SAN FRANCISCO BAY
NUTRIENT MANAGEMENT STRATEGY

San Francisco Bay Numerical Modeling: 2021-2023 Update

Project Team:

Allie King[1], Farid Karimpour[1],
Pradeep Mugunthan[1,2], David Senn[1]

[1] San Francisco Estuary Institute

[2] Four Peaks Environmental Science & Data Solutions

January 2023

SFEI Contribution #1113

Suggested Citation:

SFEI (San Francisco Estuary Institute). 2023. *San Francisco Bay Numerical Modeling: FY2021-FY2023 Update*. SFEI Contribution #1113. San Francisco Estuary Institute, Richmond, CA.

TABLE OF CONTENTS

APPENDICES

LIST OF FIGURES

LIST OF TABLES

ACRONYMS AND ABBREVIATIONS

Acronym	Definition
BAHM	Bay Area Hydrologic Model
BGC	biogeochemical
C	carbon
chl-a	chlorophyll-a
CIMIS	California Irrigation Management Information System
DEB	dynamic energy budget
DEM	digital elevation model
DFM	D-Flow Flexible Mesh
DIN	dissolved inorganic nitrogen
DO	dissolved oxygen
DWAQ	D-Water Quality
DWR	California Department of Water Resources
GAM	generalized additive model
GPP	gross primary production
HAB	harmful algae bloom
LSB	Lower South Bay
N	nitrogen
NH4	ammonium
NMS	Nutrient Management Strategy
NO3	nitrate
P	phosphorus
PO4	phosphate
POTW	publicly owned treatment works
RMP	Regional Monitoring Program
RMSE	root mean squared error
SFB	San Francisco Bay
SFB-BGCM	San Francisco Bay Biogeochemical Model
SFEI	San Francisco Estuary Institute

Acronym	Definition
Si	silicon
SOD	sediment oxygen demand
SSC	suspended sediment concentration
USGS	U.S. Geological Survey
WWTP	wastewater treatment plant
WY	water year

ACKNOWLEDGEMENTS

The work presented in this report was supported primarily with funding from the San Francisco Bay Nutrient Management Strategy (NMS). Numerous SFEI staff, external collaborators, and NMS stakeholders, have contributed valuable guidance and input during the reporting period covered in this report and in the prior years leading up to it.

1 Introduction: Nutrient Management Strategy Modeling Work, FY 2021-2023

San Francisco Bay (SFB) receives elevated inputs of the nutrients nitrogen (N) and phosphorus (P) due to human activities and ranks among the most nutrient-enriched estuaries worldwide (SFEI 2014a; SFEI 2014b; Cloern et al. 2020). Effluent from wastewater treatment plants (WWTPs)—which treat wastewater from the SFB area's 7.5 million people—account for the vast majority of N and P sources to SFB (SFEI 2014b). Despite its high nutrient inputs, SFB has not historically experienced the classic eutrophication symptoms of many other nutrient-enriched estuaries. Nonetheless, SFB's nutrient-enriched status makes severe impacts plausible, including: large phytoplankton blooms and low dissolved oxygen (DO) in open-bay and slough/creek habitats; harmful algae blooms (HABs) in open-bay and slough/creek habitats; and impacts within coastal habitats related to nutrients exiting SFB via the Golden Gate (SFEI 2014b).

The SFB Nutrient Management Strategy (NMS) Science Program is charged with developing the scientific foundation to inform decisions related to managing nutrient loads to SFB. The management questions that are being evaluated in the NMS are summarized in Table 1.1. One major focus of the NMS involves developing numerical models to simulate hydrodynamics and biogeochemical processes in SFB and undertake studies using those models to answer applied science questions related to nutrient transport and cycling and the effects of nutrients on water quality (Table 1.2). NMS modeling work is guided by a 5-year modeling work plan (2019-2024; SFEI 2019). Figure 1.1 summarizes previous and current model development activities that have been completed to date.

This report serves as a progress and status update on NMS numerical modeling work, focusing on work carried out between July 2020 and December 2022.

The modeling activities documented in this report include the following:

- *Section 2 Hydrodynamics*: Simulations of SFB for water years (WYs) 2014, 2015, 2016 and 2018; development of a combined set of tidally averaged hydrodynamic input for the aggregated grid San Francisco Bay Biogeochemical Model (SFB-BGCM) to enable a continuous simulation of biogeochemicals (BGCs) from WY2013 – WY2018 (Section 2 and Appendix A); and detailed validation of stratification and vertical mixing of the hydrodynamic model (Appendix B).
- *Section 3 Biogeochemistry:* Extension of SFB-BGCM aggregated grid model to additional WYs (2013-2018) and calibration of grazing and sediment flux parameters of the model to observed water column and benthic data; running full resolution SFB-BGCM simulations for WY2013, WY2017, and WY2018 using revised model setup and updated calibration based on WY2013-WY2018 aggregated simulation work; preliminary evaluation of model performance (6-year aggregated simulation; 3 years of full-resolution runs); and sensitivity analyses of key biogeochemical model inputs through NMS science and management-relevant performance metrics.

Table 1.1. San Francisco Bay Nutrient Management Strategy overarching management questions

1. What conditions would be considered adverse impacts or impairments that require management or mitigation?	
	1.1 Dissolved oxygen / chl in deep subtidals
	1.2 Dissolved oxygen in shallow margin habitats
	1.3 Harmful algae bloom abundance, toxin abundance, phytoplankton assemblage
	1.4 Coastal ocean
2. Monitoring and condition assessment: Are adverse impacts or impairment currently occurring?	
	2.1 Dissolved oxygen / chl in deep subtidals
	2.2 Dissolved oxygen in shallow margin habitats
	2.3 Harmful algae bloom abundance, toxin abundance, phytoplankton assemblage
	2.4 Coastal ocean
3. How do San Francisco Bay habitats respond to nutrient inputs—dose: response? Are nutrients causing or contributing to current impacts or impairment?	
	3.1 Dissolved oxygen / chl in deep subtidals
	3.2 Dissolved oxygen in shallow margin habitats
	3.3 Harmful algae bloom abundance, toxin abundance, phytoplankton assemblage
	3.4 Coastal ocean
4. What potential future impacts or impairments warrant pre-emptive management actions?	
5. What are the contributions of individual nutrient sources to nutrient levels throughout San Francisco Bay? How do these contributions vary in space and time (seasonal, under different physical forcings)?	
	5.1 Current magnitudes (± variability) of individual nutrient loads at their point of entry to San Francisco Bay?
	5.2 Anticipated load changes: environmental change, flow diversion, population, land-use, management
	5.3 Magnitudes of nutrient transformations and losses within San Francisco Bay, space/time variability?
	5.4 Contributions of individual nutrient sources to loads/concentrations in “subregions”?

6. What management actions or load reductions are needed to prevent or mitigate current future impairment?	
	6.1 What "local" reductions/changes are needed within subregions to mitigate impairments?
	6.2 What load reductions or other management actions can achieve the "local" effect(s)?
	6.3 What are plausible options for achieving the local effects?

Note:
chl: chlorophyll-a, a measure of phytoplankton biomass

Table 1.2. Modeling-related science questions that guide model development and the design of projects that apply the models

	Modeling Science Questions	Management Questions																			
		1				2				3				4	5				6		
		1.1	1.2	1.3	1.4	2.1	2.2	2.3	2.4	3.1	3.2	3.3	3.4		5.1	5.2	5.3	5.4	6.1	6.2	6.3
A	How do individual nutrient sources contribute to specific habitats, regions, and subembayments, and how do relative contributions vary as a function of physical forcings?									■	■	■					■	■		■	■
B	What are reasonable regional or subembayment boundaries for nutrient management considerations (contiguous areas having similar relative nutrient sources)?									■	■	■					■	■		■	■
C	Subembayment-level nutrient budgets: What are the magnitudes of loads, losses, transformations, and transport? What are the assimilative capacities for N and P?									■	■	■					■	■	■	■	
D	How do fluxes of N and P to the coast vary seasonally/interannually, and what are the primary controlling factors?												■	■					■		
E	What factors regulate phytoplankton blooms/productivity? Under what conditions do large blooms occur?					■	■	■		■	■	■		■					■	■	
F	Under what conditions could low DO develop in deep subtidal habitats?	■				■				■				■					■	■	
G	How do salt pond restoration and management influence nutrients, chl, DO in LSB? (current and future openings)						■			■	■			■			■	■	■	■	
H	To what extent do exchange with salt ponds, combined with high nutrient loads, contribute to DO deficits in LSB sloughs in creeks?						■			■	■			■			■		■	■	
I	What factors contribute most strongly influence, or best explain, long-term trends in chl-a over time?									■		■	■	■					■		
J	How will the Bay respond to changing N inputs? Increases due to population growth; Decreases resulting from load reduction scenarios									■	■	■	■	■			■	■	■		
K	What confidence intervals or levels of certainty are associated with management-relevant model interpretations?									■	■	■	■	■			■	■	■	■	■

Notes:
Shaded boxes indicate NMS Management Questions (Table 1.1) addressed by each Science Question.
LSB: Lower South Bay; DO: Dissolved Oxygen; N: Nitrogen; P: Phosphorus

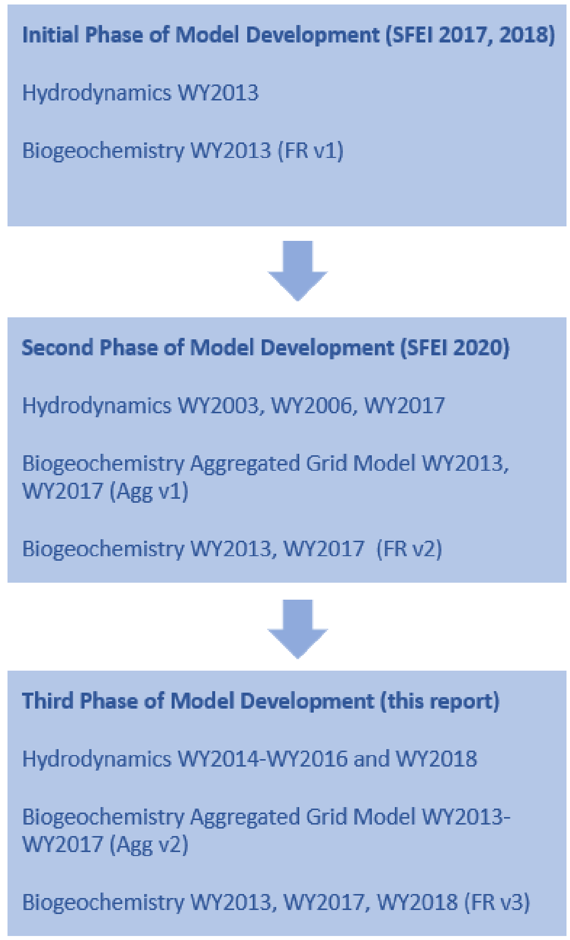


Figure 1.1. San Francisco Bay biogeochemical model development phases

2 Hydrodynamic Modeling Update

2.1 Summary

A coupled hydrodynamic-biogeochemical model of SFB has been developed as part of the NMS. This process-based, numerical model will be used to inform nutrient management decisions by:

- Improving quantitative understanding of processes that shape current conditions,
- Forecasting ecosystem response under future scenarios, and
- Evaluating potential effectiveness of management actions.

This report is an update to previous hydrodynamic modeling efforts (Nuss et al. 2018; SFEI 2020) and provides a description of the model setup and validation for additional water years (2014, 2015, 2016, and 2018).

The model inputs include tides, direct precipitation, evaporation, solar radiation, stormwater runoff, wastewater discharges, Delta outflow, and wind. From these inputs, the model calculates water levels, salinity, temperature, currents, and the force of the currents on the bed (shear stress) throughout SFB.

A wide range of observations collected throughout SFB are used to assess the model's predictive skill. This observational dataset includes tidal water levels alongside salinity readings from shipboard and moored sensors. WY2013 validation also included comparing the model's predicted velocities to measurements from a series of acoustic Doppler current profilers across the bay. A spatially variable drag coefficient was calibrated using that velocity dataset, and Nuss et al. (2018) validated modeled velocities in WY2013 using that dataset.

A complete set of validation figures for WY2013 – WY2018 are provided in Appendix A and are organized as follows:

- Appendix A1 shows the high frequency water level validation plots at the NOAA stations
- Appendix A2 shows the salinity validation plots at USGS Peterson Cruise channel locations.
- Appendix A3 shows the high frequency salinity validation plots at the USGS mooring stations.

In addition to the water level and salinity validations, a detailed evaluation of the model's ability to predict density and stratification was completed. The approach and the results of this evaluation are provided in Appendix B.

The comparisons show that there is good agreement between the model results and observations. The density verification indicated that the model can successfully simulate longitudinal and seasonal changes in density across all the water years evaluated. The strong validation performance provides confidence in the hydrodynamic model's ability to simulate the fundamental physical processes in SFB.

The remainder of this section provides details on model configuration, skill analysis, and includes discussion of potential causes for model-data disagreements.

2.2 Model Setup

The SFB hydrodynamic model is built on D-Flow Flexible Mesh (DFM). DFM is part of the Deltares suite of models, which also includes D-Water Quality (DWAQ), the platform chosen for the biogeochemical phase of the NMS modeling project (Section 3). DFM is a finite-volume, three-dimensional, unstructured hydrodynamic model (Martyr-Koller et al. 2017). The unstructured nature of the grid allows for efficient and flexible resolution of flow features ranging from small perimeter sloughs and ponds up to a regional representation of the coastal ocean. This range of features is resolved without explicit seams or nesting boundaries, which would be required should a structured grid model be applied to the same area.

The original model setup was developed by Silvia Pubben and Mick van der Wegen (Pubben 2017) as a continuation of prior modeling efforts stemming from the U.S. Geological Survey (USGS) Computational Assessments of Scenarios of Change for the Delta Ecosystem (CASCaDE)[1] and San Francisco Bay-Delta Community Model projects. We set up the model to simulate WY2013 and presented the improvements, validation, and discussion in the Nuss et al. *(*2018). We continued model development and validation by running water years 2003, 2006, and 2017, presented in SFEI (2020). In the current report, we present the model development and validation of additional water years (WY2018, and a combined model run spanning WY2014, WY2015, and WY2016).

Each water year spans October 1 through September 30, with a spin-up time of 2 months. The validation results from the years presented in the current report, in conjunction with previously validated water years, confirms that our hydrodynamic model captures a wide range of physical conditions.

The hydrodynamic model runs were performed on a Linux workstation using 16 Intel Xeon E5-2680 2.40 GHz cores, communicating over MPI. The full 426-day run takes 7.0 days of wall-clock time, for a simulation speed 61 times faster than real time. DFM was compiled from SVN revision 52184 of the source code and GCC 5.4.0.

2.3 Model Domain and Grid

The model domain covers SFB, including portions of Coyote Creek and Guadalupe River at the SFB's southernmost extent, and extending north to the Sacramento and San Joaquin rivers at Rio Vista and Jersey Point, respectively (Figure 2.1). Separate but related models include a complete Delta and detailed Lower South Bay (LSB) bathymetry, respectively; due to computational expense it is not feasible to merge these three models at this time, and the present model is optimized for the OpenBay applications excluding the Delta and much of the LSB domain. The domain extends into the Pacific Ocean, about 20 km west of Point Reyes in the north and 40 km west of Half Moon Bay in the south, roughly encompassing the San Francisco Bight. The horizontal grid resolution varies from 20 m in select sloughs of LSB, to over 2 km at Point Reyes. Nominal grid resolution in South Bay is 250 m, and 350-500 m in North Bay, for a total of 49,996 cells in the horizontal. The three-dimensional model uses a sigma coordinate in the vertical, such that all areas have 10 layers in the vertical, with uniform layer thicknesses varying in accordance with the overall depth.

Bathymetry for the model has been compiled from a combination of digital elevation models (DEMs):

- 10 m topo-bathymetry[2] from the California Department of Water Resources (Wang and Ateljevich 2012)
- High resolution USGS bathymetry[3] in LSB (Foxgrover et al. 2011)

Bathymetry is prescribed at the nodes of the grid by linear interpolation on the source DEMs, then internally extrapolated to edges and cells. All elevation data are relative to the NAVD88 vertical datum.

[1] https://cascade.wr.usgs.gov/
[2] http://baydeltaoffice.water.ca.gov/modeling/deltamodeling/modelingdata/DEM.cfm
[3] https://pubs.usgs.gov/of/2011/1315/

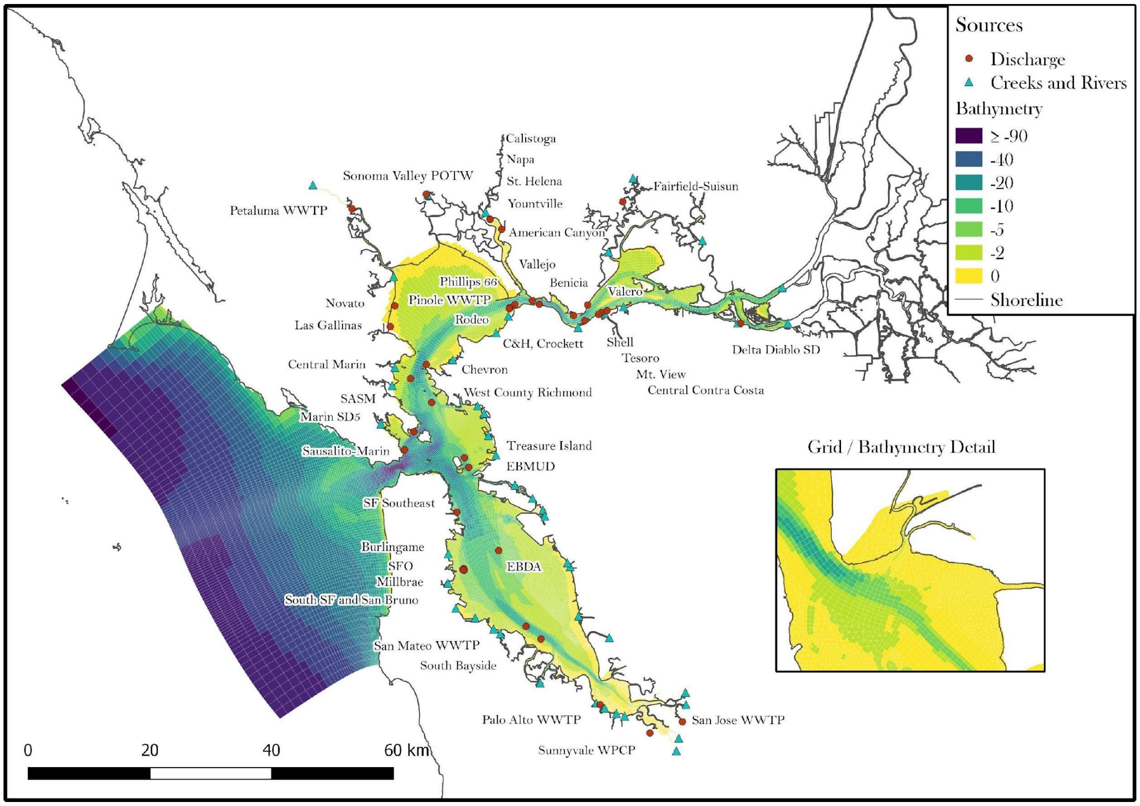


Figure 2.1. San Francisco Bay D-Flow Flexible Mesh grid

2.4 Boundary Conditions and Forcings

2.4.1 Tidal Ocean Boundary

The open ocean boundary is forced along the western edge by observed sea level data from Point Reyes, via NOAA gauge 9415020. Observed 6-minute water levels are low-pass filtered with a fourth order low pass Butterworth filter with a 3-hour cutoff period. Salinity at the ocean boundary is set to a constant 33 ppt. The shorter northern and southern edges of the ocean boundary are closed.

2.4.2 Bay Area Rivers and Stormwater

For WY2013-WY2017, freshwater flows were derived from the Bay Area Hydrologic Model (BAHM), a Hydrological Simulation Program–Fortran (HSPF)-based hydrologic model. This model was calibrated against gauge data over the 2000-2017 period, and includes 44 separate river and stormwater inputs to SFB. All river and stormwater inputs are assumed to enter SFB with negligible salinity and a constant temperature of 20°C. Appendix B in Nuss et al. (2018) describes BAHM, its inputs, and its performance in detail.

