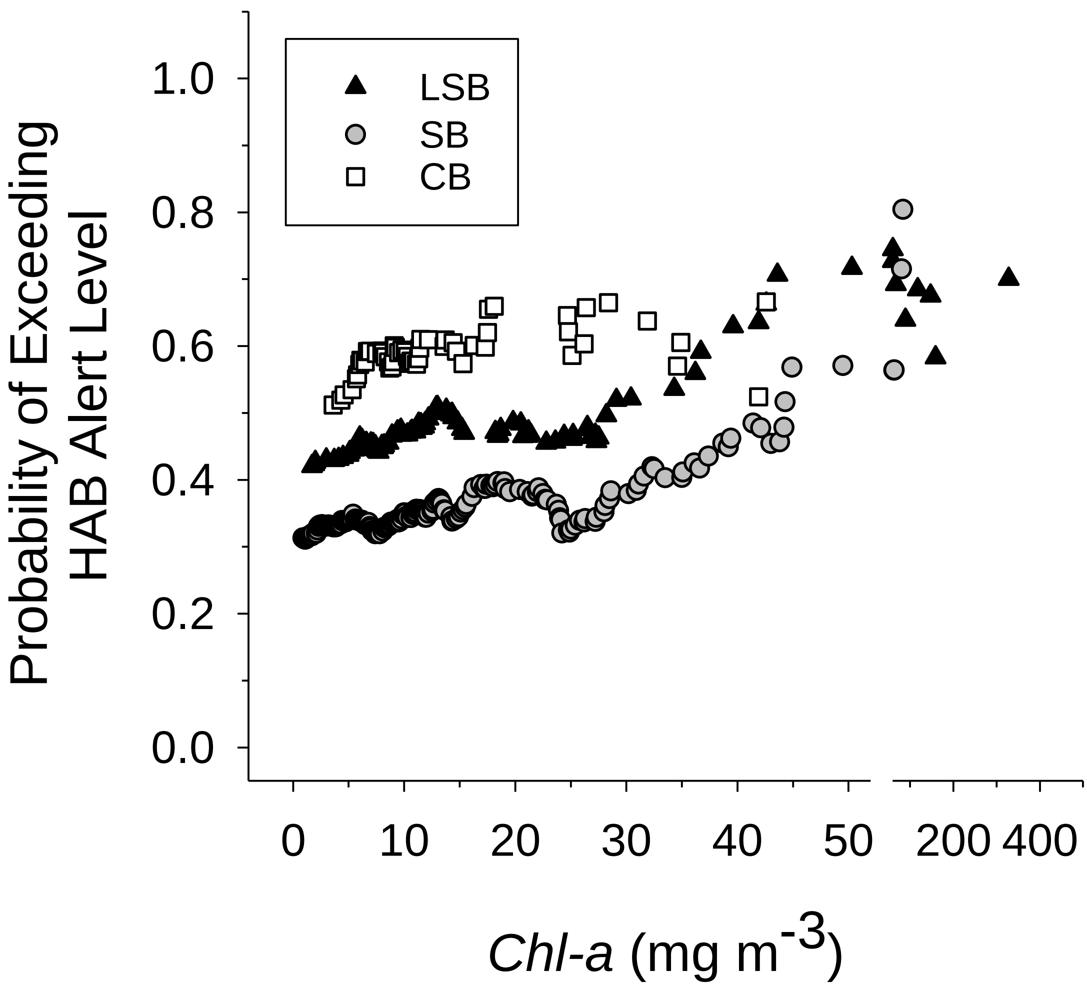

(a) All HAB Species
(b) Excluding *Alexandrium*
Probability that HABs Exceed Alert Level
1.0
0.8
0.6
0.4
0.2
0.0
0
10
20
30
40
50
500
Chl-a (mg m-3)

Fig. 6



Probability of Exceeding
HAB Alert Level
1.0
0.8
0.6
0.4
0.2
0.0
LSB
SB
CB
0
10
20
30
40
50
200
400
Chl-a (mg m-3)