For WY2018, the newly developed Watershed Dynamic Model (WDM) was used to provide the freshwater flows from local tributaries. The Watershed Dynamic Model is a Loading Simulation Program in C++ (LSPC) model developed under the Regional Monitoring Program (RMP). The hydrologic simulation has been calibrated and validated against USGS gauge data. The detailed model setup, calibration, and validation can be found in the Phase 1 modeling report (Zi et al. 2021). This model covers all the local watersheds (237 in total) that drain to SFB from 88 pour points. To be consistent with the previous (WY2013-WY2017) model settings, freshwater flows from the 88 pour points were aggregated to 44 inflow time series based on the spatial location of pour points of the BAHM model.

The hydrodynamic model does not extend into the Sacramento-San Joaquin Delta. The edge of the model domain is at Rio Vista on the Sacramento River and Jersey Point on the San Joaquin River, well within the tidal portion of the system. Boundary conditions are taken from USGS streamflow gauges at stations 11455420 (Sacramento River at Rio Vista) and 11337190 (San Joaquin River at Jersey Point) for the Sacramento and San Joaquin flows, respectively. Flows at Rio Vista and Jersey Point are assumed to have negligible salinity. Water temperature is obtained from the same USGS gauging stations and assigned to inflowing water in the model.

2.4.3 Wastewater Treatment Plants and Refineries

Given the goal to support nutrient-related studies in SFB, special attention has been given to WWTP inputs. In parts of SFB, especially South Bay, WWTP inputs are significant freshwater sources, and it is therefore essential that they are accounted for in the hydrodynamics where they can properly influence the density field (as opposed to added only as nutrient mass sources in the water quality model). Flow and load data for 37 WWTPs and 5 refineries have been compiled in a git repository that draws on observed flow rates where data exist. For dates when flow data are unavailable, a flow rate is estimated based on inter-annual trends and a seasonal flow climatology. Each of the 42 inflows have been added to the hydrodynamic model as a freshwater source located at the bed.

2.4.4 Winds

The wind field used in the WY2013 run was developed using the Ludwig wind model (Ludwig et al. 1991; Ludwig and Sinton 2000) to generate 10-m winds. To improve the accuracy of the wind field, we used the San Francisco Estuary Institute (SFEI) Wind package (King 2019) for the WY2017, WY2014-WY2016, WY2003, and WY2006 hydrodynamic model runs. SFEI Wind interpolates winds measured at 52 stations around SFB and the Delta and improves predictions of wind speed by approximately 30% compared to the Ludwig model. Wind stress over the water surface is calculated within DFlow3D-FM from 10-m wind speeds, which we specify hourly on a 1.5 km × 1.5 km grid spanning SFB and the Delta.

2.4.5 Precipitation and Evaporation

In addition to stormwater runoff, which enters the model at prescribed locations along the boundary, we also include direct precipitation and evaporation acting directly on the water surface. The model incorporates measured precipitation and evapotranspiration from the California Irrigation Management Information System (CIMIS) Union City station. Evapotranspiration is scaled to pan evaporation[4] by dividing by a factor between 0.6 and 0.7 depending on temperature. The present model configuration further scales this evaporation down by a factor of 2. Without this factor of 2 downscaling, South Bay waters become too salty in low flow conditions, i.e., summer, which may suggest that freshwater inputs to South Bay are under-represented. South Bay is more sensitive to evaporation than other regions of the bay due to long residence times and minimal local freshwater inflows. Given these conditions, using a South Bay data source across the entire domain has not been problematic.

[4] https://cals.arizona.edu/azmet/et1.htm

2.5 Validation

2.5.1 Definition of Metrics

In addition to graphical figures comparing modeled and observed quantities, we present several numeric measures of predictive skill, defined below. In these definitions a horizontal line ($\underline{x}$) indicates the arithmetic mean, and a subscript x_i indicates individual samples of the respective dataset.

Skill – Model skill is calculated according to the formula proposed by Willmott 1981:

$$SS = 1 - \frac{\Sigma_i (m_i - o_i)^2}{\Sigma_i \left(| \, m_i - \underline{o} \, | + | \, o_i - \underline{o} \, | \right)^2}$$

where o_i denotes samples in the observed data, and m_i denotes samples in the model results. A perfect model achieves a skill score of 1. Unlike the correlation coefficient, the skill score takes into account both the correlation and the relative scales of the modeled and observed data.

Bias – The average error $\underline{m_i - o_i}$.

***r*²** – The squared correlation coefficient is defined as:

$$r^2 = \frac{C_{mo}^2}{C_{mm} C_{oo}}$$

where C_{ij} is the covariance matrix computed between the model and observed data. A value of 1.0 indicates perfect correlation or anti-correlation, and a value of 0.0 indicates no correlation between the model and observed data.

RMSE – Root mean squared error $\left(\underline{(m_i - o_i)^2} \right)^{1/2}$

Lag – The time shift, in minutes, between the model time series and the observed time series. A positive value means that the model lags the observations, and a negative value means that the model leads the observations.

Amplification factor – The ratio of standard deviations between the model and the observations. A factor of 1.0 is perfect, while a factor greater than 1 indicates that the signal is amplified in the model, and a factor less than 1 indicates that the signal is attenuated or under-represented.

2.5.2 Baseline Expectations

Threshold values separating a satisfactory versus an unsatisfactory validation are difficult to generalize for these metrics. The following baseline expectations for a satisfactory validation include the caveat that the details of whether the validation is sufficient are often site- and application-specific. Water level is the most fundamental, and typically the easiest, quantity to validate, and we expect to have skill and correlation coefficients well above 0.9. Time lags should be much shorter than the period of the tides. Skill metrics for salinity at the scale of this model require nuanced interpretation due to the wide range of processes that contribute to salinity distributions. We present skill metrics for the limited number of salinity time series stations available. Additionally, wherever possible, we include descriptions of specific patterns of errors which can be related to a specific aspect of the model.

2.6 Validation Data

The following section details our data sources and the approaches used to validate the model. We present the validation plots for given model runs and discussion of validation in the later subsections.

2.6.1 Water Level

NOAA tide gauges record 6-minute water level data at numerous sites around the perimeter of SFB (Figure 2.2). Validation of water levels at tidal time scales is an essential baseline metric for a coastal hydrodynamic model. The most relevant metrics for tidal validation in open areas of SFB are the amplitude and lag. Tides are amplified and attenuated due to geometry and frictional characteristics of SFB, a process that is captured by the comparison of amplitudes. The timing or phase of water level is driven by the same characteristics, with friction generally slowing the propagation of the tidal wave, and convergent or enclosed basin geometries causing an apparent acceleration of the tidal wave (a "standing wave").

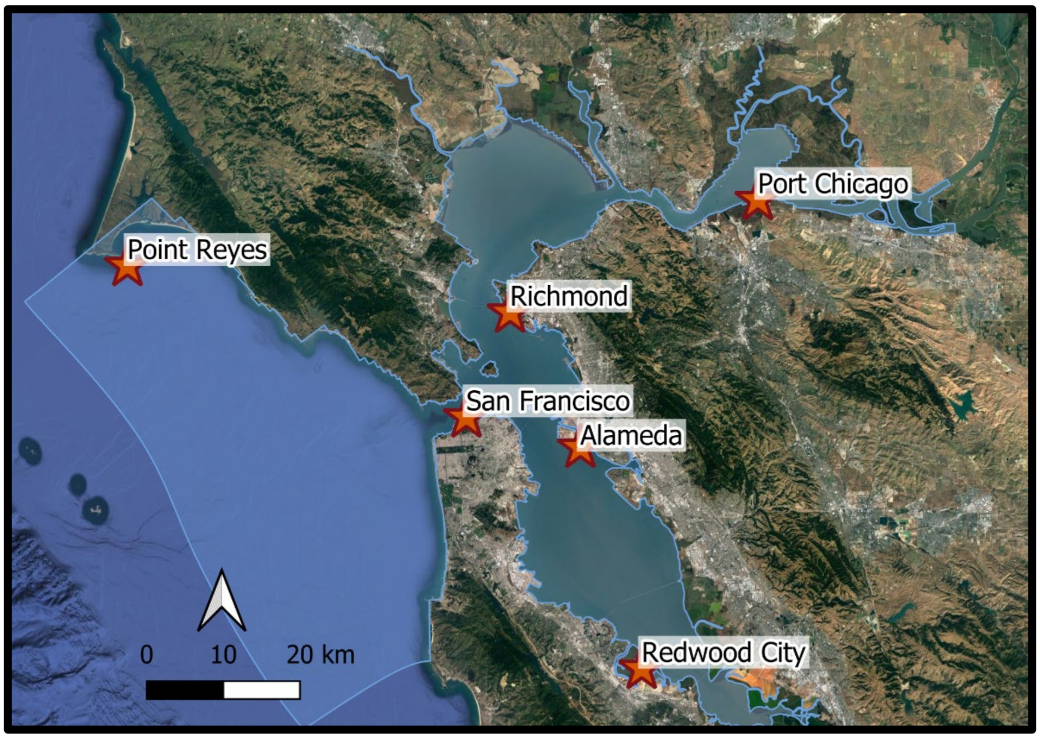


Figure 2.2. Locations of NOAA water level gaging stations

2.6.2 U.S. Geological Survey Transects

The USGS completes a cruise along the channel of SFB on a monthly to bimonthly basis, measuring vertical profiles of salinity and other parameters at 36 stations (Figure 2.3). We validate the modeled salinity field against these monthly cruises, comparing both depth-averaged salinity and estimated stratification in addition to the vertical profile. Although some USGS cruises cover only the South Bay, we include plots of model output for the full SFB in all cases for consistency and context. Depth-averaged salinity in both the model output and observations is calculated as a simple mean of the samples within each vertical profile. Stratification is calculated within each salinity profile s(z) as:

$$\frac{\partial s}{\partial z} \approx \frac{s_{max} - s_{min}}{z_{max} - z_{min}}$$

where s_{max} and s_{min} are the maximum and minimum salinity, respectively, and z_{max} and z_{min} are defined such that $s_{max} = s(z_{max})$ and $s_{min} = s(z_{min})$, essentially enforcing stable or neutral stratification. Transects are plotted south to north, with LSB on the left and the Sacramento River on the right. Thin black lines denote 5 ppt contours, with more subtle color shifts at 1 ppt intervals. The vertical dimension from the model output has been adjusted to depth below the surface. There is some variation of the profile of the bed elevation between the observations and the model due to differences in the exact location of the observations, and bathymetric details, which are not resolved in the model.

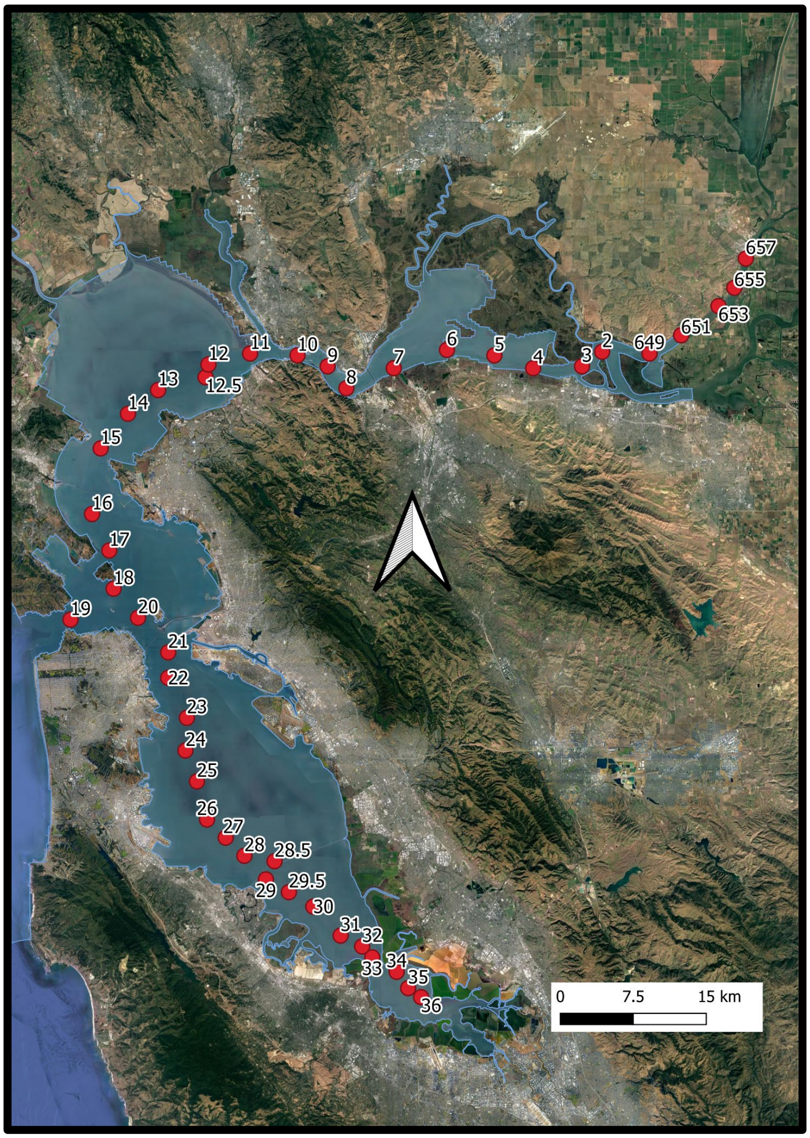


Figure 2.3. Locations of USGS stations during transect cruises

2.6.3 Salinity Time Series

In addition to the transect data shown in the previous section, USGS maintains a limited number of mooring sites within SFB (Figure 2.4). While many of the large-scale patterns of salinity in SFB can be extracted from the monthly transect data, high-frequency measurements allow for more rigorous skill metrics and analysis of daily and spring-neap time scales. Sites at the Richmond-San Rafael Bridge and the San Mateo Bridge consist of sensors at two fixed elevations. Alcatraz and Alviso each have a single sensor. Error metrics compare the depth-averaged salinity to the average of the sensor measurements.

Figure 2.4. Location of USGS moorings used for salinity time series comparison

2.7 Results

Because validation across both model runs was similar, we discuss the WY2018 validation in the text below and include additional information on the WY2014-WY2016 validation in Appendix A. Results validated include salinity and water level. Given temperature simulated within the hydrodynamics is replaced by an interpolated temperature field for our BGCM, we are not including validation plots for the modeled temperature.

2.7.1 Water Level

The model has been calibrated for tidal phase and amplitude at San Francisco by adjusting tidal conditions at Point Reyes before applying that water level as the ocean boundary condition. Water level is represented well by the model across both runs. An example time series validation at Alameda is presented in Figure 2.5. The figure shows only 7 days, but validation statistics are computed across the entire water year. Plots for the other locations can be found in Appendix A1. Metrics of model skill are presented in Table (Table 2.1). All NOAA sites except Port Chicago, (e.g., San Francisco, Point Reyes, Richmond, Alameda, Redwood City) validate very well (> 0.99). Port Chicago, similar to the previous hydrodynamic run, does not validate as well (SFEI 2020). This is due to Port Chicago's proximity to the Delta boundary. Our validation suggests model performance is strong throughout South Bay, Central Bay, and Suisun Bay, but weakens moving into Suisun Bay, where boundary effects are strong.

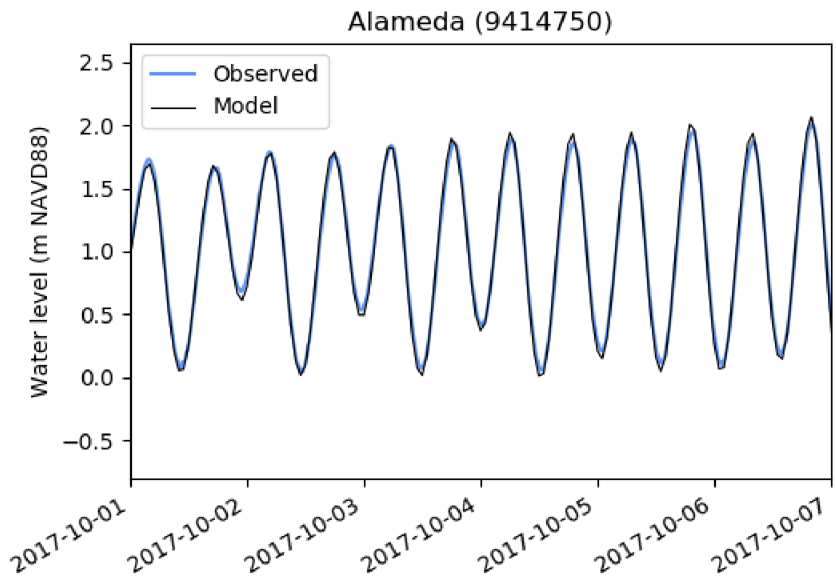


Figure 2.5. Water level time series at Alameda during the first week of October 2017

Table 2.1. Model water level performance metrics for water year 2018

Site Name	Skill	Bias (m)	r^2	RMSE (m)	Lag (min)	Amp. factor
San Francisco	0.998	-0.027	0.996	0.042	0.6	0.99
Point Reyes	0.988	0.086	0.997	0.106	6.4	0.89
Richmond	0.998	–	0.993	0.050	-4.3	1.03
Alameda	0.998	0.007	0.992	0.056	-6.8	1.02
Redwood City	0.994	–	0.976	0.117	-17.0	1.01
Port Chicago	0.903	0.001	0.887	0.374	-32.4	1.70

2.7.1.1 Longitudinal Sections

Salinity from simulations was also compared with transects captured by the USGS cruises. (Example provided in Figure 2.6, additional figures in Appendix A2). Overall, the spatio-temporal signal of salinity was reproduced well by the model. Depth-averaged salinity and stratification signals also matched observations well (see Section 3 for more detailed validation of stratification).

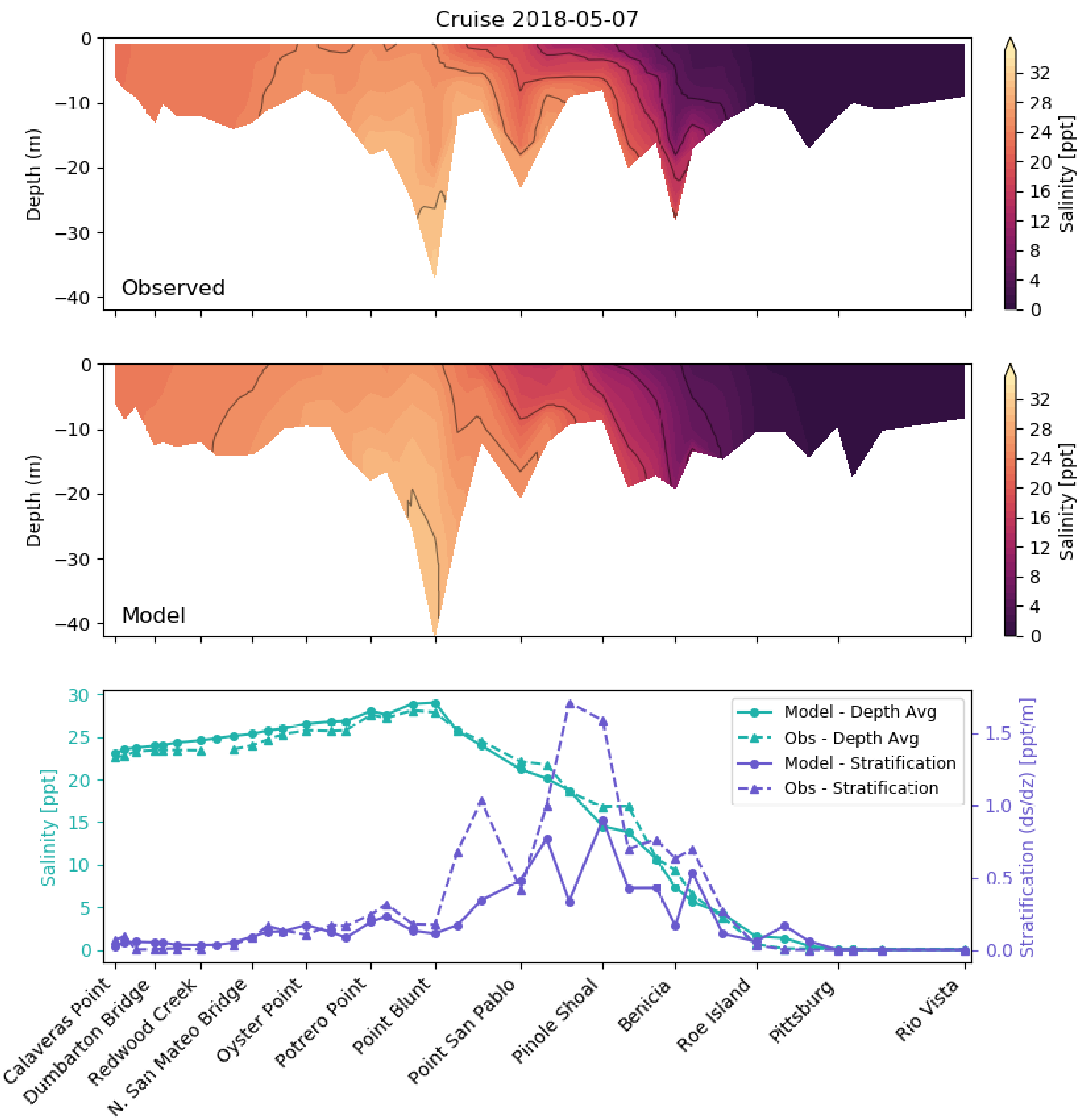


Figure 2.6. Validation with USGS cruise data, May 7, 2018

2.7.2 High-Frequency Salinity Time Series

High-frequency mooring data from USGS mooring sites were also used to validate modeled salinity. Skill metrics for WY2018 are presented in Table 2.2; plots of salinity at the Dumbarton and San Mateo Bridge USGS sensors are included in Figures 2.7 and 2.8, and plots from the additional sites and water years are included in Appendix A3. In the figures, we compare tidally filtered and max/min salinities during the tidal period, but we use the high-frequency (15 minute) time series to compute validation statistics.