Fig. 7

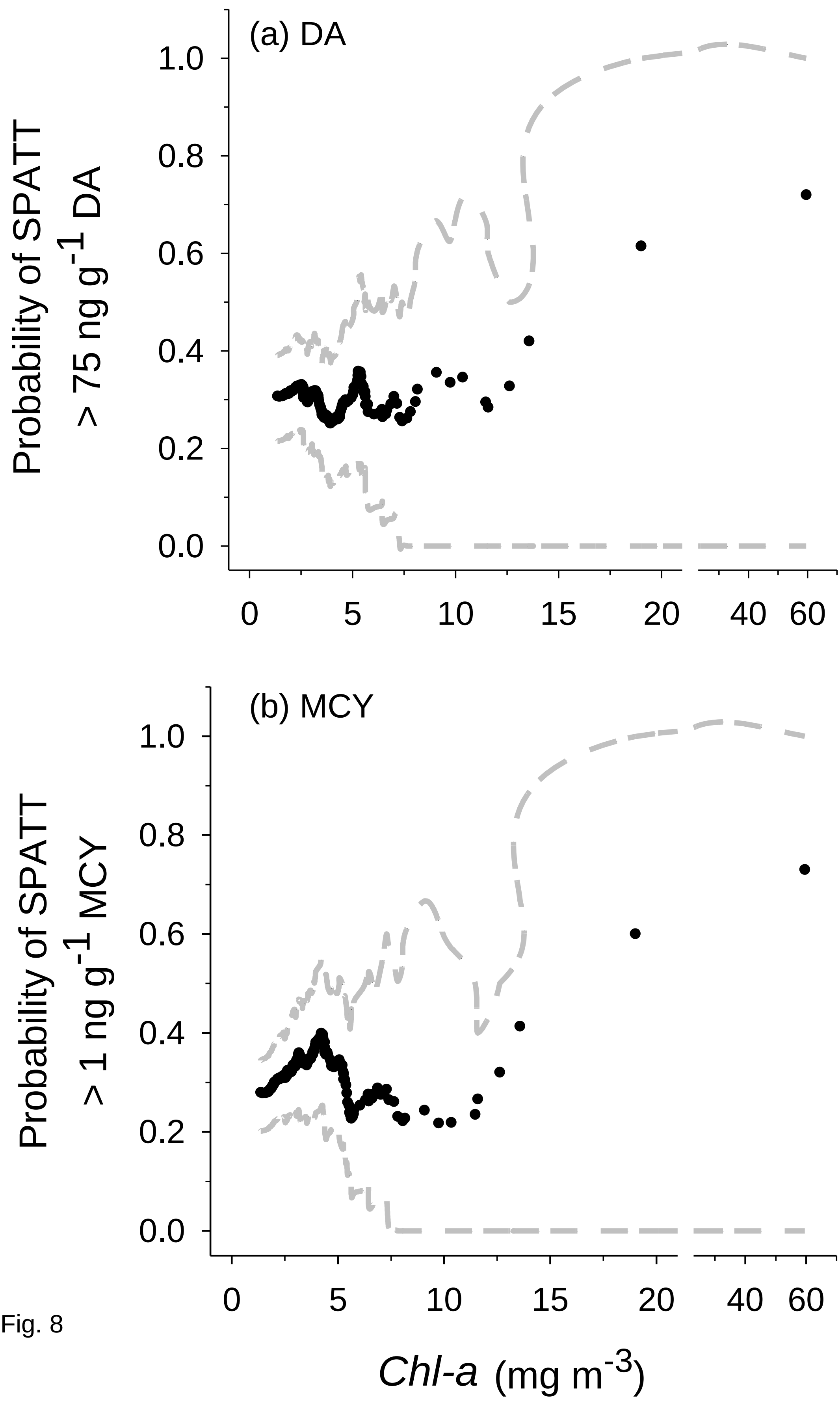

(a) DA
Probability of SPATT
> 75 ng g-1 DA
1.0
0.8
0.6
0.4
0.2
0.0
0
5
10
15
20
40
60
(b) MCY
Probability of SPATT
> 1 ng g-1 MCY
1.0
0.8
0.6
0.4
0.2
0.0
0
5
10
15
20
40
60
Chl-a (mg m-3)

Fig. 8

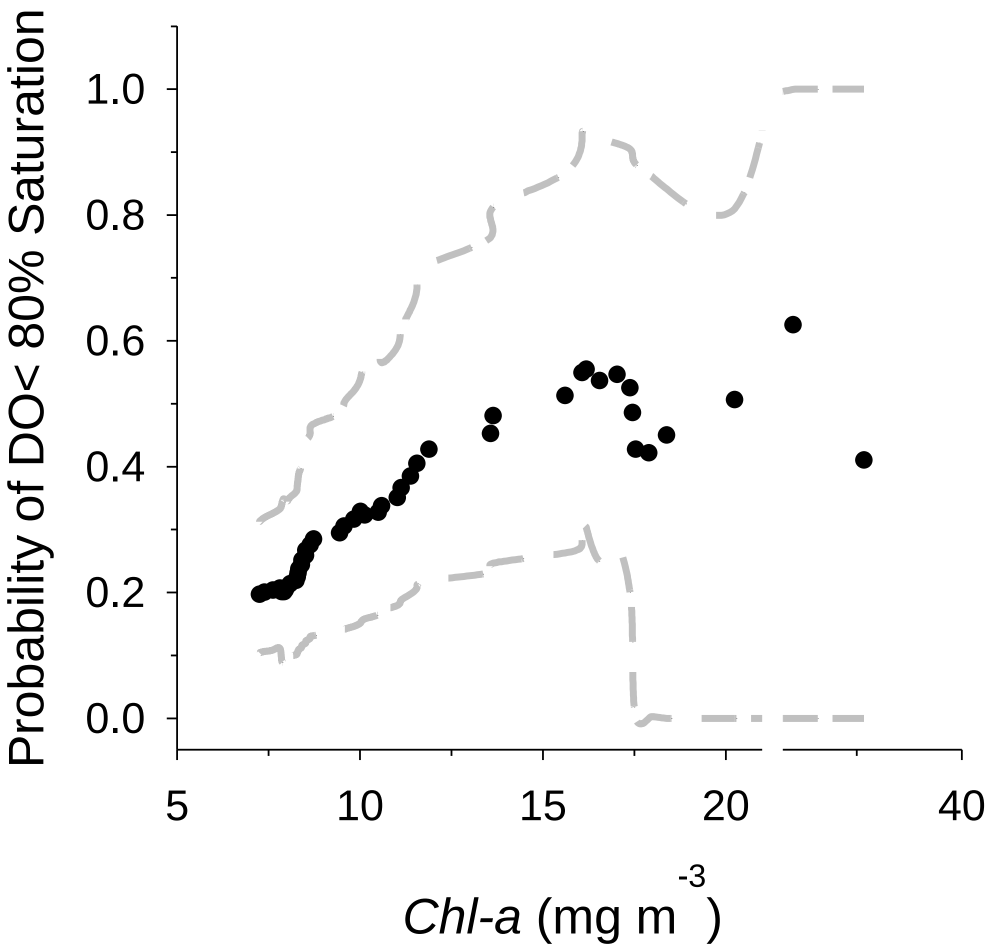

Probability of DO< 80% Saturation
1.0
0.8
0.6
0.4
0.2
0.0
5
10
15
20
40
Chl-a (mg m-3)

Fig 9

Summary of Preliminary Analysis of Historical Stratification in South San Francisco Bay
Mark Stacey, UC-Berkeley
December 2014

The hydrodynamics of South San Francisco Bay are known to force the system into and out of a stratified state, with the belief that the annual cycle of freshwater flows and the spring-neap cycle of the tides dominate the variability. The strength and duration of a stratification event is an important driver of ecosystem variability, due to the associated reduction in vertical mixing. Reduced vertical exchanges lead to retention of phytoplankton in the upper water column, providing improved light conditions and separation from benthic grazers; they also reduce the vertical fluxes of dissolved oxygen to the lower water column, increasing the risk of hypoxic conditions.

In spite of its perceived importance to the South Bay ecosystem, stratification in South Bay has received very little focused study. In the analysis described here, we perform a preliminary analysis of stratification in South Bay using historical data to establish the frequency, magnitude and duration of stratification events. Specifically, we will develop metrics to describe the persistence of stratification events and then use a simple analytical scaling to evaluate the likelihood of longer stratification events under future conditions.