Model performance is reasonable for sites within the open bay (Richmond Bridge, Alcatraz, Dumbarton Bridge, San Mateo Bridge); however, performance wanes moving southward into the sloughs of LSB. This is expected given the complex nature of the LSB complex. Note that the NMS has developed a separate hydrodynamic model focused on LSB, which includes the salt ponds and smaller sloughs.

Figures 2.7 and 2.8 show that at Dumbarton Bridge and San Mateo Bridge, during WY2018, the model underpredicts salinity during the winter (rainy season) and overpredicts salinity during the summer (drier months). This is true in other water years as well (see Appendix A3). As discussed earlier, this could possibly be attributed to an underestimation of freshwater inflow to the South Bay during the summer and an overestimation during the winter. Nonetheless, the model performance in the channel sites is acceptable with relatively small bias (< 0.5 ppt) and high-performance metrics overall. The model does not perform as well at Alviso Slough, where it was not specifically designed to capture local hydrodynamics. For the purposes of the larger Open Bay biogeochemistry the resolution of the model is adequate as evidenced by the strong performance of the model at the channel mooring locations.

Table 2.2. Model salinity performance metrics for water year 2018

Name	Skill	Bias (ppt)	r^2	RMSE (ppt)
Dumbarton Bridge	0.976	0.13	0.981	0.63
San Mateo Bridge	0.958	0.35	0.971	0.73
Alcatraz	0.898	0.21	0.918	1.16
Richmond Bridge	0.949	-0.56	0.924	1.18

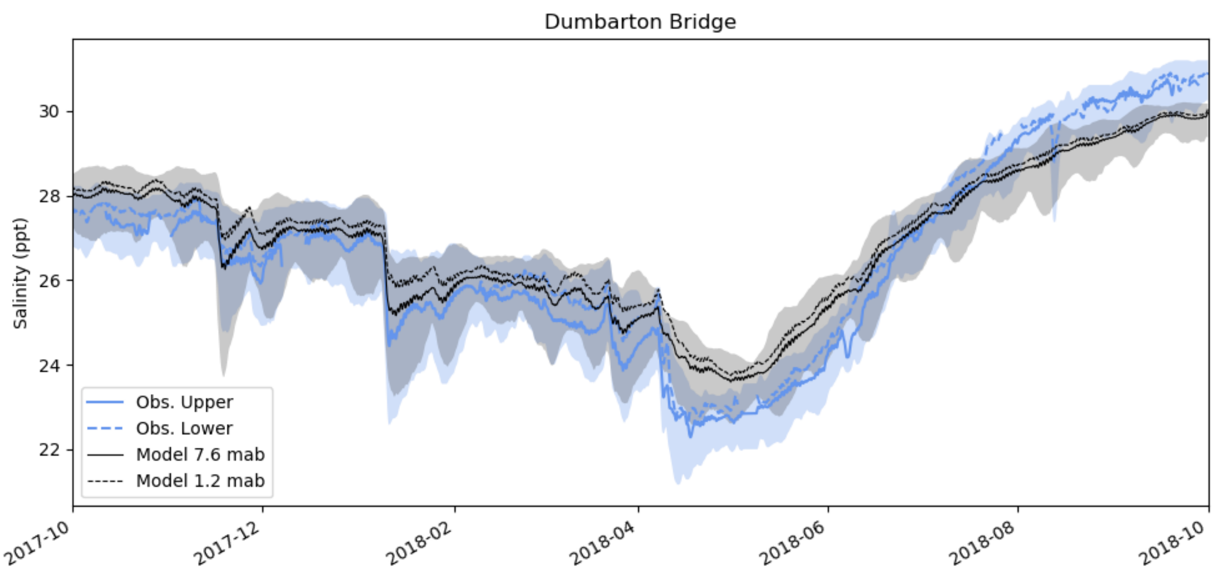


Figure 2.7. Salinity level time series at the Dumbarton Bridge (USGS) during WY2018

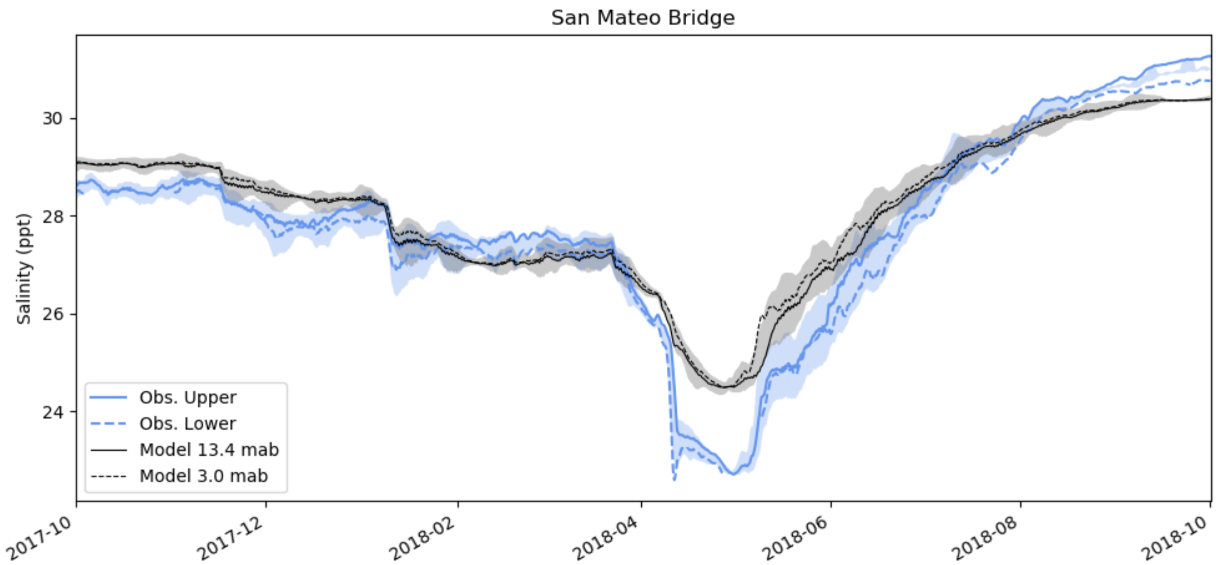


Figure 2.8. Salinity level time series at the San Mateo Bridge (USGS) during WY2018

2.8 Summary

The existing NMS Open Bay hydrodynamic model was applied to two new simulations, spanning WY2018 and water years 2014-2016, respectively, in addition to previously modeled water years 2013 and 2017, and the results were compared with available observation data. The results show that the hydrodynamic model can satisfactorily reproduce observed water levels (both phase and amplitude) over most of SFB. Similarly, the comparison of the model salinity with observations indicates that the model mostly predicts the salinity and stratification with reasonable accuracy. Model predictions near the boundaries (Suisun, for water levels and LSB for salinity) deviate notably from the observations. Including more of the Delta and LSB in the model domain could improve predictions at these locations, but the computational cost is too great at this time. The Open Bay biogeochemical model is not intended to be applied in the outer sloughs of LSB. Furthermore, a Delta-focused hydrodynamic/biogeochemical model is available for additional interpretations of the conditions near the mouths of the San Joaquin and Sacramento rivers (SFEI 2021), and development of a LSB-focused hydrodynamic model is underway, and this hydrodynamic model will serve as the basis of BCG modeling studies in future phases of the NMS. As discussed in the Appendix B, the prediction of density stratification is largely consistent with estimates of stratification derived from observations. Therefore, for the purposes of driving the Open Bay biogeochemical model, we conclude that the numeric model presented herein provides an adequate simulation of the hydrodynamic processes that have first-order effects on the biogeochemical processes.

3 Biogeochemical Model Validation

3.1 Introduction

NMS work on BGCM development began in 2016. Work carried out during 2016-2020 resulted in the development of SFB-BGCM.v2 for WY2013 and WY2017 (see Figure 1.1), and are documented in previous reports SFEI (2018, 2019, 2020). SFEI (2020) also documented the development of a fast running, aggregated grid version (Figure 3.1[5])—which takes less than half an hour for each water year simulation—was initially verified to provide adequate simulations of WY2013 and WY2017 (Agg v1) relative to the corresponding full resolution versions. The Agg v1 version was further developed in 2021 and 2022 for performing multi-year simulations and applied to further explore the parameter space resulting in an updated aggregated and full resolution model validations (Agg v2, FR v3 – see Figure 1.1).

This section describes recent updates to the SFB-BGCM, including:

- Developing multi-year model forcings for simulation of six continuous water years (WY2013-WY2018) with the aggregated grid model
- Updating WY2013 and WY2017 full resolution runs, and extending full resolution BGC simulations to WY2018, another year with significant phytoplankton activity
- Addition of a second phytoplankton group to simulate summer primary production activity. In addition to the original generic phytoplankton (called "diatoms" within DWAQ) was modified to represent the spring community, and a second generic phytoplankton (called "greens" within DWAQ) was added to represent the summer community
- Addition of benthic algae to improve model performance in shoals and margins
- Improved representation of grazing activity in the model including evaluating the sensitivity of pelagic versus benthic grazers based on additional grazer datasets for WY2013-WY2018 that became available in 2021
- Improving sediment initial conditions and sediment parameterization by comparing model performance to observed sediment fluxes from sediment incubations
- Inclusion of refractory organic matter in the water column and sediments to better represent nutrient cycling between water column and sediments
- Sensitivity analysis of key model inputs to inform uncertainties and field data collection and provide directions for further model improvements

Exploration of the importance of representing spatial and temporal variability in light extinction coefficient (K_D) was also completed. A refined approach for estimating spatially varying daily K_D will be described in a subsequent report.

3.2 Model Platform

The SFB-BGCM was developed using the DWAQ model, which is part of the Delft3D Flexible Mesh Suite of models developed by Deltares. The DFM unstructured hydrodynamic model described in Section 2 provided water surface elevations and velocities that were used as input to DWAQ for simulating transport. As described previously (SFEI 2020), two versions of the DWAQ model were developed: a

[5] Because of the large number of figures in this section, figures have been grouped at the end of this section.

high-resolution version that is capable of simulating biogeochemistry at the same resolution as the hydrodynamic model mesh, and an aggregated grid version that uses tidally averaged hydrodynamics and externally calculated horizontal dispersion coefficients (Figure 3.1). The aggregated grid model takes less than half an hour, and the full resolution model takes approximately 7 days to complete a water year simulation including spin up (14-month period) on a Linux workstation using 16 Intel Xeon E5-2680 2.40 GHz cores, communicating over MPI. Other than the differences in the grid, hydrodynamic forcings, and dispersion, the two versions use the same set of nutrient, light, temperature, and other forcing functions as described herein, with any differences in the specification of forcings or boundary conditions being due to differences in the grid resolution over which the inputs were defined.

The initial development of the high resolution DWAQ model, including the model structure, are described in SFEI (2018). Detailed descriptions of how BGC processes are implemented within DWAQ are described in the DWAQ technical reference manual (Deltares 2019a), and subsequent updates made to the model through August 2020 are documented in the previous modeling status update report (SFEI 2020). The discussions presented here are focused on the updates made to the model from August 2020-present including updates to the forcing functions, the results of the model calibration, and sensitivity analyses. These are discussed in subsequent sections, with the model development sections updated from the August 2020 report (SFEI 2020) where necessary.

3.3 Updates to the Model

3.3.1 Wind Forcing

Wind forcing was developed by interpolating hourly 10-m winds from 52 stations around the SFB-Delta region onto the model grid primarily to support the hydrodynamic model simulations. The data sources used and the approach for interpolation to the model grid are described in a prior report (SFEI 2020). The wind field interpolations were extended for BGCM simulations for WY 2014-2016 and WY2018 to support reaeration calculations within the model. In the earlier version of the model, reaeration used a fixed wind speed and a simplified representation of the reaeration coefficient. The wind interpolations were extended to both the high-resolution and aggregated grid models so that reaeration was comparably simulated at the different resolutions.

3.3.2 Light Extinction Coefficients

Spatio-temporally variable light extinction coefficients (K_D) in the FRv2 and Aggv1 models were derived from high-frequency suspended sediment concentration (SSC) data and low-frequency direct measurements of K_D (SFEI 2020). High-frequency (15 min) SSC measurements are available from 12 USGS mooring stations across the bay. Most of these stations are located in the channel. Additional high-frequency turbidity data are collected at San Mateo Bridge and Dumbarton Bridge by SFEI. The USGS also collects biweekly-to-monthly discrete measurements of suspended sediment and K_D one meter below the surface at 40 stations along the channel, from Lower South Bay through the lower reaches of the Sacramento River. These coupled SSC / K_D measurements were used to derive a formula to convert SSC to K_D at each of the high-frequency stations.

Prior to converting SSC to K_D, gaps in the SSC data, which span multiple years at some stations, were filled using a generalized additive model (GAM). GAM predictions were based on the following inputs: 1) tidally varying water surface elevation at Redwood City; 2) tidally varying velocity at San Mateo Bridge; 3) wind at the Hayward ASOS station; 4) Delta inflow; 5) Alameda Creek inflow; and 6) the station-specific seasonal trend in SSC based on USGS cruise measurements. In a final processing step, hourly K_D derived from infilled high-frequency data were forced to pass exactly through the low-frequency direct cruise measurements of K_D by applying a smoothed time-varying multiplier. The hourly mooring station K_D was subsequently interpolated and smoothed over the model grid to create the boundary condition for the biogeochemical model.

Since the light extinction coefficients are predominantly based on channel data, the K_D field does not capture the lateral variability in light extinction. Subsequently, a detailed sensitivity analysis was undertaken to assess the effect of lateral variability on the BGCM predictions in shoal and channel areas. These analyses showed that the model was moderately to highly sensitive to these changes. Capturing the lateral variability in light extinction is limited by data availability at this time.

For the FRv3 and Aggv2 models, the GAM used to infill missing high-frequency SSC measurements was improved by incorporating the following updated inputs: 1) spatially varying water levels and velocities based on tidal harmonics; 2) spatially varying 30-hour antecedent wind field; 3) Delta inflow, 4) stormwater inflow; and 5) site-specific seasonal trend. As part of these improvements, the uncertainty of the various components of the GAM was also evaluated. Future improvements in the semi-empirical approach for estimating light extinction will supplement the channel-based turbidity data set with other data such as the shoal mooring locations (installed in South Bay in 2017) and satellite imagery-based turbidity estimates to refine lateral predictions. A mechanistic sediment transport model is also being considered in the medium to long-term (~2 to 3 years) to provide spatially and temporally variable SSC for estimating light extinction.

3.3.3 Temperature

The SFB-BGCM did not include space or time varying temperatures from the hydrodynamic model, but rather used spatially and temporally varying temperature interpolated from the USGS Peterson Cruise locations. While the hydrodynamic model does simulate temperature and can provide the necessary temperature information for the biogeochemical model simulations, as discussed in Section 2 the temperature simulations from the hydrodynamic model were not the focus of the hydrodynamic calibration because stratification in the SFB estuary are predominantly salinity driven, and there are adequate temperature measurements from USGS and California Department of Fish and Wildlife to provide spatial and time-varying temperature inputs for the simulation of biogeochemistry. The interpolation methods used to derive these for the DWAQ aggregated and full resolution model grid are presented in the August 2020 model update report (SFEI 2020). The same approach was adopted for specifying the temperature for both the high-resolution and aggregated grid model simulations presented herein. The interpolated temperatures along select USGS channel stations are shown in Figures 3.2 and 3.3 for the 6-year, aggregated-grid run (WY2013-WY2018), and for the WY2018 full resolution run, respectively.

3.3.4 Updates to Water Quality Boundary Conditions

The approach used for developing nutrient loads for freshwater and publicly owned treatment work (POTW) inflows are discussed in past reports (SFEI 2018, 2020). In general, effluent monitoring data were collected by all POTWs on a once-per-month basis (flows; concentrations of major inorganic and organic N and P species), and obtained for WY2013-2018 from annual reports (BACWA/HDR 2013-2018). These data were used to calculate monthly-average loads that were assigned a mid-month date, with linear interpolation used to calculate daily loads for intervening dates. Where data gaps were encountered, values were estimated based on historical data (SFEI 2018; SFEI 2014b). Loads entering Suisun Bay from the Delta were estimated using daily discharge estimates from the Delta (California Department of Water Resources [DWR] DAYFLOW) and nutrient concentrations from long-term monitoring as previously described (see SFEI 2020 and SFEI 2015 for details). Nutrient load time-series for POTWs and the Delta boundary can be found in Appendix C.

The model uses a constant-concentration ocean boundary condition (10 uM for nitrate [NO3], 2 uM for ammonium [NH4], 0.15 uM for phosphate [PO4], 30 uM for silicon [Si] and 8 mg/L for DO). While these concentrations vary seasonally, interannually, and spatially (due to CA current and basin-scale-forcings), the constant-concentration boundary condition was deemed reasonable for the model development phase. For subsequent work (evaluating potential influence of seasonal or interannual variability in ocean forcings on nutrient fluxes; evaluating load reduction scenarios), SFB-BGCM simulations may use space-time variable ocean-boundary conditions obtained from coastal models that are being developed through a partnering coastal modeling project (Southern California Coastal Water Research Project; University of California, Los Angeles; University of California, Santa Cruz).