Data Overview

The data used in this study were collected by the U.S. Geological Survey at their San Mateo Bridge location (Figure 1). This mooring includes top (13.4 meters above the bottom) and bottom (3 meters above the bottom) conductivity-temperature-depth sensors, and has been active from 1990 through the present, with the exception of approximately a 2 year gap in 1999-2001. The data streams from the USGS were of high quality, and required no additional quality control, although there is one period in 2001 when the calibration of the two sensors seem somewhat inconsistent (showing inverted stratification in late 2001), but that period does not affect the analysis presented here.

The top-bottom salinity difference was calculated by directly differencing the two timeseries; the resulting record of stratification is shown in Figure 2. Strong annual and inter-annual variability is clearly evident, with strong stratification events typified by vertical salinity differences greater than 5 psu occurring during the wet season, but not in all years. Of particular note are the series of strong, persistent events in the period 1995-1998 and the lack of events in the period 2007-2010.

To examine the timing and duration of these stratification events, we separate the data record by calendar year in Figure 3. In this figure, it is clear that major stratification events occur only in the period from January to early May, and individual events can be of duration of a week or longer. The color-coding in Figure 3 is by calendar year and, although it is not possible in this figure to determine which year is which, it is clear that some years are characterized by regular strong events while others have limited or no events of note.

Stratification Statistics

Our goal is to understand the frequency and duration of stratification events in South Bay, which requires the specification of a threshold for the water column to be considered “stratified”. The impact of this threshold is illustrated in Figure 4, which shows the stratification timeseries from March and April of 1998. The green and red bars show the duration of a stratification event with thresholds of 2 psu and 0.25 psu respectively, and illustrate how the assumption of a stratification threshold can alter the statistics of the frequency and duration of the events. In this particular example, a 2 psu threshold results in the March 6 – March 26 period being divided into 2 separate stratification events, each of approximately 8 days duration. With a threshold of 0.25 psu, on the other hand, the period has a single stratification event with duration of 20 days. The choice of a threshold requires a subjective evaluation of the results: a lower threshold results in fewer, but longer, stratification events while a higher threshold results in more, but shorter, events.

To specify the stratification threshold, we wanted to ensure that the threshold is both above the detection limit and dynamically significant. Examining the two salinity timeseries, and comparing with available Polaris CTD profiles, we concluded that 0.5 psu was a minimum threshold to meet these criteria. To be slightly more conservative, we chose a threshold of 0.75 psu. The results for the March-April 1998 period are illustrated in Figure 5, which shows the period being divided into 4 significant stratification events (as well as many more very short ones that are not highlighted in the figure). These four events ranged from 3.5 to 9 days, and all four had stratification that greatly exceeded the 0.75 psu threshold.

Extending this threshold analysis to the entire data record allows us to count the number of events of particular duration. This frequency analysis would allow us to define the “return period” of particular stratification events in the same way as is done for flood forecasts, although the 20 year record here is not long enough to establish converged statistics. Nonetheless, the frequency distribution of stratification events is shown in Figure 6, which simply presents the number of events (height of bars) of a given duration (horizontal axis, in hours). The two left-most bars represent stratification events of less than 12 and 24 hours, which represent tidally-periodic stratification events that are not relevant to the analysis here. In this 20 year record, only 1 event exceeded 240 hours, but 6 events were in the range between 168 and 240 hours. It is therefore not abnormal to have a significant stratification event of duration 7 days or longer, and the “20-year event” is approximately 12 days.

Temperature Stratification

In particular HAB events, it has been noted that temperature stratification was associated with the HAB. In order to consider the role that temperature stratification plays in persistent stratification events, we repeated the analysis described in the previous section, but considering temperature instead of salinity. In Figure 7, the annual variation of temperature stratification is shown, again color-coded by year (as in Figure 3). Temperature stratification events are associated with the spring months, when air temperatures begin to warm, but salinity stratification is still present. Once salinity stratification is reduced in the summer (see Figure 3), temperature stratification events are also reduced. Our interpretation is that the temperature stratification is a *response* to the combined effects of salinity stratification (which reduces vertical mixing) and atmospheric warming.

To confirm that temperature is not an important driver of persistent stratification events, Figure 8 presents the USGS San Mateo Bridge data in T-S space, with temperature difference on the vertical axis and salinity difference on the horizontal. There is a general upwards trend, with increasing temperature stratification associated with increasing salinity stratification (with slope that varies between events, probably due to differences in air temperature). More importantly, the green line shows the level of temperature stratification that would be required for the temperature to effect density stratification at a level comparable to salinity. In all of the persistent events identified in either the salinity or temperature record, the data falls well below this line, which means salinity dominates the density dynamics.

We conclude that temperature is not an important driver of persistent water column stratification, although it should be noted that temperature stratification that is above the upper USGS sensor would not be detected, and may be playing an additional role in shaping the South Bay ecosystem.

Drivers

With the stratification events identified using the 0.75 psu threshold, we performed preliminary analysis of what the key external drivers were in creating persistent stratification events in South Bay. The data were aggregated by Water Year (October 1 through September 30), which meant that a variety of metrics were possible to describe the stratification including maximum event duration, the total time spent stratified, number of events greater than a particular duration and others. In Figure 9 (panels c and d), we present two of these metrics as a function of water year: maximum event duration in a water year and the number of events longer than 24 hours. Regardless of which stratification metric we used, we found that local precipitation, as measured in San Francisco, was the best predictor. In the top panels of Figure 9, this precipitation data is aggregated across the entire water year (panel a) and for the period October-January (panel b). It is clear in this comparison of timeseries that persistent stratification events are associated with increased local precipitation; they are not as strongly correlated with the major freshwater flows into the Bay, which are dominated by Sierra snowmelt.

As a preliminary evaluation of the relevant drivers for persistent South Bay stratification events, we present a direct comparison of the two precipitation metrics with the stratification metrics in Figure 10. Here, the stratification response metrics (maximum event duration within a water year in the upper panels; total time in events longer than 24 hours within a water year in the lower panels) are directly compared to the precipitation data (total water year precipitation in the left panels; early water year precipitation in the right panels). Although not quantified, there is a better positive correlation between South Bay stratification events and early season precipitation than with total water year precipitation.