3.3.5 Incorporation of Spatio-Temporally Variable Temperature Adjusted Phytoplankton Growth Rate

DWAQ uses the Arrhenius formulation for simulating the effects of temperature on phytoplankton growth. Because the Arrhenius formulation is monotonic, it allows an indefinite increase in the growth rates with increasing temperature and does not provide appropriate correction at the upper end of the temperatures observed within SFB where metabolic demands cannot be met. To account for this, an asymmetric double exponential temperature correction model was adopted to account for the decline in growth rates at low and high temperatures (Thornton and Lessem 1978):

$$\lambda_T = \frac{K_1 e^{\gamma_1 (T-T_l)}}{1+K_1 e^{\gamma_1 (T-T_l)}-K_1} \frac{K_4 e^{\gamma_2 (T_4-T)}}{1+K_4 e^{\gamma_2 (T_4-T)}-K_4} \text{ for } T_1 < T < T_4$$

$$\lambda_T = 0 \text{ for } T \leq T_1 \text{ or } T \geq T_1$$

where:

$$\gamma_1 = \frac{1}{T_2 - T_1} ln \frac{K_2(1-K_1)}{K_1(1-K_2)}$$

$$\gamma_2 = \frac{1}{T_4 - T_3} ln \frac{K_3(1-K_4)}{K_4(1-K_3)}$$

In the equations above, λ_T is the temperature correction multiplier for maximum growth rate; $T_1$and T_4 represent the lower and upper temperature limits beyond which there is no growth; $T_2$and T_3 are used to define the optimal growth range; $K_1$and K_2 are factors that range from 0 to 1 with K_1< K_2 and controls the shape of the rising limb; and $K_3$and K_4 are factors that range from 0 to 1 with K_4< K_3 and controls the shape of the falling limb of the temperature function. Figure 3.4 shows the temperature correction functions used for the two groups of phytoplankton that are formulated to peak in spring and summer. These two groups are generically called diatoms and greens within DWAQ (and throughout this report) to represent spring and summer algae respectively. The parameters T1 through T4 and K1 through K4 were determined through calibration. The model presented in the August 2020 report had only a single phytoplankton group and used the Arrhenius formulation and is also shown in Figure 3.4 for reference.

3.3.6 Salinity-Based Mortality for Phytoplankton

Salinity-based mortality was included for the spring algae ("diatoms" in the model) to provide a competitive advantage under high runoff conditions in early and mid-spring. The salinity function for mortality allows a low mortality rate below a lower bound of salinity and increases it linearly with increasing salinity until a user-specified upper bound of salinity, beyond which the mortality rate does not change. The upper and lower salinity bounds, and the corresponding mortality rates for diatoms were determined through calibration (see Section 3.5). In combination with the double exponential temperature function, this helped improve the magnitude and timing of the spring bloom, particularly in years with high freshwater runoff (for example, WY2017). Salinity-based mortality was not simulated for the summer algae ("greens" in the model).

3.3.7 Sediment Initial Conditions

The initial concentrations of carbon (C), N, and P in the sediment bed were estimated using measured data from the RMP program[6]. This dataset spans 1993-2017 and contains estimates of percentage dry weight of total organic carbon and total nitrogen. Sediment P data are only available for February 2010, and sediment silica data are not available. The initial condition was developed by using a weighted average around October 2012 (model start date) and interpolating the data across the model grid. The percentage dry weight values were converted into mass concentration units (g/m^2) according to

$$w_{sed} = p_{DW} * (1 - n) * \rho_{sediment} * h_{sediment} * 1{,}000$$

where w_{sed} is the g/m^2 concentration of C/N/P, p_{DW} is the percentage dry weight of C/N/P, n is the porosity of the sediment (assumed to be 0.75), $\rho_{sediment}$ is the dry density of the sediment (2500 kg/m^3), $h_{sediment}$ is the thickness of the sediment layer (assumed to be 0.11 m), and 1,000 is a conversion factor from kg to grams. The resulting maps of the surface sediment concentrations are shown in Figure 3.5.

For DWAQ, a surficial layer was assumed to be top 2 cm for SFB-BGCM and the bottom layer was assumed to be 9 cm thick. The labile (Det C/N/P) and refractory (OOC/OON/OOP) fractions were assumed to be 1% and 99% respectively, which was assumed for other estuaries (DiToro 2001). For silica a spatially uniform initial concentration of 3 g/m^2 and 300 g/m^2 were assigned for the surficial and bottom layers during model calibration.

3.3.8 Other Biogeochemical Model Components

A benthic algal group was included in the model to better represent primary production in the shoal. Introducing the benthic algal group was initially determined through sensitivity analysis to determine whether the benthic algae played a significant role in primary production. This analysis showed that during certain times of the year, when the photoperiod is long, and the tides provide sufficiently calm conditions for sufficient light penetration to the bottom, benthic algae could contribute significantly (~10%) to primary production in the shoals. Therefore, it was included in the model as an additional state variable. There is little data available to accurately specify the initial biomass. Therefore, the initial concentration of benthic algae was assumed to be 0.5 gC/m^2 based on the sensitivity analysis.

For grazing in Central and South Bay, work during previous phases of the model development focused on developing a dynamic energy budget module to simulate primarily a pelagic grazer, guided by

[6] Downloadable from: https://cd3.sfei.org/

observations that clam biomass in South Bay decreased to de minimus levels between the late 1990s and mid-2000s (Cloern et al. 2007).

Clams were found to be important to simulate the dynamics of primary production within Delta and Suisun Bay (SFEI 2021). Unlike benthic algae, which appears to only exert a small influence on the overall primary production and nutrient cycling, the Suisun Delta model (SFEI 2021; Kimmerer and Thompson 2014; Lucas and Thompson 2012) showed that clams could have a major effect on the nutrient cycling in North Bay. Through collaboration with USGS, we recently obtained a limited set of benthic clam biomass data for portions of South Bay for some periods during 2013-2018, but clam data were still unavailable for large portions of South Bay and Central Bay, which makes the determination of initial clam biomass highly uncertain at this time. As a first step, the importance of benthic grazers in the Open Bay model was evaluated through sensitivity analysis. Future phases of model development may further explore the available clam data to determine if the uncertainty in the initial biomass estimates could be reduced either through additional data collection or by processing additional historical data sets that have been archived, if the preliminary sensitivity analyses show clams to be important drivers of primary production and nutrient cycling.

3.4 Model Calibration Approach

3.4.1 Setup and Focus

The aggregated model was used for exploring the parameter space through multi-year simulations. This model was used to run 6-year simulations (WY2013-WY2018) to explore the parameter space, both during sensitivity analysis (below) and for refining the calibration. In all, over 250 runs were performed to evaluate different parameters relative to the datasets and metrics identified below, through which an updated calibration was identified, i.e., the best-combination of processes and coefficients for capturing seasonal and spatial variability over the 6-year simulation.

Full resolution runs were then launched for WY2013, WY2017, and WY2018, using the updated model setup. WY2013 and WY2017 simulations were also presented in SFEI 2020 using the prior setup and calibration; thus, the WY2013 and WY2017 presented herein represent updates to those results.

3.4.2 Datasets for Evaluating Model Performance

Comparisons between modeled and observed data were developed for sites in SFB's deep channel (see Figure 2.3 for locations; Schraga et al. 2020) where biweekly to monthly data are available for nutrients, chlorophyll-a, and DO. In addition, data from the DWR Environmental Monitoring Program, Interagency Ecological Program (Martinez et al. 2021) were used for model-data comparisons in Suisun Bay.

High-frequency measurements of DO and chlorophyll-a are available at San Mateo Bridge, Dumbarton Bridge, and the east shoal of South Bay (Figure 2.4). At San Mateo Bridge and Dumbarton Bridge, data are available from WY2014-WY2018. At the SFEI mooring station on the east shoal of South Bay, chlorophyll-a, NO3, and DO data were collected for portions of WY2017 and WY2018.

A series of field studies were carried out in 2020-2021 (SFEI, Stanford University, University of Maryland) to collect and measure sediment chemistry and biogeochemical fluxes (N, O2) at channel and shoal locations in South Bay and LSB (Figure 3.34). Sediment samples collected in fall (Oct 2020 and Nov 2021), winter (Feb 2021), spring (May 2021), and summer (July 2021) were incubated in the lab to evaluate the flux of nutrients from SFB's sediments under dark and light conditions. The Oct 2020 sampling event was treated as a pilot test and was limited to LSB, but included more locations (3 channel locations, 2 on the east shoal, and 4 on the west shoal) relative to the two LSB locations shown in Figure 3.34. The flux estimates from these studies, while not directly paired to the model simulation periods, provide a useful dataset to assess spatial and seasonal ranges in the sediment fluxes simulated in the model. At the time this report was compiled, data from July and November 2021 sediment incubations were not yet available.

3.4.3 Approach for Evaluation of Model Performance

Direct comparison of hourly model outputs to semi-monthly and monthly grab data through traditional goodness of fit measures can be misleading in estuarine systems where mismatches in timing at a particular location can lead to poor correlations between model and data. Moreover, the NMS science and management questions are generally focused on a higher-level assessment that require reasonable simulation of nutrient cycling, primary production, and DO at a subembayment level (for example South Bay) or at a sub-subembayment level (for example, channel versus shoals) over relatively longer time periods (monthly/spring-neap to seasonal). Therefore, rather than relying purely on traditional goodness of fit measures based on point data, multiple lines of evidence were used to assess model performance as described below.

1. Model performance was assessed by visual comparisons of goodness-of-fit of the time series plots based on the USGS and DWR discrete data at nine locations spanning stations from all subembayments. The visual comparisons focused the model's ability to simultaneously capture the timing of key processes and the model's ability to reproduce the spatial gradients observed between the locations. Except for chlorophyll-a and DO, most of the nutrient data are collected only in the top 2 m. Therefore, model predictions from the surface layer (which is a few to several meters thick at the channel locations) were used for these comparisons.
2. Key model state variables and derived model outputs were compared by grouping USGS Peterson Cruise stations by sub-embayment at longer time scales that reflect the seasonal dynamics within the bay and that are more relevant to the NMS science and management questions of interest. Gross primary production from model outputs were compared to estimates derived from observed chlorophyll-a concentrations and light extinction at the USGS Peterson Cruise locations using the relationship from Cloern et al. (2007). Subembayent level box and whisker plots were used to assess performance over multiple years during the typical spring bloom period (Feb, Mar, Apr) and summer (July, August) conditions. In addition, traditional statistical goodness-of-fit performance measures were developed for these parameters at a subembayment level. The goodness of fit measure was computed on an instantaneous (paired with the time of measurement) and spring-neap time scale (best fit within a +/- 7-day window from the time of measurement). The latter helps overcome uncertainties in timing of model predictions due to uncertainties or simplifications in model inputs (for example, loads are monthly averages) and the potential effects of tides at the time of measurement.
3. At some (South Bay) stations where high-frequency data were available visual comparisons of the goodness-of-fit were performed. These comparisons focused on the model's ability to reproduce large and smaller bloom events through the growing season and the corresponding changes in DO and dissolved inorganic nitrogen (DIN).
4. Finally, sediment fluxes from the model, were averaged for the corresponding months where the data were collected over the water years where the aggregated grid and full-resolution

simulations were performed. The model simulation period (WY2013-WY2018) does not coincide with the period over which the sediment cores were collected (2020-2021), so the model results are not directly comparable to the years when the data were collected. The objective of this comparison was to assess whether the range of fluxes simulated by the model are comparable to the observations, under the assumption that the inter-annual variability in sediment fluxes are small enough to render these comparisons useful.

3.5 Model Calibration Results

The current major model parameters used in the Agg v2 and FR v3 model runs are shown in Table 3.1. The calibration results are discussed in the sections below.

Table 3.1. Biogeochemical parameters used in the D-Water Quality Model of San Francisco Bay

Process[1]	Parameter Name[2]	Parameter Value[3]	DELWAQ Default[4]	Basis[5]	Chesapeake Bay Model[6]	CE-QUAL-W2 Model Default[7]	Literature Range[8]
Water Column Nitrification	First-order rate (k1nit20) (RcNit)	0.1 per day	0.1 per day	Default	0.062 - 0.125 gN/m^3/d ~ 0.12 to 0.25/d at 0.5 gN/m^3	0.12/d	0.001 – 1.3/d
	Temperature correction factor (ktnit) (TcNit)	1.07 (36% of max @ 5°C to 140% @ 25°C)	1.07	Default	0.003 /°C[2] (15% of max @ 5°C to 93% @ 25°C)	10% of max @ 5°C to 99% @ 25°C	
Water Column Denitrification	First-order rate (k1den20) (RcDenWat)	0.005/d	0.1 per day	Calibration	Not simulated		0.05 - 0.15/d
	Temperature correction factor (ktden) (TcDenWat)	1.07 (36% of max @ 5°C to 140% @ 25°C)	1.07	Default	Not simulated	10% of max @ 5°C to 99% @ 25°C	
Sediment Denitrification	First-order rate (k1den20) (RcDenSed)	0.05 m/d	0.1 m/day	Calibration	0.1 – 0.3 m/d for aerobic layer; 0.25 m/d for anaerobic layer	0.001 m/d	1.6e-06 - 2.1 /d (Paraska et al. 2014)
	Temperature correction factor (ktden) (TcDen)	1.12	1.12	Default	1.08 (32% of max @ 5°C to 147% @ 25°C)	10% of max @ 5°C to 99% @ 25°C	
Diatoms Growth	Max. growth rate (PPMaxDiat)	1.70/d	2.3/d	Calibration	4/d for "spring" algae	2/d at 25°C	0.28/d - 3.4/d (marine diatoms)
	Temperature dependence (TcGroDiat)	56% at 5°C to 122% at 25°C	1.04	Default	80% at 5°C and to 62% at 25°C on either side of optimal temp = 16°C for "spring" algae	Not provided but recommends calibration of this parameter	
	Half-saturation for light (RadSatDiat)	22 W/m^2	25 W/m^2	Calibration	95 W/m^2 (formulation is slightly different; light saturation derived from slope and max. growth rate)	100 W/m^2	100 - 170 W/m^2 (Cole and Wells 2015)
	Half-saturation for N (KMDINdiat)	0.01 gN/m^3	0.005 gN/m^3	Calibration	0.025 gN/m^3 for "spring" algae	0.014 gN/m^3	0.009 to 0.2 gN/m^3
	Half-saturation for P (KMPdiat)	0.003 gP/m^3	0.001 gP/m^3	Calibration	0.0025 gP/m^3 for "spring" algae	0.003 gP/m^3	0.001 to 1.52 gP/m^3 with most values on the lower end
	Half-saturation for Si (KMSidiat)	0.02 gSi/m^3	0.027 gSi/m^3	Calibration	Not modeled	0	0.005 to 0.03 gSi/m^3
Diatoms Respiration and Mortality	Basal (maintenance) respiration rate (MRespDiat)	0.036/d at 20°C	0.036/d	Default	0.01/d	0.04/d	0.01 - 0.92/d

Process[1]	Parameter Name[2]	Parameter Value[3]	DELWAQ Default[4]	Basis[5]	Chesapeake Bay Model[6]	CE-QUAL-W2 Model Default[7]	Literature Range[8]
	Photorespiration (growth respiration) (GRespDiat)	11% of growth rate	11%	Default	25% of growth rate	0.04/d = 2% of default max growth rate	0.014 - 0.044/d which at a growth rate 2/d = 0.7% to 2.2% of growth rate
	Mortality rate (Mort0Diat)	Base mortality of 0.02/d	0.25/d	Calibration	Calculates a salinity related mortality and adds to basal metabolism which ranges from 0.1/d at 2 ppt and goes down to 6% of this value at 30 ppt; a first-order mortality rate is not included	0.1/d	0.03 to 0.3/d
	Mortality rate (MortSDiat)	Salinity dependent mortality of 0.07/d	0.25/d	Calibration			
	Temperature dependence for respiration and mortality (TcDecDiat)	1.07 -> 56% at 5°C to 122% at 25°C	1.07	Calibration	Varies from 62 to 117% from 5°C to 25°C for basal metabolism	Not provided but recommends calibration of this parameter	
Diatoms Settling	Sedimentation velocity for diatoms (V0SedDiat)	0.5 m/d	0.1 m/d	Calibration	0.5 m/d for "spring" algae	0.2 m/d	0.02 to 30 m/d
Diatoms Stoichiometry	N:C ratio (gN/gC): (NCRatDiat)	0.15	0.16	Calibration	0.135 for "spring" algae	0.178	0.01 - 0.36
	P:C ratio (gP/gC): (PCRatDiat)	0.02	0.02	Default	0.0125for "spring" algae	0.011	0.002 - 0.067
	Si:C ratio (gSi/gC): (SCRatDiat)	0.49	0.49	Default	Not modeled	0.4	0.4 - 0.7
	Chla:C ratio (g Chla/gC)	0.0225 (this is not a model input, but used during post-processing)	N/A	Literature	0.013 for "spring" algae	0.044	0.01 - 0.1
Green Growth	Max. growth rate (PPMaxGreen)	1.6/d	1.8/d	Calibration	7.5/d for "green" algae	2/d at 25°C	0.7/d – 9.0/d (“greens” in Cole and Wells)
	Temperature dependence (TcGroGreen)	56% at 5°C to 122% at 25°C	1.04	Default	25% at 5°C and to 92% at 20°C and 100% at 25°C for "green" algae	Not provided but recommends calibration of this parameter	
	Half-saturation for light (RadSatGree)	17 W/m^2	30 W/m^2	Calibration	113 W/m^2 (formulation is slightly different; light	100 W/m^2	100 - 170 W/m^2 (Cole and Wells 2015)

Process[1]	Parameter Name[2]	Parameter Value[3]	DELWAQ Default[4]	Basis[5]	Chesapeake Bay Model[6]	CE-QUAL-W2 Model Default[7]	Literature Range[8]
					saturation derived from slope and max. growth rate)		
	Half-saturation for N (KMDINGreen)	0.005 gN/m^3	0.005 gN/m^3	Default	0.025 gN/m^3 for "green" algae	0.014 gN/m^3	0.009 to 0.2 gN/m^3
	Half-saturation for P (KMPGreen)	0.001 gP/m^3	0.001 gP/m^3	Default	0.0025 gP/m^3	0.003 gP/m^3	0.001 to 1.52 gP/m^3 with most values on the lower end
	Half-saturation for Si (KMSiGreen)	0.02 gSi/m^3	0.027 gSi/m^3	Calibration	Not modeled	0	0.005 to 0.03 gSi/m^3
Green Respiration and Mortality	Basal (maintenance) respiration rate (MRespGreen)	0.045/d at 20°C	0.045/d	Default	0.02/d for “green”	0.04/d	0.01 - 0.92/d
	Photorespiration (growth respiration) (GRespGreen)	11% of growth rate	15%	Calibration	25% of growth rate	0.04/d = 2% of default max growth rate	0.014 - 0.044/d which at a growth rate 2/d = 0.7% to 2.2% of growth rate
	Mortality rate (Mort0Green)	Base mortality of 0.07/d	0.35/d	Calibrated	Not sensitive to salinity	0.1/d	0.03 to 0.3/d
	Mortality rate (MortSGreen)	Salinity dependent mortality of 0.07/d	0.35/d	Calibration			
	Temperature dependence for respiration and mortality (TcDecGreen)	1.07 -> 56% at 5°C to 122% at 25°C	1.07	Default	Varies from 62 to 117% from 5°C to 25°C for basal metabolism	Not provided but recommends calibration of this parameter	
Green Settling	Sedimentation velocity for greens (V0SedGreen)	0.3 m/d	0 m/d	Calibration	0.1 to 0.5 m/d for "green" algae	0.2 m/d	0.02 to 30 m/d
Green Stoichiometry	N:C ratio (gN/gC): (NCRatGreen)	0.16	0.16	Default	0.175 for "green" algae	0.178	0.01 - 0.36
	P:C ratio (gP/gC): (PCRatGreen)	0.02	0.02	Default	0.0125 for "green" algae	0.011	0.002 - 0.067
	Si:C ratio (gSi/gC): (SCRatGreen)	0.25	0.00	Calibration	Not modeled	0.4	0.4 - 0.7
	Chla:C ratio (g Chla/gC)	0.0225 (this is not a model input, but used	N/A	Literature	0.017 for "green" algae	0.044	0.01 - 0.1