Our interpretation of this result is that it is local freshwater flows into South Bay, which are strongly forced by local precipitation, that drive persistent, strong, stratification events in South Bay. These flows can have the largest effect on stratification early in the season, when the South Bay is still relatively saline. In the late spring and summer, large flows entering the Bay through the North Bay and Delta freshen the entire Bay to some extent, including South Bay, so that late season precipitation events have a weaker effect on the local stratification. A more complete evaluation of this dynamical description would require additional analysis, including idealized and realistic modeling and would benefit from a longer data record to more completely evaluate a range of conditions and forcing.

Future Conditions

Finally, we wish to explore the prospects for a significant change in the frequency or duration of stratification events under the influence of climate change. The balance between stratifying and destratifying forces is captured by the Simpson number:

$Si = BH/u^{*3}$

In this expression, B represents the stratifying influence of freshwater flows and the associated density gradients; u^{*3}/H is the destratifying effects of tidal mixing. In the coming century, both B and u* may be modified, either through changes in precipitation or in tidal forcing (due to the combined effects of sea level rise and new inundation).

To examine how much adjustment from current conditions would be required to create significant changes in the stratification regime, we present in Figure 11 the tidal velocity cubed from observations at the San Mateo Bridge location (the data is from September, but tidal forcing is similar in March), including both the instantaneous (blue) and tidally-averaged (red) currents. The top panel shows current conditions, with the 20-year event (12 day duration) illustrated with the green bars. The idea is that as tidal mixing decreases into the neap tides, it drops below some threshold and the water column stratifies (the start of the green bars); the stratification then persists until the tidal mixing increases to the point that the water column is mixed (the end of the green bars). The upper green bar illustrates this dynamic based on the instantaneous currents; the lower green bar is based on the tidally-averaged currents.

In the lower two panels, we present schematically where this threshold would be if there is a 5% (panel b) and 10% (panel c) adjustment in the relative strength of tidal mixing as compared to buoyancy. That is, panel b represents the case where tidal currents decrease by 5% or buoyancy forcing increases by 16% (because the Simpson number depends on the velocity cubed but is linear with buoyancy). Panel c represents the case where tidal currents decrease by 10% or buoyancy forcing increases by 37%. In each case, it is clear that the 20-year stratification event would increase in duration significantly (consider extending the green bars left and right until they intersect with the velocity data), and in the case of a 10% reduction in the tidal currents, the stratification may persist across the spring tides as well as the neaps, leading to stratification that will vary with freshwater flows, rather than the spring-neap cycle.

Summary

In summary, we found that strong, persistent stratification is common in South Bay, with regular events of magnitude greater than 5 psu that extend for 7 or more days. At the same time, the historical record does not include events that last longer than 12 days, although this may not be the case under future conditions. The key drivers of stratification appear to be early-season local precipitation, although this conclusion would require further analysis to establish it firmly.

It is important to note that this analysis is limited to the central portion of South Bay. The paucity of data in the Lower South Bay (south of the Dumbarton Narrows) made it impossible to evaluate the variation and dynamics of stratification there, and it is quite possible that longer-duration events are more typical in that embayment.

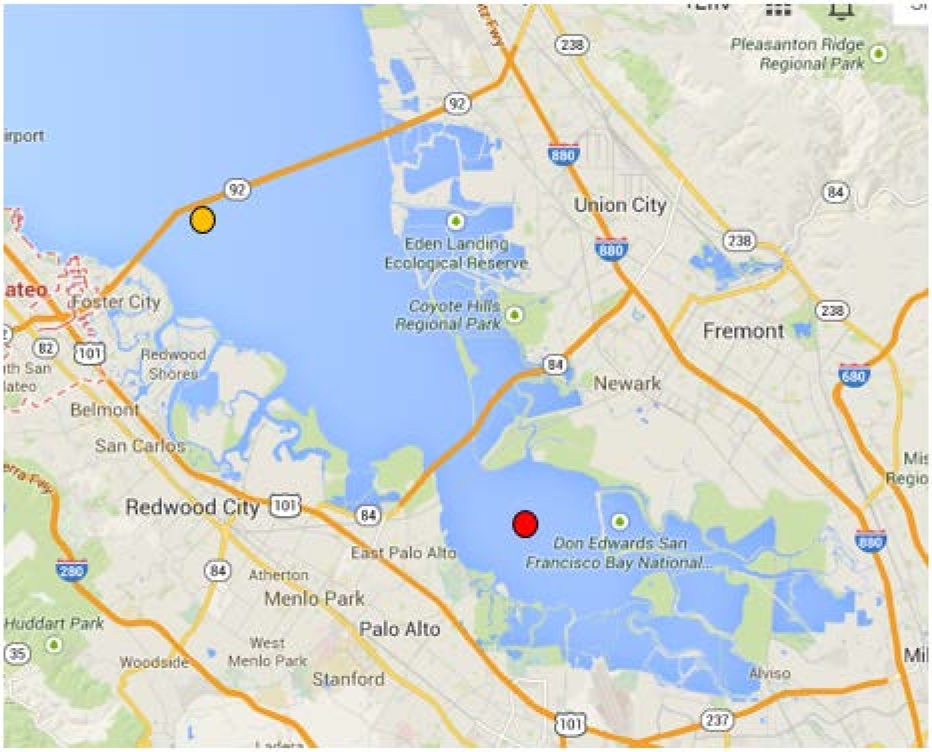


Figure 1: Locations of data sets used in analysis. USGS long-term (1990-present) San Mateo Bridge mooring location marked by yellow marker. Short-term record from lower south bay (red) was found to be of insufficient length or quality for detailed stratification analysis.