Process[1]	Parameter Name[2]	Parameter Value[3]	DELWAQ Default[4]	Basis[5]	Chesapeake Bay Model[6]	CE-QUAL-W2 Model Default[7]	Literature Range[8]
		during post-processing)					
Benthic algae in layer S1 Growth	Max. growth rate (PPMaxDiaS1)	1.60/d	-	Calibration		2.0/d	
	Half-saturation for light (RadSatDiS1)	12 W/m^2	2.5 W/m^2	Calibration		75 W/m^2	
	Half-saturation for N (KMDINDiaS1)	0.005 gN/m^3	0.001 gN/m^3	Calibration		0.014 gN/m^3	
	Half-saturation for P (KMPDiaS1)	0.001 gP/m^3	0.001 gP/m^3	Default		0.003 gP/m^3	
	Half-saturation for Si (KMSiDiatS1)	0.027 gSi/m^3	0.027 gSi/m^3	Default		-	
Benthic algae in layer S1 Respiration and Mortality	Basal (maintenance) respiration rate (MRespDiaS1)	0.036/d at 20°C	0.036/d	Default		0.04/d	
	Photorespiration (growth respiration) (GRespDiaS1)	11% of growth rate	11%	Default			
	Mortality rate (MrtSedDiat)	Base mortality of 0.5/d	3.0/d	Calibrated		0.1/d	
Benthic algae in layer S1 Stoichiometry	N:C ratio (gN/gC): (NCRatDiatS)	0.16	0.16	Default		0.18	
	P:C ratio (gP/gC): (PCRatDiatS)	0.02	0.02	Default		0.01	
	Si:C ratio (gSi/gC): (SCRatDiatS)	0.49	0.49	Default		0.4	
Zooplankton	Max. Ingestion Rate (Z_JXm)	170 $J/cm^2/d$	58.5 $J/cm^2/d$	Calibration	Neither CE-QUAL-W2 nor the Chesapeake Bay model use DEB. Grazers can be modeled as state variable in CE-QUAL-W2 but is only specified as a loss term on algae and is essentially a calibration parameter. Zooplankton is not modeled in the 2015 Chesapeake Bay Model (older models did have two classes of zooplankton).		
	Food half-saturation constant (Z_Xk)	0.2 gC/cm^3	0.04 gC/cm^3	Calibration			
	Min. food threshold (Z_minfood)	0.15	0	Calibration			
	Mortality rate (Z_rMor)	0.07/d	0.2/d	Calibration			
Water Column POM Mineralization	First-order rate ($kmin_{min,20}$ and $kmin_{max,20}$) (ku_dFdcM20 and kl_dFdcM20)	0.12/d and 0.18/d	0.12/d and 0.18/d	Default	0.12/d for labile and 0.005/d for refractory	0.08/d for labile and 0.001/d for refractory	0.001 - 0.11 /d

Process[1]	Parameter Name[2]	Parameter Value[3]	DELWAQ Default[4]	Basis[5]	Chesapeake Bay Model[6]	CE-QUAL-W2 Model Default[7]	Literature Range[8]
	Temperature coefficient for mineralization (ktmin) (kT_dec)	1.047	1.047	Default	27% @ 5°C to 141% @ 25°C	10% of max @ 5°C to 99% @ 25°C	
Sediment POM Mineralization	First-order mineralization rate in DetC/N/P/Si S1/S2 (RcDetMS1/2)	0.01/d in S1; 0.03/d in S2	0.03/d for C/N/P; 0.015/f for Si	Calibration	0.01 - 0.035/d for labile; 0.0018/d for refractory	0.1/d	
Sediment POM Mineralization	Temperature coefficient for mineralization (RcDetC/N/P/ S1/Si S2) (TcBMDetM)	1.09	1.09	Default	1.1 for layer 1 and 1.15 for layer 2	10% at 5°C and 99% at 25°C	
POM Settling	Sedimentation velocity for POC/N/P (V0SedPOC1)	1 m/d POC1	0.5 m/d	Calibration		0.1 m/d	

Notes:
1. Not all processes and parameters are listed here. Only the key biogeochemical processes/parameters that were evaluated in the sensitivity analyses are included in this table. Sensitivity to temperature correction factors were not evaluated but are listed here for completeness.
2. Parameter names as used in the DELWAQ technical reference manual (Deltares 2019a).
3. As used in the base case model.
4. Default value provided in DELWAQ process reference tables (Deltares 2019b).
5. Basis for establishing the base case parameter.
6. Water column parameters are from the 2017 Chesapeake Bay Modeling Report (Cerco and Noel 2017); sediment flux model parameters are from Testa et al. (2013) and Brady et al. (2013).
7. From CE-QUAL-W2 User's Manual (Cole and Wells 2015).
8. Unless otherwise noted, literature range were derived from the range of literature values reported in Cole and Wells (2015).
DEB: dynamic energy budget
$J/cm^2/d$: joules per square centimeter per day
W: watts

3.5.1 Comparison to Discrete Data at Select Channel Locations

3.5.1.1 Aggregated Grid Model Multi-Year Runs

Time series plots of DIN, NH4, orthophosphate, and silica are shown in Figure 3.6 through Figure 3.9, respectively. Figures 3.10 through 3.13, show total chlorophyll-a, chlorophyll-a components for two algal groups, zooplankton, and DO, respectively. Zooplankton data were collected only in WY2018 and are presented in Section 3.5.1.2. In all these figures, except for Figure 3.12, results obtained using the 2020 Aggregated Model (BGCM Agg v1) are shown to illustrate changes.

The 2022 model (Agg BGCM v2) predictions of both DIN and NH4 are within the range of the data, and consistently captures the seasonal, and spatial patterns of both of these parameters (Figures 3.6 and 3.7). Observations show that both DIN and NH4 levels are highest in LSB and decrease progressively northwards until San Pablo Bay and increases slightly thereafter towards the Delta boundary. The data also show that at the channel locations there are seldom N limiting conditions for primary production. The model predictions reproduce these spatial magnitudes well. Observed DIN levels show an annual pattern with a decline in spring and summer, and an increase from late-summer through winter. Upon comparison with the chlorophyll-a (Figures 3.10 and 3.11), it is evident that the magnitude of the spring decline in DIN is often related to the magnitude of the chlorophyll-a levels. For example, WY2018 shows a prominent decline in DIN levels throughout South Bay. The model reproduces these patterns fairly well, although the spring decline in DIN levels are of relatively shorter duration compared to the data, particularly in LSB and lower portions of South Bay. In Central Bay, San Pablo Bay, and Suisun Bay, the range of the observed DIN are significantly smaller (typically 0.2 - 0.6 mg/L in most years), and are captured well in the model, particularly in Suisun Bay. While the 2020 model (Agg v1) reproduced these spatial patterns in DIN from LSB to Suisun Bay, it did not capture the seasonal variations as well as the 2022 model (Agg v2). Most notably, the Agg v1 model underestimated the spring/early summer uptake and the subsequent nitrogen cycling.

NH4 also exhibits a similar spatial pattern as DIN but is much smaller in magnitude because of the rapid nitrification within SFB (Figure 3.7). In LSB, the observed NH4 levels fluctuate substantially during the spring and summer months reflecting the primary production and subsequent detrital recycling of N both within LSB and the sloughs and salt ponds therein (the sloughs and salt ponds are not explicitly simulated in the model). The 2022 model (Agg v2) can simulate these patterns largely well. The Agg v1 model does not exhibit these variabilities in the NH4 levels and largely underestimated the observed NH4 levels. Both models underestimate NH4 levels within Central Bay in summer and early fall. Since significant bloom events are not evident from the data in late-summer and fall in most years (Figures 3.11 and 3.12), this may be indicative of sediment flux and/or uncertainty in nitrification rather than pelagic recycling from phytoplankton.

Predicted and observed dissolved orthophosphate concentrations (as P) are shown in Figure 3.8. The data show a consistent cyclical pattern of low P levels in late winter, progressively increasing through summer and fall reaching a maximum during late fall and declining in winter. The magnitude of observed P throughout SFB is much greater (> 20 μg/L, and often several 100s μg/L in summer and fall months) than what would be considered P limiting (typically 1 - 3 μg/L). Considering that the magnitude of observed chlorophyll-a levels are relatively small (5 - 30 μg/L), the P uptake during primary production has only a modest impact on observed P levels within SFB. Thus, the cyclical patterns in observed P are indicative of a strong sediment signature with P release peaking during summer and fall months concurrent with elevated sediment benthic activity. The Agg v2 model reproduces a similar cyclic pattern, but the model predictions of the late-summer/early-fall P levels are lower than the data in South Bay and LSB, in line with the data in Central Bay, and higher in San Pablo and Suisun Bay. The Agg v1 model consistently underestimated the P everywhere particularly later in the simulation indicating that the refinements in the sediment initial conditions played a major role in improving the P predictions in the Agg v2 model. Since P data were available only from February 2010, the overestimation in San Pablo and Suisun bays is an indication of the uncertainty in the interpolation of fewer locations with P data in those subembayments. Regardless of these uncertainties in initial conditions it is evident that a mechanism other than uptake of phytoplankton and subsequent recycling of detrital P is necessary to explain the rather large P levels in the water column in LSB and South Bay. One possible explanation is that the P bound to inorganic solids in SFB may be delivering P to the sediments over winter and spring (periods with high freshwater inflow and watershed solid influx). This mechanism is presently not included in the model although it can be simulated if necessary. Simulation of this mechanism will require inclusion of sediment transport, which significantly increases the complexity of the modeling effort. Because P is not a limiting nutrient and as such is unlikely to be a regulated parameter relative to N these additional complexities are not being pursued at this time.

Model predicted and observed silicate (as silicon) show large inputs from freshwater inflows during winter, and in LSB and South Bay, a prominent increase in late summer and fall, presumably from sediment recycling of biogenic silicon (Figure 3.9). The prominent signals in the data that are captured well by the model include uptake of silicate by spring algae (diatoms) in most years, but most notably in 2018 when both the data and model show a precipitous decline in silicate levels in LSB and South Bay (coinciding with a large bloom event - see also Figures 3.11 and 3.12), and a large runoff signature in the spring of 2017 throughout the system. As with P and N, the model underestimates the presumed sediment signature in the observed silicate levels in late summer and fall in LSB and South Bay. Unlike C, N and P sediment initial conditions for silicate are not based on observed sediment silicate levels since those data are not available, which renders the prediction of silicate even more challenging particularly in LSB and South Bay where there is an apparent sediment signature. Nonetheless, similar to P, water column silicate levels are much larger than what would be considered nutrient limiting for diatoms, thereby making it less of a resource management parameter of interest. While it is possible to improve the calibration of silicate through improved sediment data, the model’s ability to capture the overall seasonal signals (except summer sediment cycling in LSB and South Bay) is sufficient to meet the intended applications of the SFB-BGCM in answering the NMS’s science and management questions.

The total chlorophyll-a levels predicted by the model (sum of the spring and summer algae) are shown compared to the observed chlorophyll-a levels in Figure 3.10. Since the observed chlorophyll-a does not track individual algal populations, and detailed algal species taxonomy data are not available for the entire bay, it is not possible to break down the observations into spring and summer algae. However, it is still useful to compare the model predicted spring and summer algae levels to the total chlorophyll-a (Figure 3.11). Overall, there is a distinct spring chlorophyll-a peak in most years, with large bloom events observed in 2013 (LSB), 2016 (Suisun Bay), 2017 (South Bay and Central Bay), and 2018 (South Bay). There is also a summer bloom in most years, but the summer chlorophyll-a levels are generally substantially lower than the spring chlorophyll-a levels, and in all 6 years there were no large bloom events in summer. The 2022 BGC Model (Agg v2) does a reasonable job of capturing these dynamics, most notably: the model consistently captured the spring and summer blooms (Figure 3.11); and the model captured large South Bay bloom events in 2017 and 2018. The model does not capture the large bloom event in LSB in 2013. The model also overestimated spring chlorophyll-a levels in LSB in most years (except in 2013 as noted above), and in spring and summer in San Pablo Bay and Suisun Bay. However, the over estimations in LSB do not appear to influence predictions in South Bay significantly as the model provides a reasonable reproduction of the seasonal patterns within South Bay. As discussed in Sections 3.5.2 and 3.6, production in the shoal may have a substantial influence on the channel chlorophyll-a levels, at least within LSB and South Bay. Recognizing the uncertainties in light extinction, which are predominantly based on channel observations, it is possible that in some years (2013 and 2017) conditions in shallow areas (LSB and shoals of South Bay) may have been more favorable than what is modeled. The predominantly channel-based light extinction coefficient is a key uncertainty in the model and will continue to be an active area of model improvement efforts including improvements to the approach for calculating light extinction inputs. The most notable improvements in the Agg v2 model relative to the Agg v1 model are a better representation of the timing and magnitude of the spring bloom in most years and a better simulation of the dynamics (low algal levels) observed in late summer through late fall. The addition of a summer algae and the improved representation of the effects of temperature and salinity on algal growth likely contributed to these improvements.

A close examination of zooplankton levels (Figure 3.12) relative to the algal levels show that zooplankton play a key role in the model in keeping algal levels low (despite longer photo-period in summer months, and apparently no nutrient limiting conditions at least in the channel locations). In particular, the model predicted zooplankton shows distinct peaks that respond to algal blooms, and these peaks are proportional to the magnitude of the chlorophyll-a levels (larger in LSB and South Bay, and substantially smaller in North Bay). These patterns are comparable in both the 2020 model (Agg v1) and 2022 model (Agg v2) because the representation of the top-down control in both models is simulated in the water column (i.e., pelagic grazing only) with relatively rapid first-order grazing rates in both. The extent to which the pelagic versus benthic grazers exert top-down control in the model is a key determinant in explaining the observed chlorophyll-a levels. As discussed in Section 3.6, inclusion of clams not only affect the pelagic to benthic grazing balance within the system, but also can shift the timing of the simulated blooms particularly in shallow subembayments such as LSB. Suisun Bay and the Delta have received considerable attention on assessing the importance of pelagic versus benthic grazer communities on primary production (Kimmerer and Thompson 2014; Lucas and Thompson 2012). A parallel Delta-Suisun Bay modeling effort incorporated some of these findings and continues to assess the importance of pelagic versus benthic grazing in Suisun Bay (SFEI 2021). Findings from the Delta-Suisun modeling efforts have informed, and will continue to inform, SFB-BGCM.

DO levels simulated in the channel locations are shown compared to observations in Figure 3.13. DO levels within SFB show a seasonal signal directly related to temperature, which are more pronounced in South Bay and LSB (where waters are warmer) compared to Central Bay and North Bay (colder waters with smaller seasonal range). In addition, there are some substantially large DO swings in summer particularly in LSB and South Bay relating to primary production and (pelagic and/or benthic) organic matter degradation. The model captures the seasonal patterns but tends to overestimate the summer DO everywhere except at Suisun Bay, where the model underestimated the DO. It is possible that the low DO levels in the channel may reflect organic matter produced from outside of the channel from areas not represented well in the model (for example, in the salt ponds and sloughs of LSB and South

Bay). Future LSB-focused modeling work is planned to help address some of these uncertainties. Similarly, in Suisun the higher DO may be reflective of the primary production in shoals and/or the Delta, which appear to subside by Carquinez Strait (Station 9). The Delta-focused model described earlier (SFEI 2021) is better equipped to capture these dynamics.

3.5.1.2 Full Resolution Model: WY2018

The full resolution discussions below are focused on WY2018, which was previously not presented in the August 2020 modeling report (SFEI 2020). Updated full resolution results (FR v3) for WY2017 and WY2013 are presented in Appendix D and Appendix E respectively, along with the calibrations presented in 2020 (FR v2).

Full resolution model predictions of DIN and NH4 largely follow the aggregated grid model (Figures 3.14 and 3.15 respectively). The focused presentation of a single water year illustrates the intra-annual seasonal patterns more distinctly than what was apparent in the 6-year simulation results presented in Section 3.5.1.1. The large declines in both DIN and NH4 within South Bay in April 2018 follow the bloom event, which is evident in both model predictions and data. The full resolution model shows the range of concentrations resulting from diurnal tidal fluctuations (shaded area in Figures 3.14 through 3.20 illustrate the range, and the line represents the daily average). Even though the diurnal tides were averaged in the aggregated grid model, it nonetheless provides comparable simulations to the full resolution model illustrating its usefulness in simulating biogeochemistry within SFB. The largest differences between the two simulations occur in LSB, where the spatial resolution of the aggregated grid model is likely too coarse to accurately represent subtle changes in bathymetry and also where the overall flows are lower compared to the inflows from Delta.

The full resolution model provides better predictions of dissolved orthophosphate particularly in LSB and South Bay, but is otherwise generally comparable to the Aggregated Grid model (Figure 3.16). As with the aggregated grid model the full resolution model overestimates orthophosphate in Suisun Bay, which most likely results from a higher sediment reservoir (uncertain initial conditions). The aggregated and full resolution models also show a similar pattern in dissolved silicate levels (Figure 3.17). Both models show declines of both nutrients in April associated with spring bloom.

Figures 3.18 and 3.19 show the total chlorophyll-a (sum of both algal groups) and zooplankton simulated in WY2018. Both models show a distinct spring bloom that is brought under control by the grazers. Both models overestimate the bloom in central portions of South Bay (for example, SF Airport), but both do a reasonable job of predicting the timing and magnitude of bloom overall.

SFEI collected zooplankton data at several channel locations in WY2018. These data were analyzed for zooplankton taxa (which included both meso and micro zooplankton) and the biomass estimates were reported in dry weight units which were converted to mgC/L carbon using a dry weight to carbon conversion factor of 0.4 (as used in the model, which is the default value in DWAQ). The model results were compared to the total zooplankton (sum of all taxa on a given date and location). The simulated (sum of greens and diatoms), and the observed total chlorophyll-a are also shown on Figure 3.19 for reference. As discussed above, the model predicted zooplankton are closely linked to the timing of the bloom, and the highest zooplankton levels are simulated in LSB and lower portions of South Bay, which also corresponds to the highest observed and simulated chlorophyll-a concentrations. However, the observed zooplankton levels are far lower than the model in general in South Bay and LSB; however, zooplankton data were not collected during peak bloom time in South Bay. In San Pablo and Suisun bays, the model predicted zooplankton biomass are more in line with the data and also capture the timing of the observed increases in zooplankton biomass (in July). These results suggest that grazing is either not a dominant mechanism in LSB and South Bay channel locations, or if there is top-down control, then it occurs at the off-channel locations. As discussed in the August 2020 report, top-down control was essential to keep the phytoplankton in check through the growing season, and this was again the case in the sensitivity analyses conducted over the 6-year period (see Section 3.6). These results point to the potential importance of off-channel grazing (either pelagic or benthic). These are further explored in Section 3.6 through simulations with a benthic grazer.

Predicted DO levels are fairly consistent with data across LSB and South Bay, capturing the higher DO levels associated with primary production and the DO declines thereafter from decaying detritus (Figure 3.20). However, both models consistently overestimate DO in San Pablo Bay in summer and underestimate DO in Suisun. Considering that the model-predictions or data do not show an apparent bloom in the summer months either in San Pablo Bay or Suisun Bay, the higher DO appears to be an artifact of the Delta boundary condition for DO. This will be investigated further in subsequent model improvement efforts.

Full resolution model predictions for WY2017 and WY2013 are largely like the aggregated grid model overall (Appendices D and E).

3.5.1.3 Subembayment Level Summary of Biogeochemical Model Performance

Figures 3.21 and 3.22 show the subembayment level box and whisker plots of the aggregated grid model (Agg v2) relative to USGS Peterson Cruise measurements from WY2013-WY2018 for late winter/spring period (Feb – Apr) and mid-summer (Jul-Aug) period, respectively. The corresponding full resolution model predictions from WY2013, WY2017, and WY2018 are shown compared to data for late winter/spring and mid-summer period in Figures 3.23 and 3.24, respectively.

These boxplots show that model predictions of key state variables align with the distribution of observations reasonably well. Observations for spring months (Figure 3.21 and Figure 3.23) show that chlorophyll-a, NO3, and PO4 concentrations are generally highest in LSB and decrease progressively through Central Bay. Nutrient concentrations increase towards the Delta, but the chlorophyll-a levels remain low in North Bay. The model reproduces the observed gradients in nutrients, but overestimates PO4 levels in San Pablo Bay and Suisun Bay, which, as indicated earlier, is most likely an artifact of the sediment initial condition. Chlorophyll-a concentrations in the model capture the variability in the data reasonably well in LSB, South Bay, and Central Bay, but the model substantially overestimates concentrations in North Bay. Gross primary production (GPP) predictions from model, align well with observations, except at San Pablo Bay. The largely overlapping boxes for DO suggests that model predictions of DO align well with the data. The model underestimates DO in Suisun Bay relative to data, which indicates DO levels are relatively supersaturated within Suisun Bay. Recognizing that observed chlorophyll-a levels within Suisun Bay were low, the high DO must either be due to productivity within the Delta or within the shoals of Suisun Bay where there is no data.