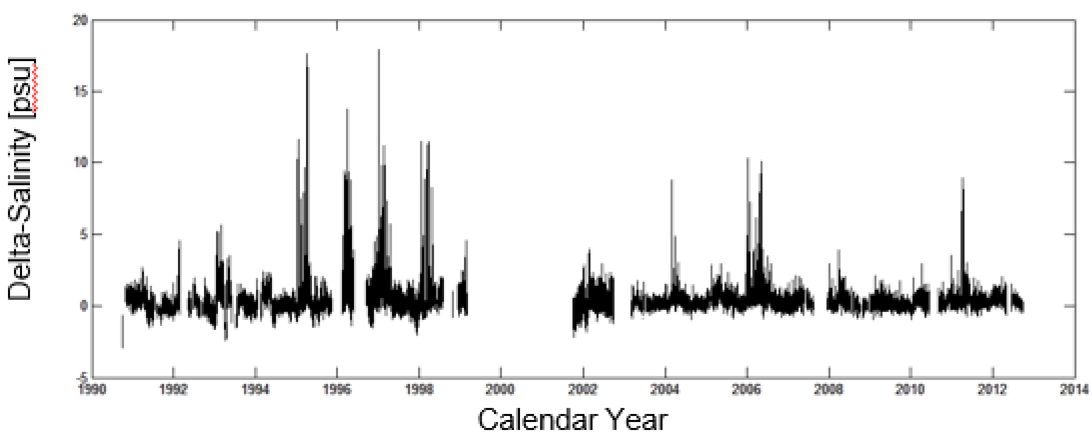


Figure 2: Overview of stratification record at USGS San Mateo Bridge mooring; calculated as bottom salinity minus top salinity

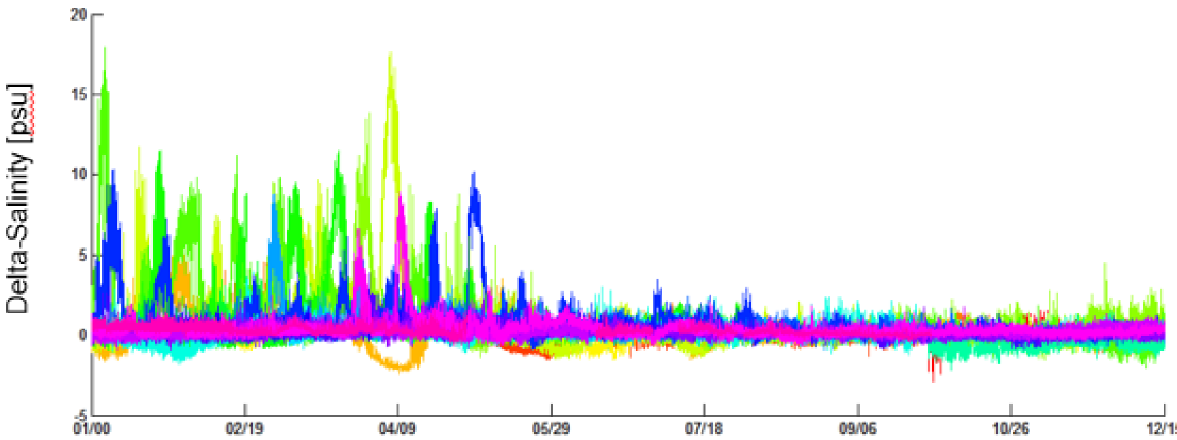


Figure 3: Overview of San Mateo Bridge stratification timeseries as a function of the day of the year, color coded by calendar year.

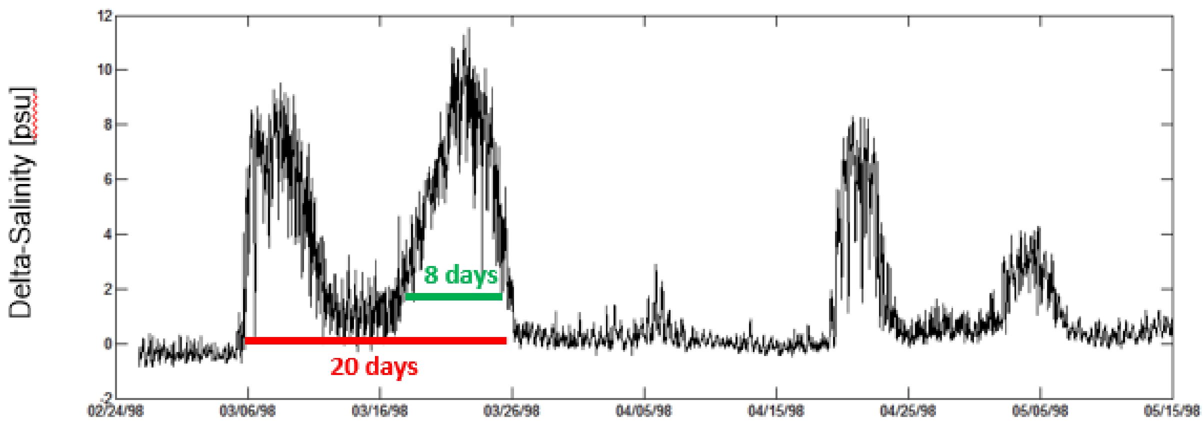


Figure 4: Example of stratification event from March-April 1998. Green and red bars indicate lengths of 8 and 20 days, respectively.