Conditions in summer show a similar spatial pattern for NO3, P, and DO, but NH4 levels are highest in LSB, and GPP and chlorophyll-a in Central Bay are higher relative to other subembayments (Figure 3.22 and 3.24). Model predictions match the observations for all parameters and reproduce the observed spatial patterns well. The model also reproduces the observed seasonal differences of high chlorophyll-a and higher DO in spring versus lower chlorophyll-a and lower DO in summer at a subembayment level (as it did at the station-by-station comparison).

Goodness of fit measures computed on an instantaneous (i.e., paired with the date/time of measurement) and the best fit within +/- 7 days of the time of measurement are summarized in Table 3.2. Corresponding scatter plots showing the comparisons for model state variables and derived metrics (DIN, GPP) are shown in Appendix F. For both the aggregated grid and full resolution models, performance on chemical parameters is substantially better than the biological parameters (chlorophyll-a and GPP). Similarly, for both models the performance based on +/- 7 days is significantly better than the paired comparisons, which presents the most stringent challenge for any model. Notably model metrics for DO (RMSE generally < 1 mg/L) and DIN, which are two key management variables of interest, are reasonably good considering the spatial and temporal complexities of the SFB estuary and uncertainty in the ocean boundary and freshwater boundary conditions (see Section 3.3.4).

Table 3.2. Statistical goodness-of-fit metrics for Aggregated Grid and Full Resolution models

Parameter	Metric	Subembayment				
		LSB	SB	CB	SPB	Suisun
Agg Grid (WY2013-WY2018), paired model and data						
DIN	RMSE	0.43	0.16	0.1	0.12	0.13
	SS	0.5	0.72	0.6	0.72	0.83
	r^2	0.07	0.34	0.14	0.27	0.65
NH4	RMSE	0.08	0.06	0.03	0.05	0.07
	SS	0.49	0.42	0.45	0.59	0.63
	r^2	0.02	0.01	0.03	0.12	0.24
PO4	RMSE	0.12	0.05	0.04	0.07	0.09
	SS	0.67	0.84	0.57	0.38	0.4
	r^2	0.26	0.52	0.37	0.39	0.35
DO	RMSE	1.14	0.6	0.66	0.93	1.44
	SS	0.67	0.85	0.63	0.55	0.49
	r^2	0.24	0.54	0.2	0.12	0.53
Chl-a	RMSE	18.76	6.94	3.85	7.79	6.87
	SS	0.24	0.44	0.46	0.16	0.1
	r^2	0.01	0.07	0.09	0.03	0
GPP	RMSE	0.95	1.41	1.63	0.85	0.7
	SS	0.58	0.62	0.57	0.63	0.28
	r^2	0.13	0.34	0.25	0.24	0
Agg Grid (WY2013-WY2018), best fit within ± 7 days of data						
DIN	RMSE	0.36	0.13	0.28	0.1	0.12
	SS	0.64	0.81	0.69	0.81	0.87
	r^2	0.26	0.54	0.28	0.42	0.73
NH4	RMSE	0.05	0.04	0.04	0.03	0.05
	SS	0.8	0.64	0.53	0.73	0.73
	r^2	0.45	0.27	0.09	0.29	0.41
PO4	RMSE	0.11	0.05	0.03	0.07	0.09
	SS	0.72	0.6	0.44	0.42	0.34
	r^2	0.34	0.6	0.44	0.42	0.34
DO	RMSE	0.6	0.36	0.46	0.65	1.28
	SS	0.89	0.94	0.8	0.72	0.52
	r^2	0.66	0.83	0.49	0.37	0.52

Parameter	Metric	Subembayment				
		LSB	SB	CB	SPB	Suisun
Chl-a	RMSE	13.39	5.33	2.21	5.31	6.29
	SS	0.36	0.56	0.74	0.22	0.15
	r^2	0.04	0.16	0.35	0.04	0
GPP	RMSE	0.69	1.22	1.45	0.36	0.63
	SS	0.73	0.72	0.65	0.9	0.38
	r^2	0.46	0.56	0.44	0.7	0.06
Full Res. (WY2013, WY2017, WY2018), paired model and data						
DIN	RMSE	0.44	0.2	0.1	0.11	0.15
	SS	0.48	0.73	0.56	0.72	0.75
	r^2	0.04	0.3	0.15	0.29	0.63
NH4	RMSE	0.06	0.05	0.05	0.05	0.1
	SS	0.63	0.48	0.44	0.56	0.51
	r^2	0.17	0.04	0	0.12	0.3
PO4	RMSE	0.14	0.08	0.04	0.06	0.07
	SS	0.65	0.75	0.53	0.38	0.34
	r^2	0.21	0.46	0.39	0.2	0.47
DO	RMSE	0.77	0.55	0.69	0.93	1.5
	SS	0.83	0.9	0.65	0.6	0.51
	r^2	0.61	0.71	0.25	0.19	0.65
Chl-a	RMSE	16.2	7.2	3.73	6.53	2.87
	SS	0.19	0.47	0.57	0.22	0.25
	r^2	0	0.1	0.17	0.09	0
GPP	RMSE	1.17	1.67	1.64	0.8	0.33
	SS	0.38	0.61	0.57	0.67	0.62
	r^2	0.04	0.4	0.28	0.33	0.13
Full Res. (WY2013, WY2017, WY2018) , best fit within ± 7 days of data						
DIN	RMSE	0.39	0.17	0.09	0.08	0.12
	SS	0.57	0.8	0.6	0.82	0.81
	r^2	0.15	0.42	0.23	0.45	0.72
NH4	RMSE	0.04	0.04	0.04	0.04	0.09
	SS	0.81	0.67	0.5	0.67	0.55
	r^2	0.57	0.31	0.03	0.22	0.32
PO4	RMSE	0.12	0.07	0.04	0.05	0.07
	SS	0.71	0.79	0.58	0.42	0.35
	r^2	0.29	0.52	0.43	0.25	0.42
DO	RMSE	0.66	0.44	0.56	0.77	1.37
	SS	0.87	0.93	0.77	0.67	0.54
	r^2	0.73	0.82	0.47	0.4	0.64
Chl-a	RMSE	15.52	6.45	2.32	4.4	2.21
	SS	0.29	0.58	0.76	0.27	0.38
	r^2	0.05	0.23	0.39	0.05	0.02
GPP	RMSE	1.05	1.51	1.48	0.38	0.25
	SS	0.46	0.7	0.65	0.89	0.74
	r^2	0.24	0.53	0.46	0.64	0.33

Notes:
Concentrations for RMSE are in mg/L, except for chlorophyll-a (µg/L) and GPP (gC/m2/d). Section 2.5.1 provides a definition of model skill metrics.

3.5.2 Comparison to High-Frequency Data

3.5.2.1 South Bay Channel Moorings

The aggregated grid model predictions of chlorophyll-a and DO are shown compared to high frequency measurements at San Mateo Bridge and Dumbarton Bridge in Figures 3.25 and 3.26, respectively. The observed (discrete data from USGS cruise) and modeled NO3 concentrations are also shown to illustrate N uptake during primary production; high frequency NO3 data are not available at either of these locations. The high-frequency measurements at both locations are generally comparable to USGS cruise data, except for occasional deviations in chlorophyll-a. For example, at San Mateo Bridge in summer of 2017, the cruise data do not show the minor summer bloom indicated in the high-frequency data, and in spring of WY2018 the high-frequency measurements appear to exhibit substantially lower chlorophyll-a concentrations relative to the cruise data even though the high-frequency DO data shows super-saturated conditions (consistent with the cruise data, which also shows a sharp decline in NO3 levels around the same time providing further evidence of the bloom condition).

The 2022 BGC Model (Agg v2) does a reasonable job of reproducing these dynamics; although, the tidally aggregated grid does not show the high frequency fluctuations in the model outputs. Notable features captured in the model include spring bloom events in WY2014, WY2015, and WY2018. The model also reproduces the magnitude and seasonal variations, including supersaturation conditions in DO observed during bloom events. The model overestimates the summer bloom events in WY2014 and WY2015 (with corresponding DO supersaturation not evident in the high-frequency data) and does not capture the LSB and South Bay early spring bloom in 2013 (high frequency data were not available for this year). Improvements in the Agg v2 model over the Agg v1 model are the substantially better reproduction of the high-frequency fluctuations (even with tides averaged). While the Agg v1 model did a reasonable job of capturing the spring blooms of 2013 and 2017, it missed the spring blooms in other years, most notably in WY2018. Finally, the Agg v2 model captures the NO3 dynamics of spring decline (associated with uptake) and subsequent summer increase (recycling of detrital organic matter and sediment N release) significantly better than the Agg v1 model, indicating the effects of the nutrient cycling from the small but frequent bloom events in the observations. However, both models do not reproduce the larger increase in some years in late-summer and fall at both locations. These may be indicative of primary production and subsequent detrital recycling outside of the channel areas, possibly in the sloughs and salt ponds of LSB, which are not explicitly modeled in the SFB-BGCM. Overall, these comparisons illustrate that the aggregate grid model is reasonably successful in simulating the dynamics of primary production observed in the main channel.

Full resolution model outputs are compared to high-frequency measurements from San Mateo Bridge and Dumbarton Bridge in Figures 3.27 and 3.28 and in Figures 3.29 and 3.30 for WY2018 and WY2017, respectively. As noted earlier, the high-frequency chlorophyll-a measurements in San Mateo Bridge in 2018 appear to be inconsistent with the channel cruise data and the Dumbarton Bridge high frequency measurements where a bloom was evident in late March through early April. Like the aggregated grid model, the full resolution model also reproduces the dynamics of primary production, N uptake, and DO well in both WY2017 and WY2018. The diurnal tidal ranges provide an indication of the fluctuations in chlorophyll-a (~ ±5 ug/L), NO3 (~ ±100 ug/L), and DO (~ <±0.25 mg/L) simulated in the model, and generally lower than those evident in the high frequency data for both DO and chlorophyll-a. The high frequency data show that these fluctuations are far greater in Dumbarton Bridge compared to San Mateo Bridge, where the channel is deeper than, and farther from, the sloughs and salt ponds of LSB.

3.5.2.2 San Fransisco Estuary Institute Shoal Mooring

High frequency measurements are available from spring of 2017 and for two brief periods in spring and summer of 2018. Paired high frequency NO3 measurements were also obtained at the SFEI mooring. The aggregated grid model predictions for both water years are shown in Figure 3.31, and the full resolution model results for WY2017 and WY2018 are shown in Figures 3.32 and 3.33, respectively. As with the main channel locations, the SFEI mooring data also shows a distinct spring bloom in both years, with a bigger bloom in WY2017 compared to WY2018. Also, similar to the channel locations, the model captures WY2018 bloom, the corresponding N uptake and DO supersaturation observed in the data but misses these conditions during the WY2017 bloom. These patterns are similar between the full resolution and aggregated grid model. It is noteworthy that the WY2017 simulations presented in the August 2020 report captured the spring bloom excellently (SFEI 2020), but those same parameters did not provide good simulations for the multi-year aggregated grid runs, or WY2018. Notable differences between the 2020 and current runs are 1) a change in the light extinction inputs to hourly, and 2) removal of an adjustment factor that scaled down the older light inputs by a factor of two (as discussed in the August 2020 report, the unscaled light inputs were not successful in simulating blooms for both WY2013 and WY2017 without unreasonably high phytoplankton growth rates). Further improvements to the light extinction parameter are ongoing and can provide additional insights on the conditions that led to the WY2017 bloom.

3.5.3 Sediment Flux Results

Sediment oxygen demand (SOD) measured from 2020 and 2021 sediment incubations from South Bay and LSB cores (Figure 3.34) are shown compared to the range of SOD predicted by the aggregated grid model for the corresponding season when the sediment cores were collected (Figure 3.35). Modeled SOD is shown with and without the nitrogenous oxygen demand. The range of modeled sediment fluxes represent the minimum and maximum of the averages over the corresponding season for the simulation period (i.e., WY2013-WY2018). Modeled SOD are generally in the range of the measurements, although with few exceptions, the modeled fluxes are largely lower than the data particularly in the east shoal of South Bay (S7 through S11). Model predictions at the channel locations (C2, C5, and C6) are in-line with data in summer and fall, but the model overestimates SOD in spring and underestimates SOD in winter at nearly all locations. It is interesting that even in the fall and winter months the sediment cores exerted significant SOD. The model's underestimate in late fall and winter months is probably reflective of the averaging period, which includes months that are cooler than when the cores were collected (December [fall] and January [winter] were typically cooler than the range of temperatures indicated in the data for fall and winter[7] - see Figure 3.2 for the range of temperatures in the model). The model's consistent underestimation of SOD may be indicative of uncertainties in sediment organic matter content, lower primary production and associated organic matter loading to the sediments, or both. The 2021 sediment cores in the east shoal of South Bay (S7, S8, and S9) contained significant numbers of clam shells. This probably also suggests that clam activity may have sequestered more organic carbon locally than what might have otherwise occurred through settling of phytoplankton alone. As indicated earlier, the base case run did not explicitly model clams.

[7] Average water temperatures at the time of sediment core collection in Feb, May, Aug, and Nov of 2021 were 13°C, 20°C, 22°C and 17°C respectively.

Model predicted and observed range of N mineralization fluxes are shown in Figure 3.36. The data are predominantly limited to winter and spring of 2021, except for LSB sediment cores from October 2020. The figure shows that the model predictions align with the data quite well in spring and moderately well in winter. Similarly, model predicted denitrification fluxes are also well aligned with the data in spring (Figure 3.37). Denitrification flux is high in the LSB cores (C2, S2) in both spring and fall, which also corresponds to higher SOD. However, the correlation between SOD and denitrification fluxes are not as consistent in the cores from the east shoal of South Bay (S7, S8, and S9). The denitrification flux is very small in fall at S8 and S9, but the highest denitrification flux in the fall cores was observed at S7. This suggests that local conditions (at a sub-subembayment scale) can be considerably variable. The modeled denitrification captures the larger sub-embayment level differences between LSB and South Bay and is generally within the range of the observed denitrification fluxes.

Modeled sediment fluxes of P follow a similar pattern as N diagenesis flux, which is not surprising since they are produced through the same simulated process (as first-order transformation of sediment organic matter) within the model (Figure 3.38). Experimental P flux data from incubations were not available for comparison with modeled fluxes.

3.6 Sensitivity Analysis

The 2020 modeling report presented an initial analysis of the parameter sensitivities for WY2013 (SFEI 2020). During the validations in the previous sections, several parameter combinations were evaluated, which yielded insights into some of the more influential parameters. Based on the sensitivities observed during the validation, three sets of key sensitivities are explored in detail in this section (Table 3.3). The results presented in this section are not a formal sensitivity analysis aiming to identify parametric uncertainty. Rather, these are intended to prioritize continued validation efforts and provide guidance for areas of improvement in model inputs (for example, light extinction coefficients). These results are also intended to assess the propagation of these key sensitivities when answering the NMS management and science questions outlined in the 5-year modeling plan (SFEI 2019).

Table 3.3. Parameter ranges evaluated for sensitivity analysis

Group	Objective	Parameter	Range Evaluated[1]
Phytoplankton	Assess effect of growth rate	Maximum growth rate	+/- 25%
	Evaluate the importance of channel vs shoal in driving primary production	Growth rate in South Bay channel and shoal	Channel growth rate set to 0 vs. shoal growth rate set to zero (evaluations limited to South Bay and LSB)
	Assess sensitivity to light extinction coefficient	Light extinction coefficient	+/- 50%
Top-down Control on	Evaluate effect of grazing pressure	Maximum ingestion rate	+/- 25%

Group	Objective	Parameter	Range Evaluated[1]
Primary Production	Evaluate importance of benthic versus pelagic grazing	Model runs with and without benthic grazers (clams)	Addition of benthic grazers
Sediments	Evaluate importance of sediment initial conditions on water column nutrient concentrations	Initial C/N/P/Si concentration of sediments in Layers 1 and 2	+/- 50%
	Evaluate importance of diagenesis in freshly deposited materials in the surface sediments	C/N/P/Si diagenesis rates in Layer 1	+/- 50%
	Evaluate importance of diagenesis in slow reacting (legacy) materials in the sediments	C/N/P/Si diagenesis rates in Layer 2	+/- 50%

Note:
1. Ranges are relative to base case parameter value.

All figures for the model sensitivity analyses are presented in Appendix G. Model sensitivity within the parameter ranges in Table 3.2 were evaluated based on the following:

- 9-panel temporal plots of key model state variables (DIN, DO, chlorophyll-a, and zooplankton) in the form presented earlier for model validation
- Changes in net primary production rate during the growing and non-growing season
- Changes in oxygen consumption during the growing and non-growing season
- Changes in export of DIN to the coastal ocean assessed at Golden Gate Bridge and net assimilation of DIN within SFB

The time series plots provide the sensitivity of model validation to these parameter changes. The latter three are useful to provide an estimate of the sensitivity of the model predictions in answering higher level NMS management and science questions.

All results presented herein are based on the aggregated grid model from WY2013-WY2018. These analyses provide an indication of how key sensitivities propagate across different hydrological and meteorological conditions. In the following sections, results are presented for the metrics above in the form of summarized tables (color coded for sensitivity based on the definitions below), time series of select model state variables and as spatial maps for WY2018. For the purpose of this discussion, the propagation of parameter/input sensitivity on key model output metrics was evaluated based on the following definitions:

- A model output was considered highly sensitive if it showed a greater than a 1:1 response (~1.5x or more) relative to the perturbation in the input. For example, if phytoplankton growth was increased 25% and the model outputs for water column DO consumption showed a 50% change relative to the base case in major portions of one or more subembayments, DO consumption was evaluated to be highly sensitive to phytoplankton growth rate.
- A model output was considered moderately sensitive if it showed an approximately 1:1 response (~ 0.5x to 1.5x) relative to the perturbation in the input. For example, if phytoplankton growth was increased 25% and the model outputs for water column DO consumption showed a 12.5% to 37.5% change relative to the base case in major portions of one or more subembayments, DO consumption was evaluated to be moderately sensitive to phytoplankton growth rate.
- A model output was considered slightly sensitive if it showed an approximately milder than a 0.5x response relative to the perturbation in the input. For example, if phytoplankton growth was increased 25% and the model outputs for water column DO consumption showed a lower than 12.5% change relative to the base case in major portions of one or more subembayments, DO consumption was evaluated to be slightly sensitive to phytoplankton growth rate.

In cases where the input did not vary by a fixed percentage (for example, inclusion of a benthic grazer) output sensitivity was estimated using professional judgment based on how extensively the sensitivity propagated.

3.6.1 Phytoplankton

The overall summary of phytoplankton input sensitivity analysis is provided in Table 3.4. The paragraphs below discuss the findings in greater detail.