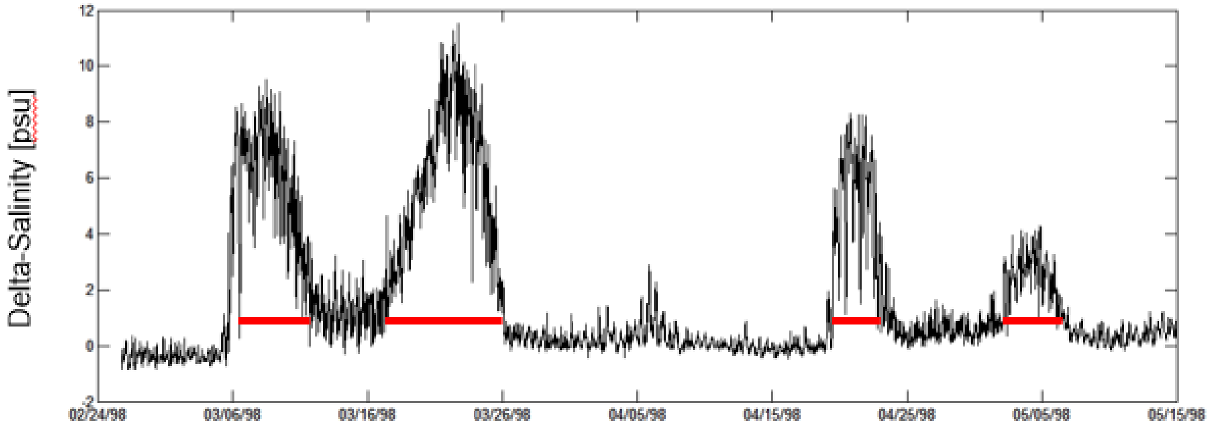


Figure 5: March-April 1998 stratification illustrating 0.75 psu threshold

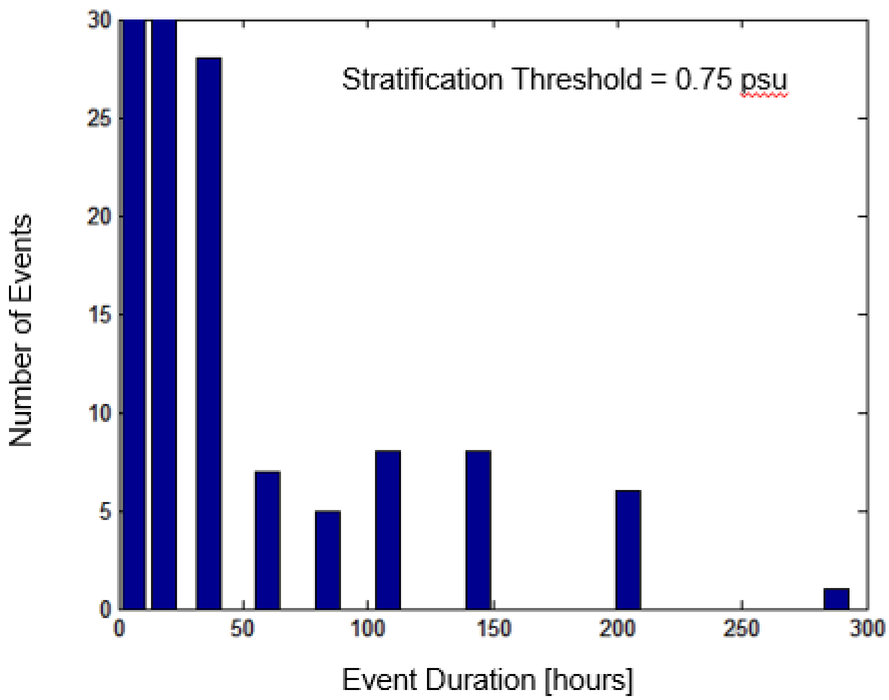


Figure 6: Number of events in stratification record (20 years of data) of a given duration.

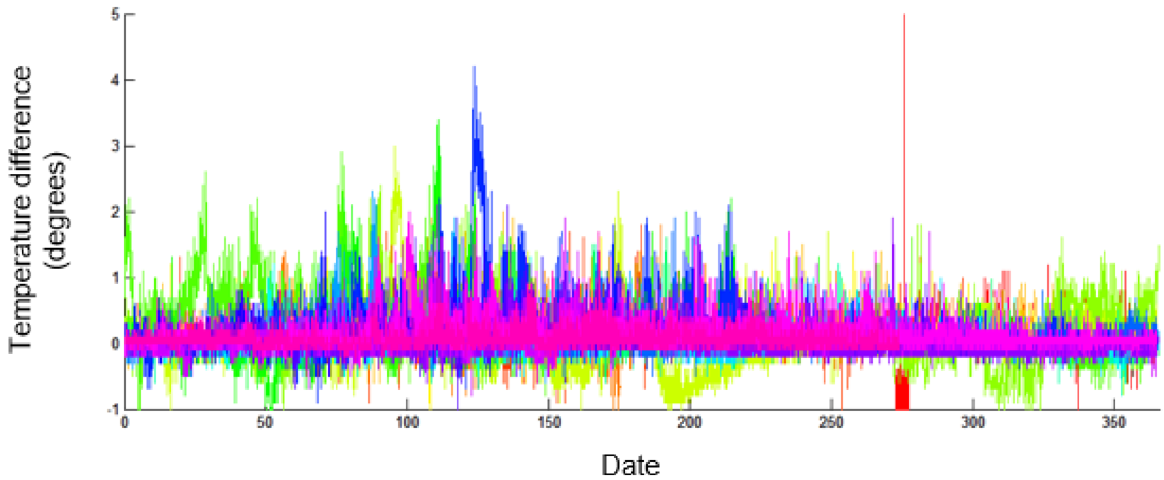


Figure 7: Annual variability of top-bottom temperature difference, color-coded by year (as in Figure 3)

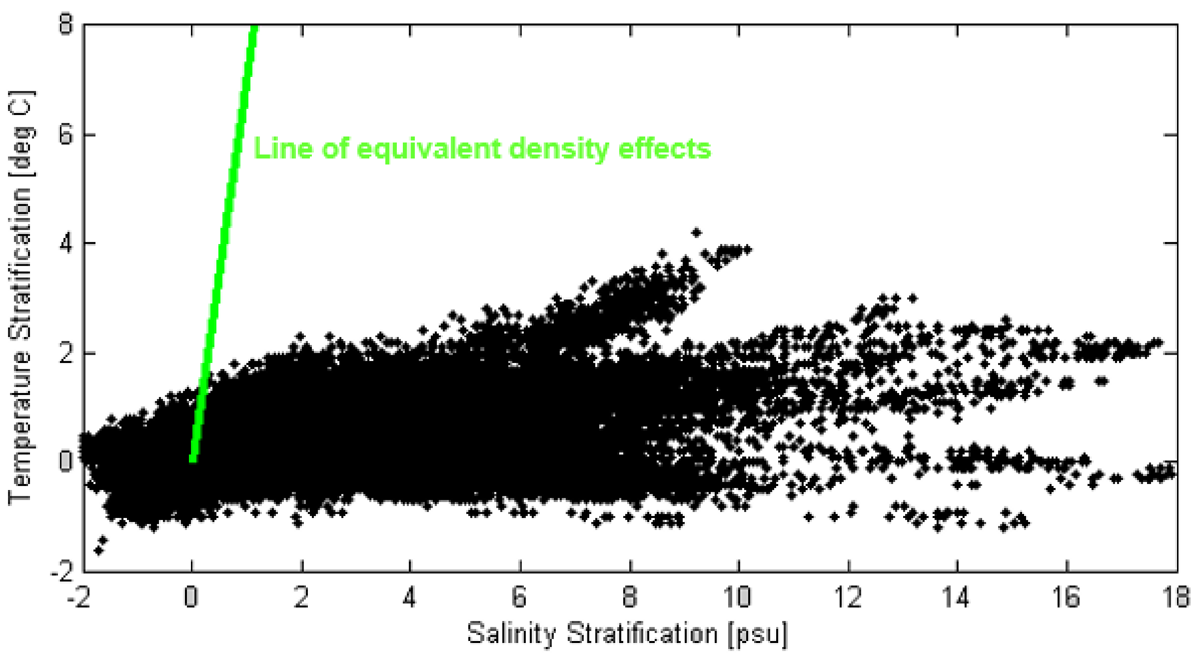


Figure 8: Comparison of salinity and temperature effects on stratification. Data is shown as black dots in T-S space; Green line shows equivalent density effects

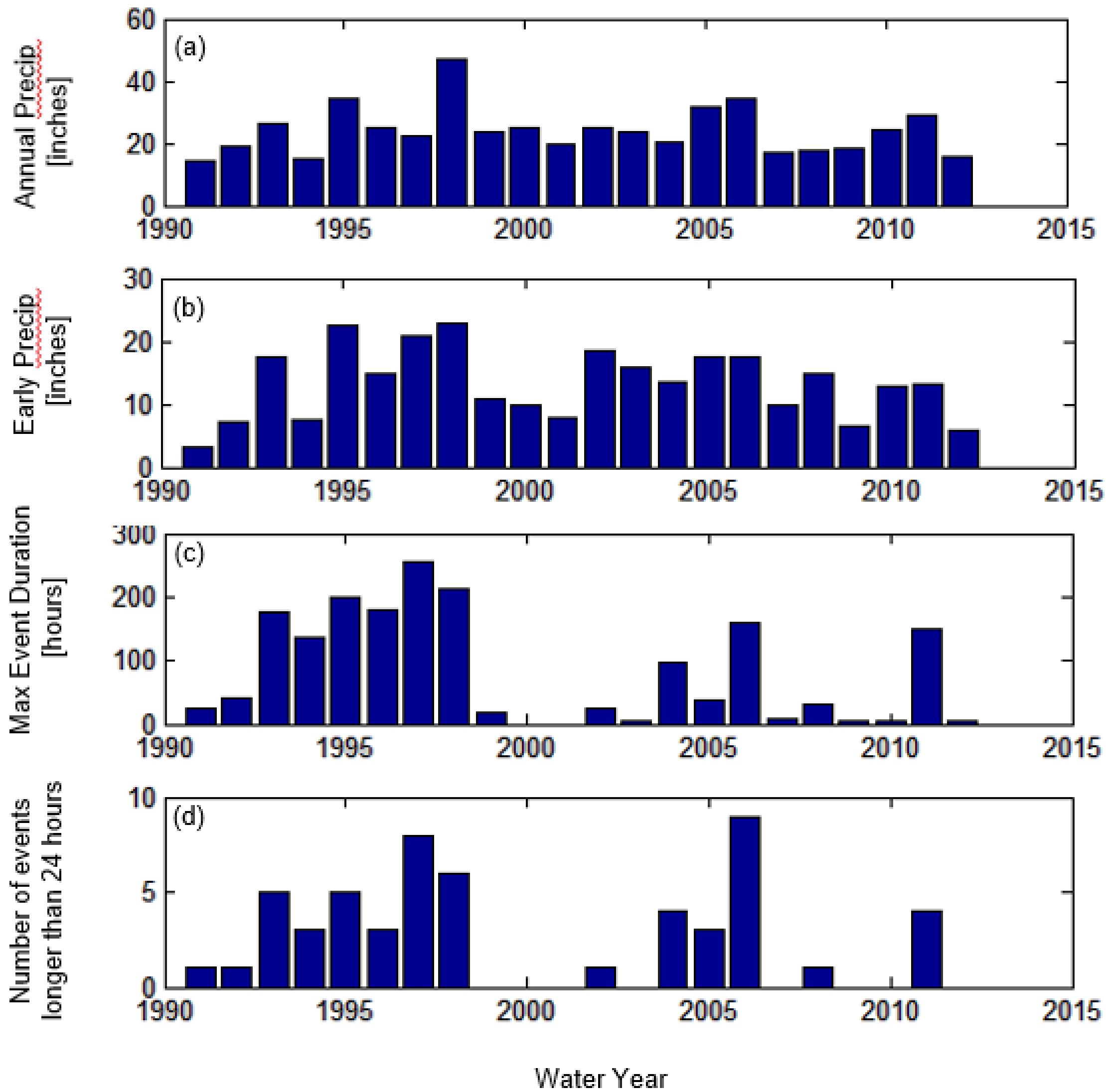


Figure 9: Variation by water year (October 1 – September 30). (a) Total annual precipitation; (b) Total precipitation in October through January; (c) Maximum duration of stratification event (threshold of 0.75 psu) for the water year; (d) Number of stratification events longer than 24 hours in the water year.