Table 3.4. Model sensitivity to phytoplankton parameters and processes

Input/Process	Perturbation/ Range	State Variable/Output						
		Chl-a	ZP	Prim. Prod.	DIN	DIN Export	DO	WC DO Consum.
Maximum Growth Rate	+25%	Highly Sensitive	Highly Sensitive	Highly Sensitive	Moderately Sensitive	Slightly Sensitive	Moderately Sensitive	Slightly Sensitive
	-25%	Highly Sensitive	Highly Sensitive	Highly Sensitive	Slightly Sensitive	Slightly Sensitive	Slightly Sensitive	Slightly Sensitive
Light Extinction	+50%	Highly Sensitive	Highly Sensitive	Highly Sensitive	Slightly Sensitive	Slightly Sensitive	Slightly Sensitive	Moderately Sensitive
	-50%	Highly Sensitive	Highly Sensitive	Highly Sensitive	Highly Sensitive	Moderately Sensitive	Moderately Sensitive	Highly Sensitive
Shoal/Channel Growth in South Bay	No Growth in Channel	Moderately Sensitive	Moderately Sensitive	Moderately Sensitive	Slightly Sensitive	Slightly Sensitive	Slightly Sensitive	Slightly Sensitive
	No Growth in Shoal	Highly Sensitive	Highly Sensitive	Highly Sensitive	Moderately Sensitive	Slightly Sensitive	Moderately Sensitive	Highly Sensitive

Slightly Sensitive
Moderately Sensitive
Highly Sensitive

Sensitivity of the predicted phytoplankton and zooplankton concentrations to perturbations in algal growth rate (both pelagic groups) are shown Figures G.1 and G.2, respectively. These figures show that a higher growth rate results in shifting the spring peak earlier in the season and results in a larger spring bloom in most years, and vice-versa. Zooplankton levels are also correspondingly higher or lower depending on the direction of perturbation of the growth rate. The growing season primary production was also higher/lower throughout SFB corresponding to an increase/decrease in phytoplankton growth rate (Figure G.3). Regardless of the direction of perturbation, the response was more proportionally prominent in inter-tidal areas, most notably along the eastern shoal of South Bay, Berkeley Marina in Central Bay, Richardson Bay, San Pablo shoal, and in portions of Suisun Bay. DIN levels were only moderately sensitive (Figure G.4). The cumulative DIN assimilation within SFB and the coastal export of DIN at Golden Gate Bridge were only slightly sensitive (Figure G.5) although it is apparent that with higher phytoplankton growth there is a consistent decline in the export overall in all years and vice-versa. The effect of phytoplankton growth on the water column DO is moderate, with the largest changes occurring during bloom events with the higher growth rate (Figure G.6). At the lower growth rate, DO concentrations followed a more seasonal pattern that was primarily reflective of temperature. The overall DO consumption from the concomitant nutrient recycling of the detrital algal biomass is slightly to moderately sensitive at the baseline algal levels predicted in these simulations (Figure G.7). Spatially the largest changes are apparent in the channel where it would take longer to replenish the DO lost from decaying detrital organic matter because of the water depth.

Both phytoplankton and zooplankton concentrations are highly sensitive to light extinction (Figures G.8 and G.9). At lower light extinction the spring bloom shifted earlier in the year in nearly all the years, which also resulted in the zooplankton levels increasing sooner than in the base case. The overall phytoplankton levels were also generally higher in both spring and summer. Similarly, a higher extinction shifts the spring bloom later in the year with overall lower phytoplankton and zooplankton levels. Spatial distribution of primary production also shows proportionally higher/lower response relative to the lower/higher light extinction (Figure G.10). The largest changes were observed within Suisun Bay. Some of the shallower locations (LSB, shoals in South Bay, Central Bay, and San Pablo Bay) were not as sensitive to changes in light extinction. DIN was moderately sensitive at the lower light extinction, with the late summer/early fall concentrations of DIN tracking observations significantly better (Figure G.11). Coastal export of DIN and net DIN assimilation were both moderately sensitive at the lower light extinction, but not as sensitive at the higher light extinction (Figure G.12). Coastal export was lower/higher, and concomitantly the DIN assimilation was higher/lower, relative to lower/higher light extinction. Water column DO concentration was moderately sensitive for the lower light extinction, particularly during and immediately after bloom periods (Figure G.13). DO consumption was substantially more sensitive throughout SFB at lower light extinction, but only moderately sensitive at higher light extinction (Figure G.14).

The channel and shoal phytoplankton growth sensitivities were evaluated primarily to assess uncertainty in light extinction in the shoal areas where only limited suspended sediments/turbidity data are available to derive the light extinction coefficient. Values for the shoal are largely based on channel measurements of turbidity. Thus, eliminating production selectively in the channel and shoal provides an assessment of which signature dominates in the channel where most of the model validation datasets are based. Figures G.15 and G.16 show that model predictions of phytoplankton and zooplankton in the USGS Peterson cruise locations in the channel were highly sensitive to eliminating algal production in South Bay shoal, but only moderately sensitive to eliminating production in the South Bay channel. Eliminating production in shoals also nearly eliminated the zooplankton, suggesting the tight relationship between the two state variables in the model. Note that the model does not adequately represent other sources of organic matter due to limitations in input or process resolutions (for example, from watershed runoff and organic matter resuspension from sediments) that could also potentially sustain zooplankton. Therefore, the phytoplankton and zooplankton relationship in the model is probably oversensitive relative to the real world lower trophic level dynamics in SFB. The sensitivities in the phytoplankton and zooplankton concentrations were also largely limited to South Bay and Central Bay with San Pablo Bay and Suisun Bay showing nearly no change in concentrations, which suggests limited exchange of organic matter between the South Bay and North Bay. These patterns were also evident in the spatial maps of primary production (Figure G.17). DIN concentrations (Figure G.18), and DIN export at Golden Gate and DIN assimilation within SFB (Figure G.19) were only slightly sensitive overall to channel/shoal production changes within in South Bay, with shoal production having a bigger influence on these processes. DO concentration and DO consumption in water column were moderately to highly sensitive to shoal production and only slightly sensitive to channel production (Figures G.20 and G.21). As with the other state variables/outputs the changes were largely local to South Bay (and LSB) and did not extend significantly into North Bay.

3.6.2 Top-Down Control

A summary of top-down control sensitivity runs is shown in Table 3.5. Sensitivity of model outputs to zooplankton ingestion rates are shown in Figures G.22 through G.28. Water column phytoplankton and zooplankton concentrations, and primary production are all highly sensitive to ingestion rate (Figures G.22 through G.24). Reducing grazing pressure resulted in substantially higher phytoplankton levels throughout. While increasing grazing did reduce algal concentrations and primary production it was only moderately sensitive relative to the corresponding reduction scenario. DIN concentrations were highly sensitive while DIN assimilation within SFB and export at Golden Gate Bridge were moderately sensitive to reduction in ingestion rate (Figures G.25 and G.26). DIN was generally only slightly sensitive to the increase in ingestion. Water column DO and DO consumption were moderately sensitive to reduction in ingestion, but only slightly sensitive to an increase in ingestion rate (Figures G.27 and G.28).

Table 3.5. Model sensitivity to parameterization and processes for top-down control on phytoplankton

Input/Process	Perturbation/ Range	State Variable/Output						
		Chl-a	ZP	Prim. Prod.	DIN	DIN Export	DO	WC DO Consum.
Ingestion Rate	+25%							
	-25%							
Clams	Included							

Slightly Sensitive
Moderately Sensitive
Highly Sensitive

Model response to inclusion of a benthic grazer in addition to the zooplankton are shown in Figures G.29 through G.35. As discussed earlier, only limited benthic biomass data are available in South Bay and Central Bay. Therefore, there is considerable uncertainty in the spatial interpolations used to initialize benthic grazer biomass. Nonetheless, including a benthic provides an opportunity to evaluate potential shifts in top-down control. Chlorophyll-a and zooplankton concentrations were moderately to highly sensitive depending on location (Figures G.29 and G.30). There were substantial reductions in chlorophyll-a and zooplankton in North Bay locations, whereas in South Bay and Central Bay there were increases or shifts in the timing of the spring and summer blooms. On a growing season basis, the differences in primary production were greatest in the shoals of South Bay, Central Bay, and San Pablo Bay (Figures G.31). These figures indicate that pelagic grazers can exert considerable pressure on water column phytoplankton and zooplankton particularly in shallow areas. The changes in DIN concentrations and DIN export to Golden Gate were only slightly to moderately sensitive to the addition of benthic grazer (Figures G.32 and G.33). The shifts in phytoplankton timing in South Bay and Central Bay resulted in moderate sensitivity in DO, but the changes were in North Bay were not prominent (Figures G.34). The water column DO consumption is also similarly moderately sensitive (Figure G.35).

3.6.3 Sediments

A summary of top-down control sensitivity runs is shown in Table 3.6. Figures G.36 through G.42 show the results of perturbations in sediment initial conditions. Figures G.43 through G.49 show the results of perturbations in the sediment layer 1 (freshly deposited material) diagenesis rate. Figures G.50 to G.56 show the corresponding results for the sediment layer 2 (legacy material) diagenesis rate. In general, water column model outputs were only slightly sensitive to the sediment processes and ranges of inputs that were evaluated. This suggests that while sediment nutrient cycling may be important, it is not exerting a strong influence in the dynamics of primary production. In other words, water column nutrient concentrations are sufficiently high throughout the year and seldom get to nutrient limiting conditions to make sediments an important nutrient source under current conditions. These dynamics could change if water column nutrient sources are reduced—so getting the sediment-water column coupling right would still be necessary to adequately simulate long-term system response. Section 3.5.3 showed that even though the model underestimated sediment fluxes at some shoal locations, particularly in the late summer/fall months, the fluxes were generally within the range of observation. In combination with the reduced sensitivity to the input parameters, these findings indicate that the uncertainty in sediment flux model inputs can be considered small at this time.

Table 3.6. Model sensitivity to parameterization of sediment flux processes

Input/Process	Perturbation/ Range	State Variable/Output						
		Chl-a	ZP	Prim. Prod.	DIN	DIN Export	DO	WC DO Consum.
Initial Condition	+50%							
	-50%							
Freshly Deposited Material (Layer 1) Diageneis	+50%							
	-50%							
Legacy Material (Layer 2) Diageneis	+50%							
	-50%							

Slightly Sensitive
Moderately Sensitive
Highly Sensitive

3.7 Biogeochemical Model Summary

Continued improvements to the NMS BGC model were completed including:

- Incorporation of a second phytoplankton group
- Addition of refractory organic matter class to both the water column and sediments; inclusion of a benthic algae
- Spatially variable sediment initial conditions, and
- Improvements to light extinction coefficients.

Consecutive BGCM simulations were completed over 6 water years using the aggregated grid model to refine model parameters including phytoplankton growth and mortality, top-down control, and sediment diagenesis rates.

Model performance was also verified using parallel lines of evidence including USGS Peterson Cruise data, high frequency moorings in South Bay, GPP, and sediment fluxes measured from sediment incubations. The aggregated grid model reproduced the observed seasonal dynamics of nutrient cycling and primary production reasonably well for the WY2013-WY2018 simulations. Furthermore, multi-year simulations illustrated that the sediment organic matter in the surface layer reflected the seasonal dynamics, and the larger pool of refractory organic matter in the second layer remained relatively stable indicating that the efficacy of the spatially variable initial conditions and diagenesis rates, particularly for C and N. The model's ability to reproduce observed water column P concentrations is still sensitive to the sediment source (initial conditions) and diagenesis rates. The improved parameters also provided comparable or better performance for WY2018, WY2017, and WY2013 when used in the full resolution model.

Sensitivity analyses conducted with the aggregated grid model indicated high sensitivity to phytoplankton growth, light extinction, and top-down control. The model is not as sensitive to uncertainties in sediment initial conditions or the diagenesis rates (except for P), which is a notable improvement over the sensitivity analysis presented previously in the Aug 2020 report, wherein the water column concentrations were moderately sensitive to sediment initial conditions.

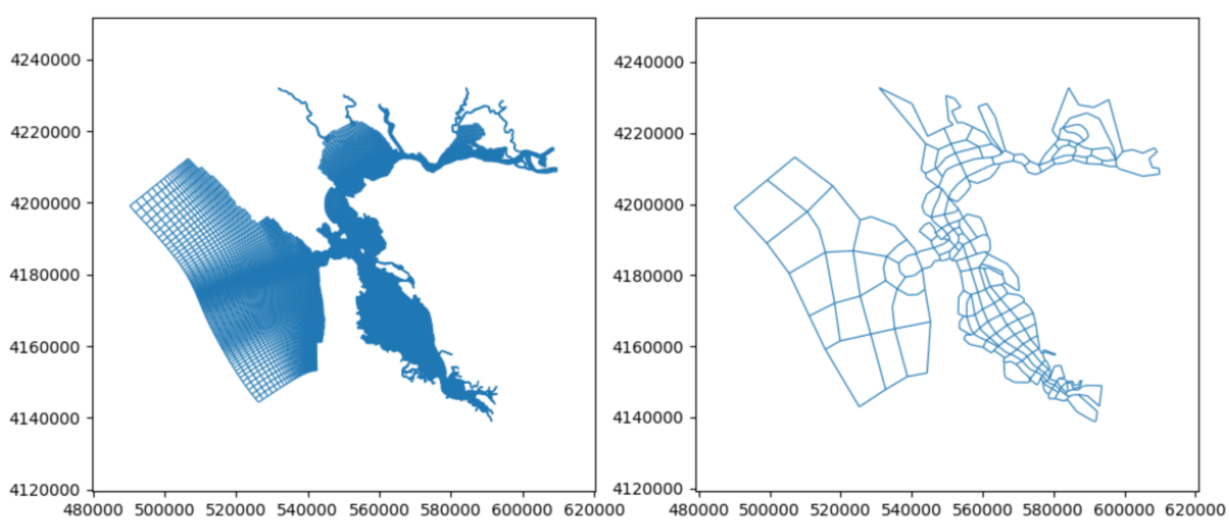


Figure 3.1. Full-resolution model grid (left) and aggregated model grid (right)

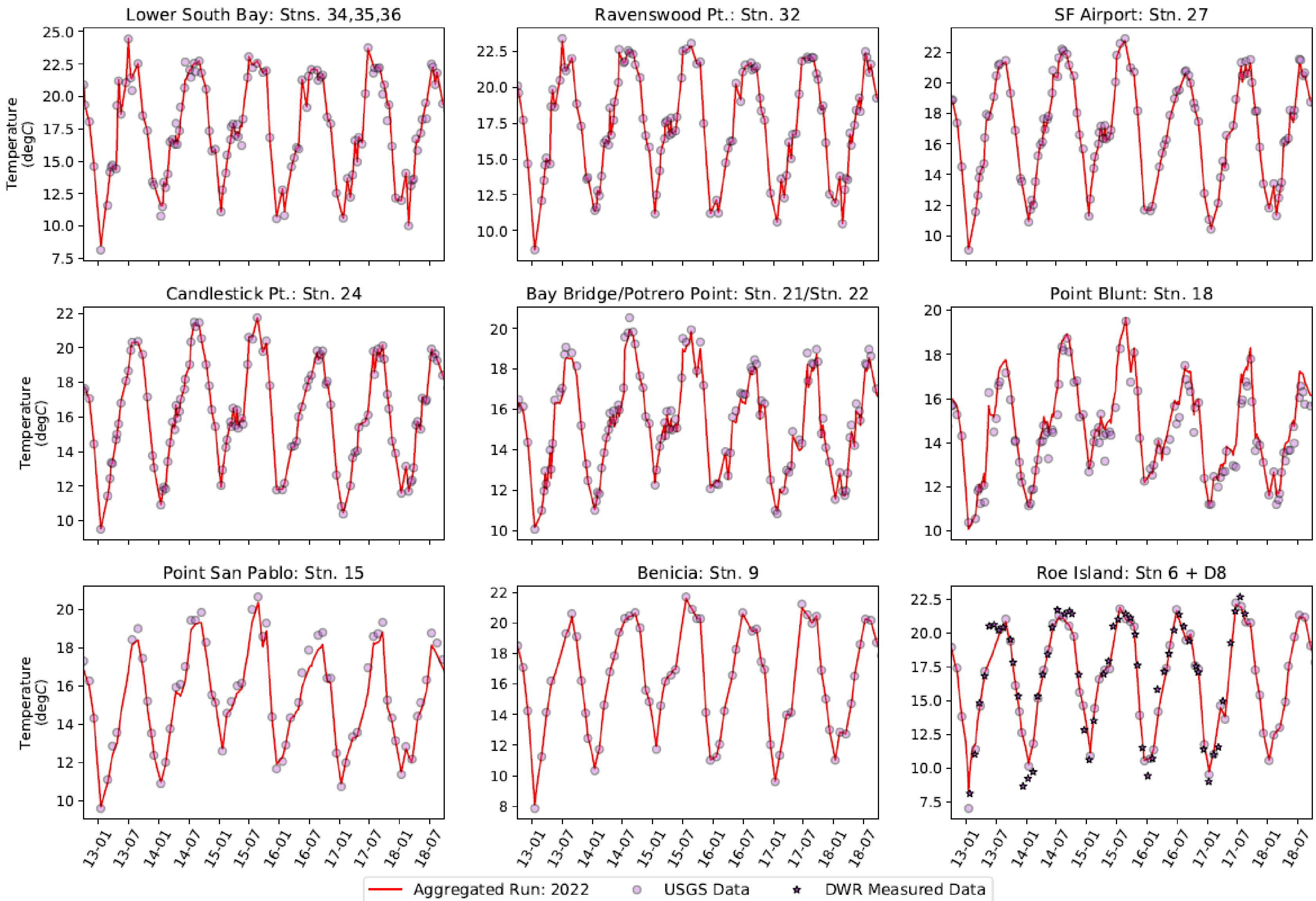


Figure 3.2. Interpolated temperatures specified for water year 2013-2018 simulations for the aggregated grid model

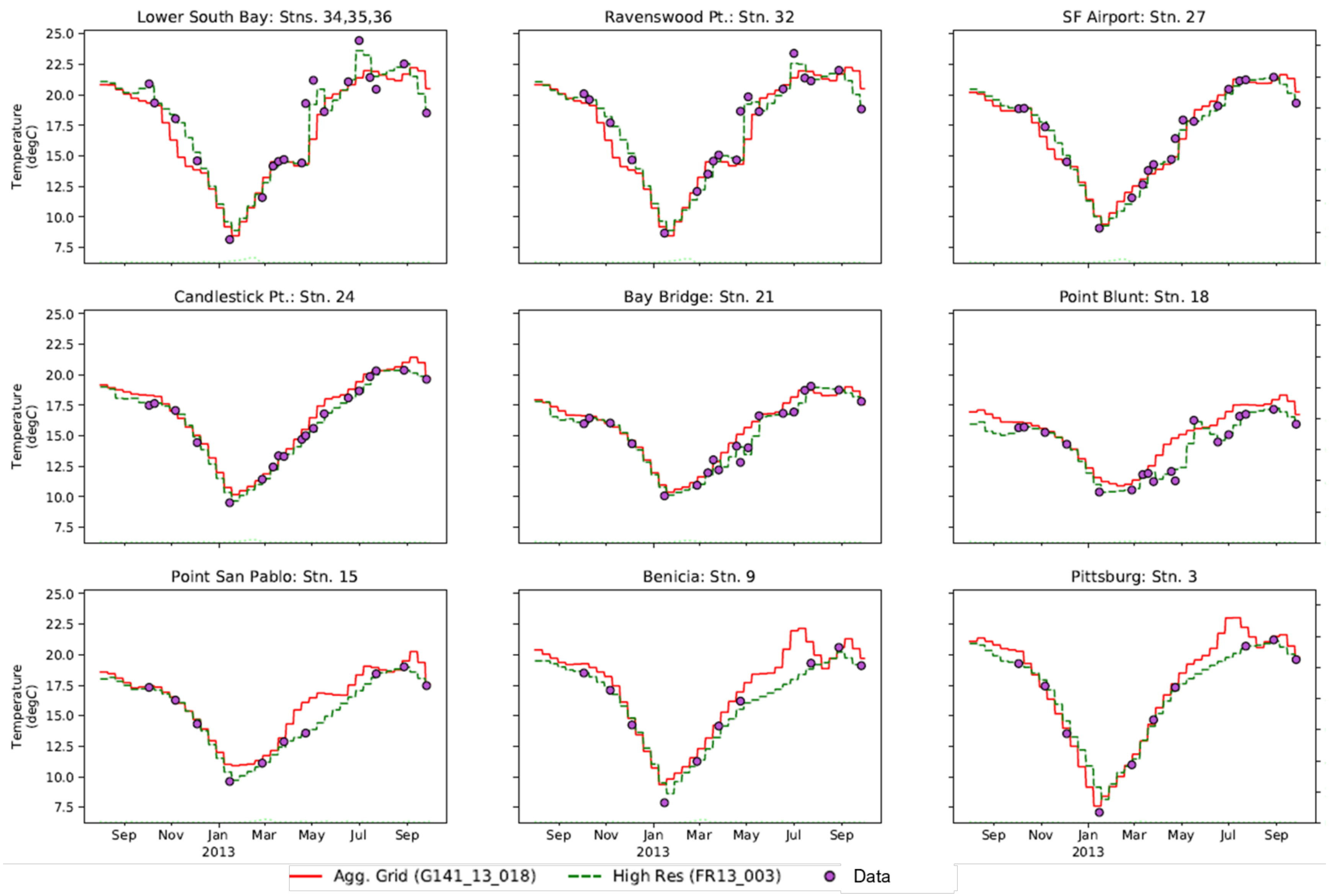


Figure 3.3. Interpolated temperatures specified for water year 2018 simulations for the full-resolution model