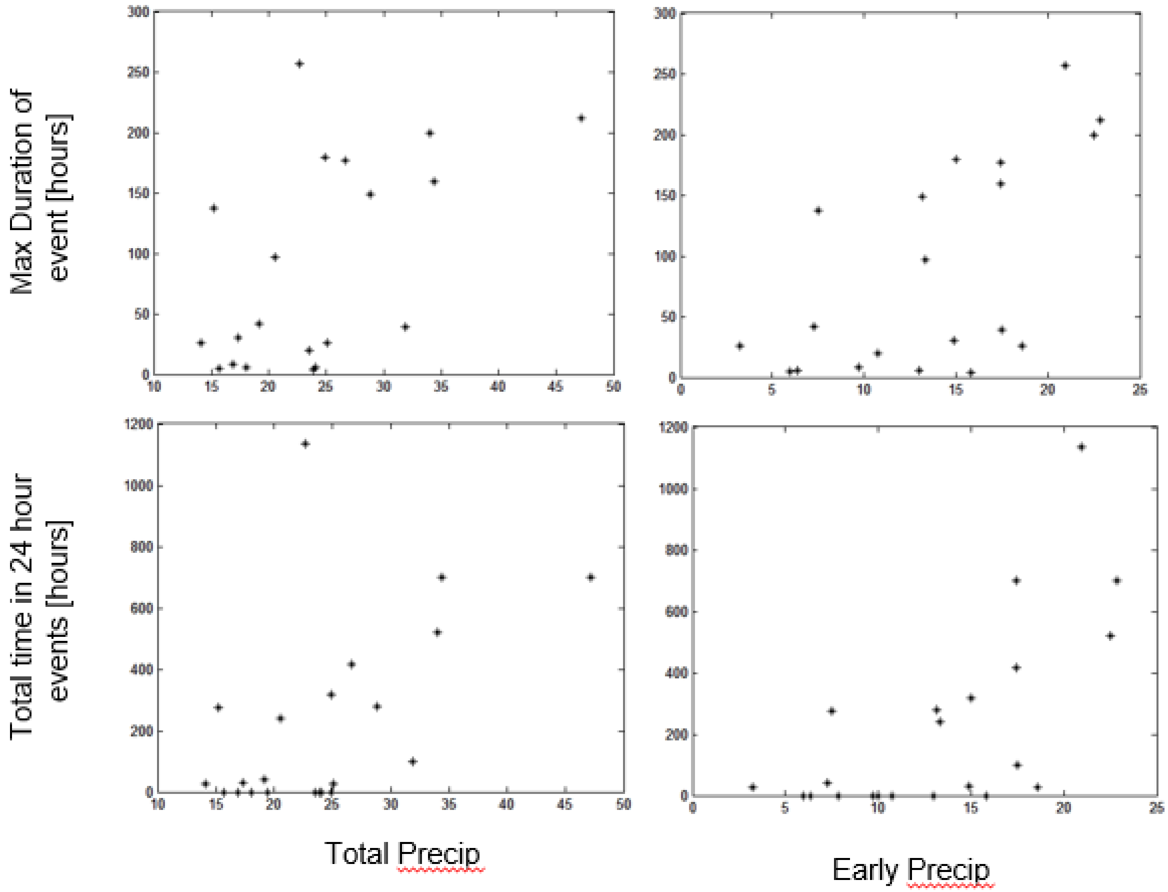


Figure 10: Preliminary evaluation of drivers for South Bay stratification. Candidate drivers (Total Water Year Precipitation and Early Season Precipitation) are on the horizontal axes; Response metrics (Maximum event duration and total time in stratification events) are on the vertical axes.

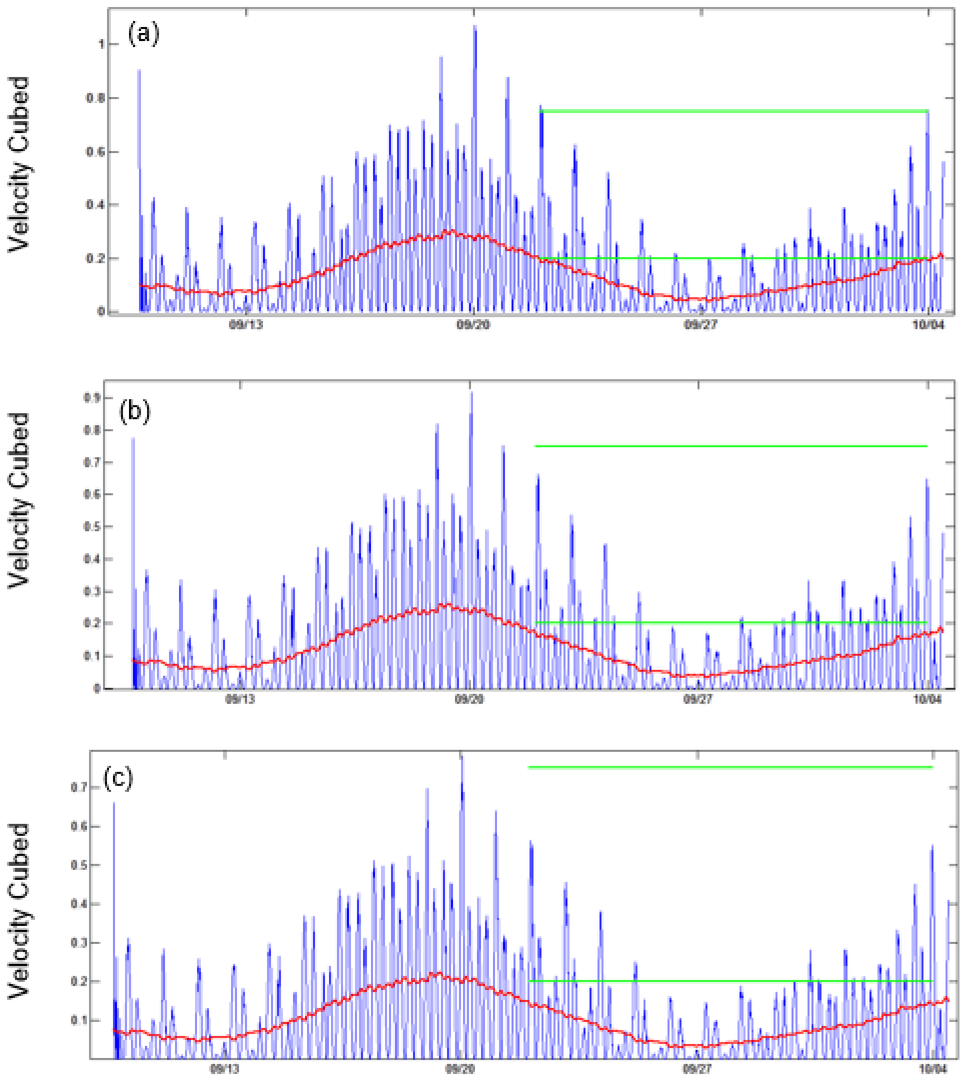


Figure 11: Schematic of effects of reduction of tidal velocities. Blue line is (absolute value of) velocity cubed at San Mateo site; Red line is tidally-averaged velocity cubed. Green bars show 12 day event (approximately the 10 year event under current conditions). In panel (a), current conditions are shown. In panels (b) and (c), the vertical position of the green bars represent the relative level of buoyancy forcing under a 5% and 10% reduction in tidal energy, respectively.