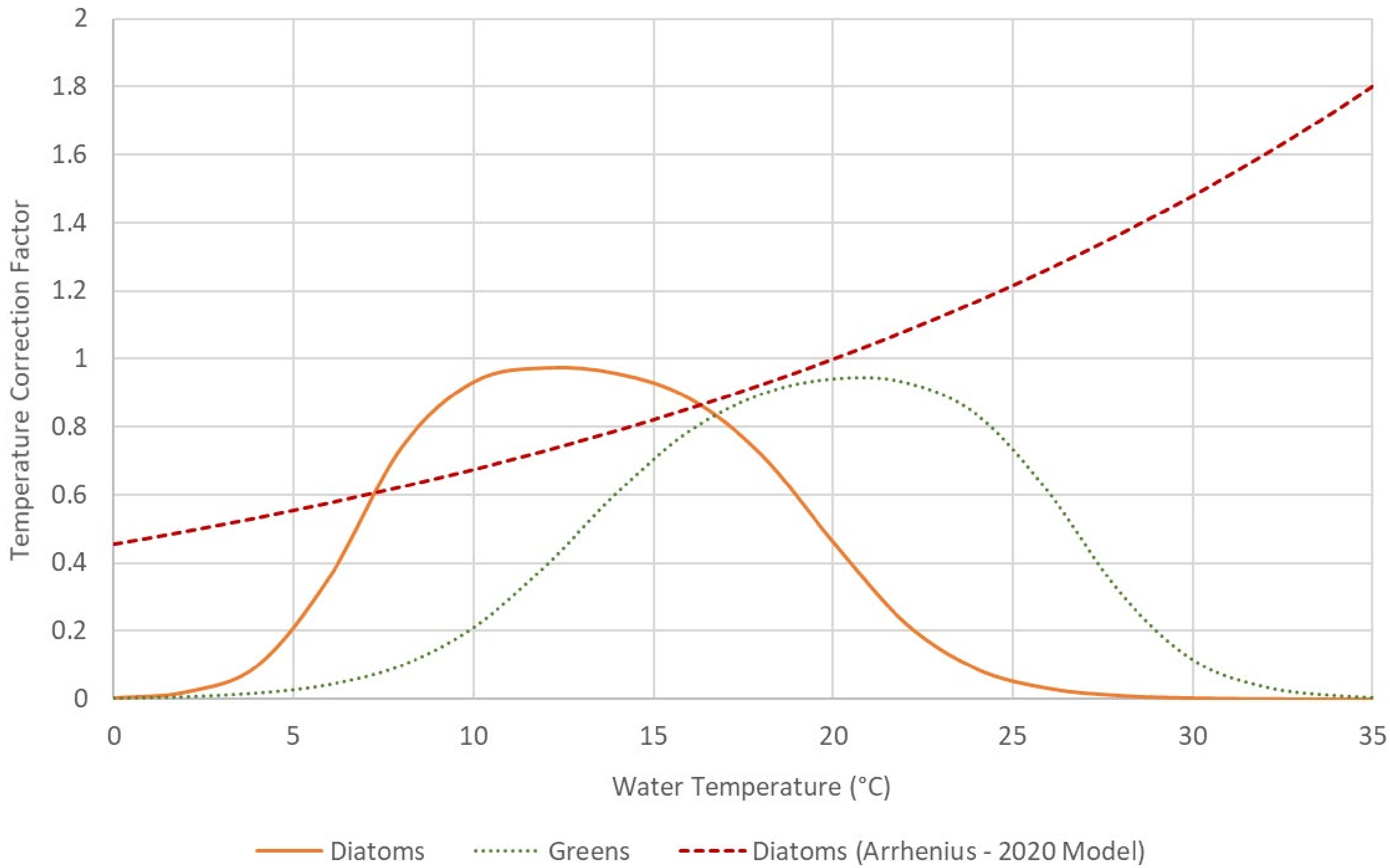
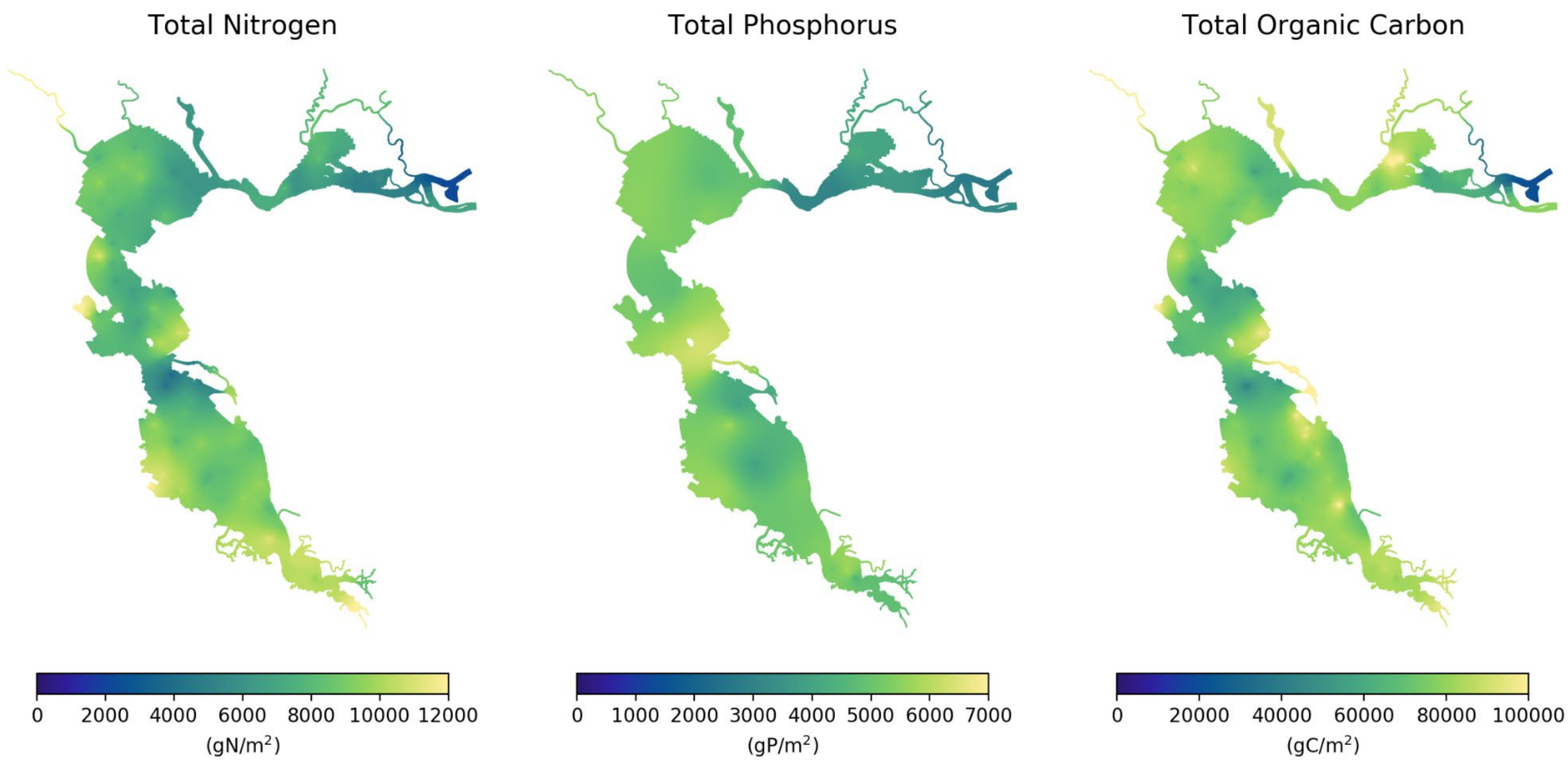


Figure 3.4. Temperature correction functions for phytoplankton growth rate



Note: Data were collected as part of the Regional Monitoring Program and represent surface sediment data collected from 2010-2020.

Figure 3.5. Initial sediment concentrations developed using measured surface sediment data

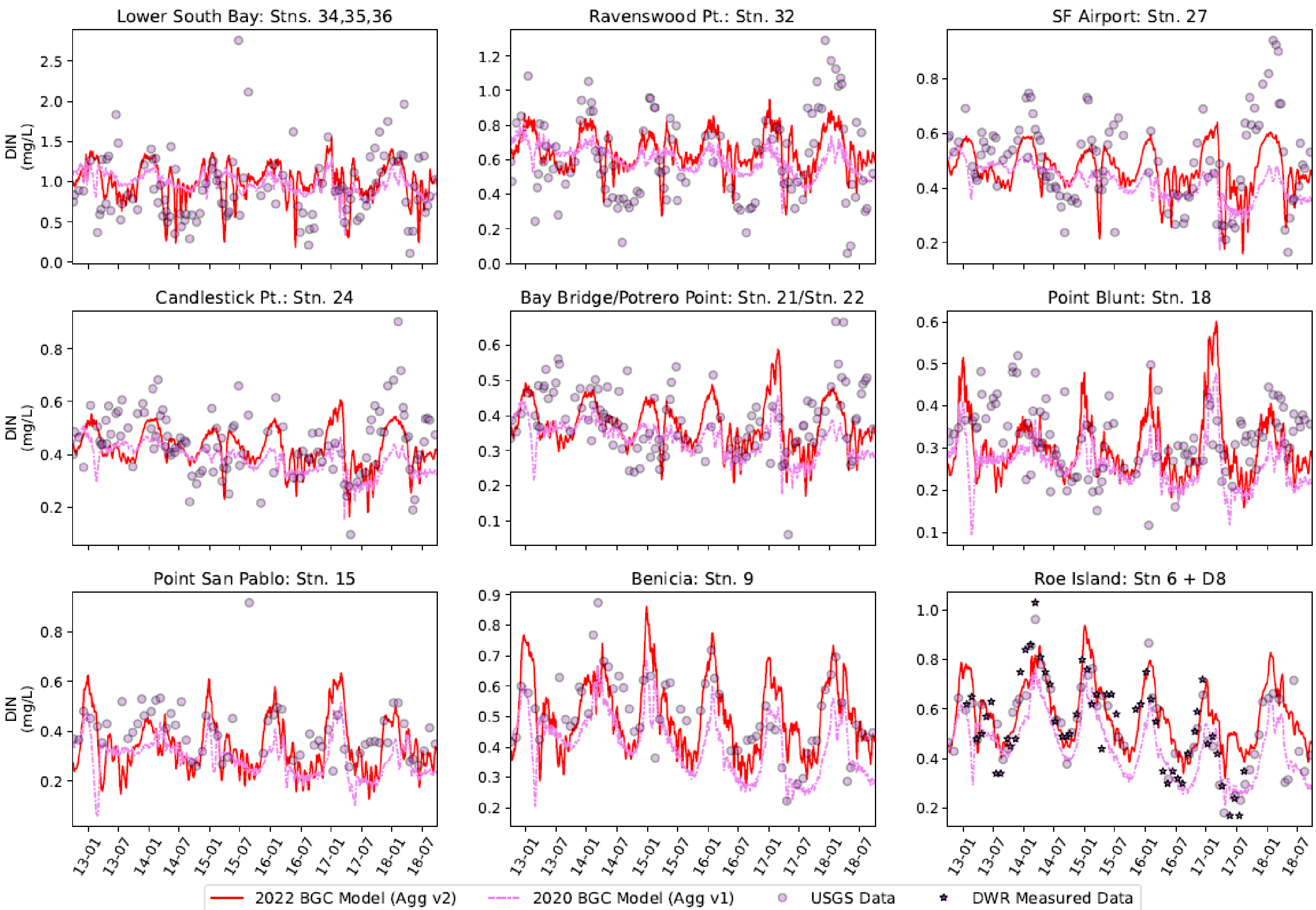


Figure 3.6. Predicted and observed dissolved inorganic nitrogen (DIN) concentrations at select USGS stations along the channel for the aggregated grid model WY2013-WY2018 simulation

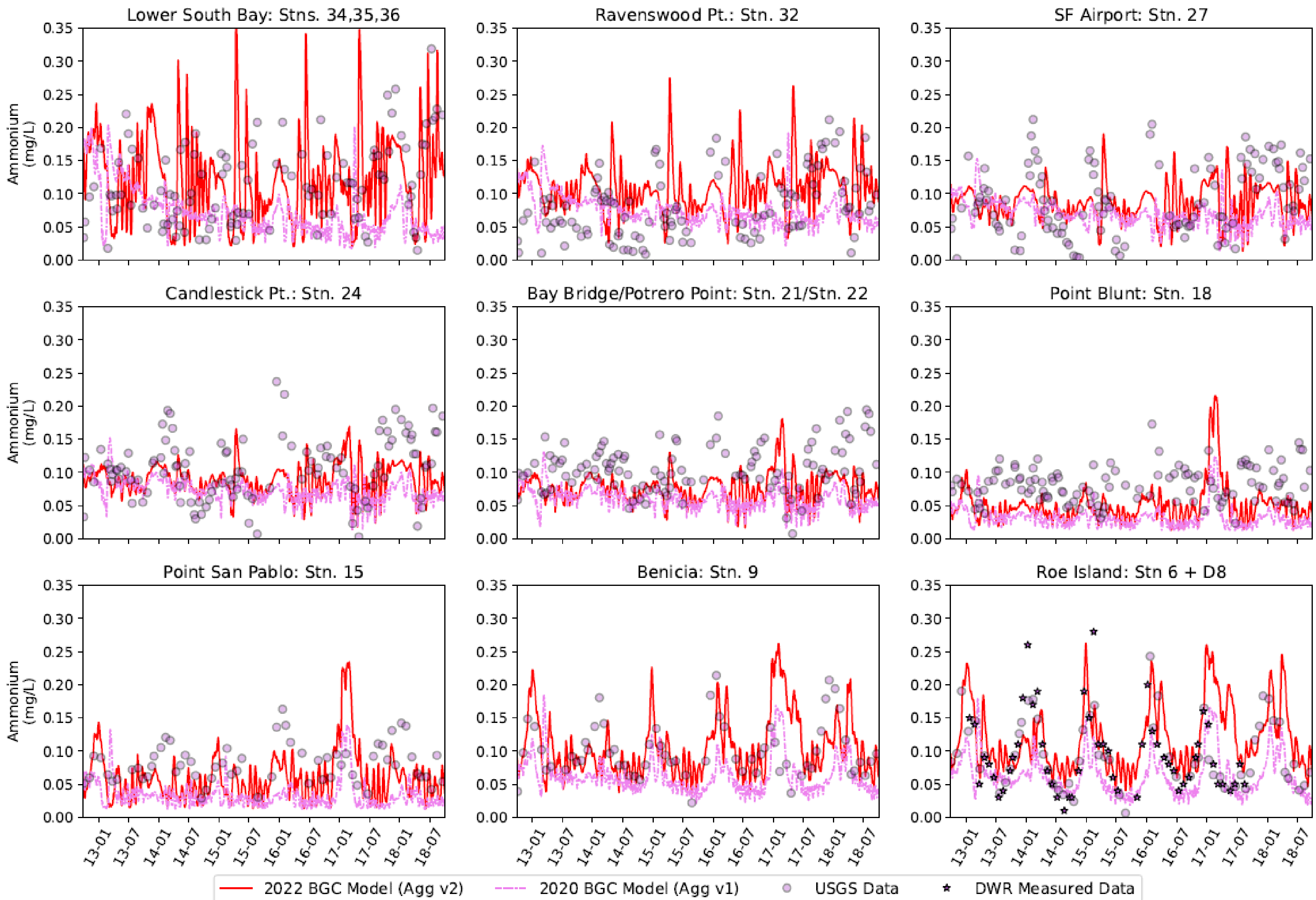


Figure 3.7. Predicted and observed ammonium concentrations at select USGS stations along the channel for the aggregated grid model WY2013-WY2018 simulation

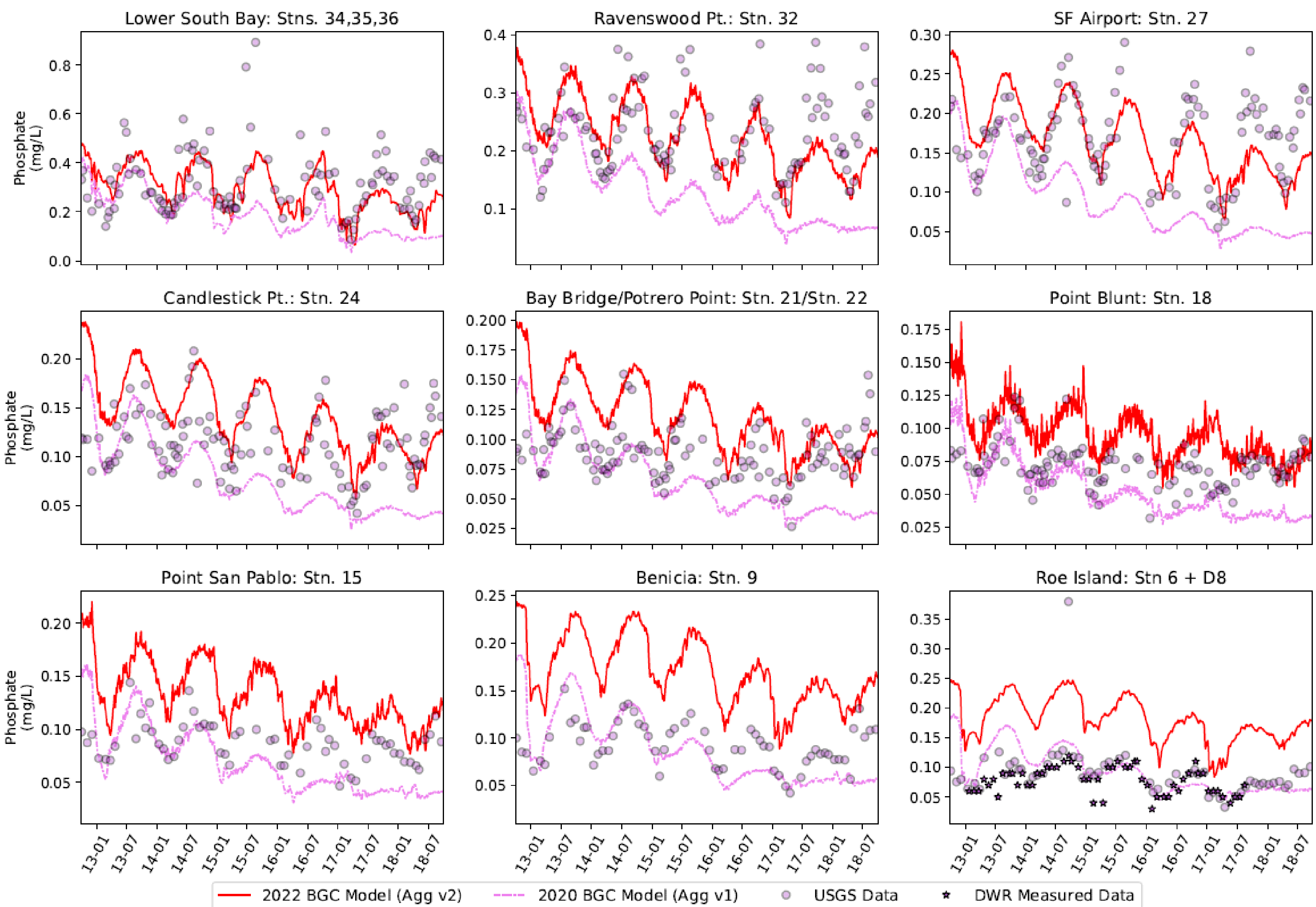


Figure 3.8. Predicted and observed dissolved orthophosphate concentrations at select USGS stations along the channel for the aggregated grid model WY2013-WY2018 simulation

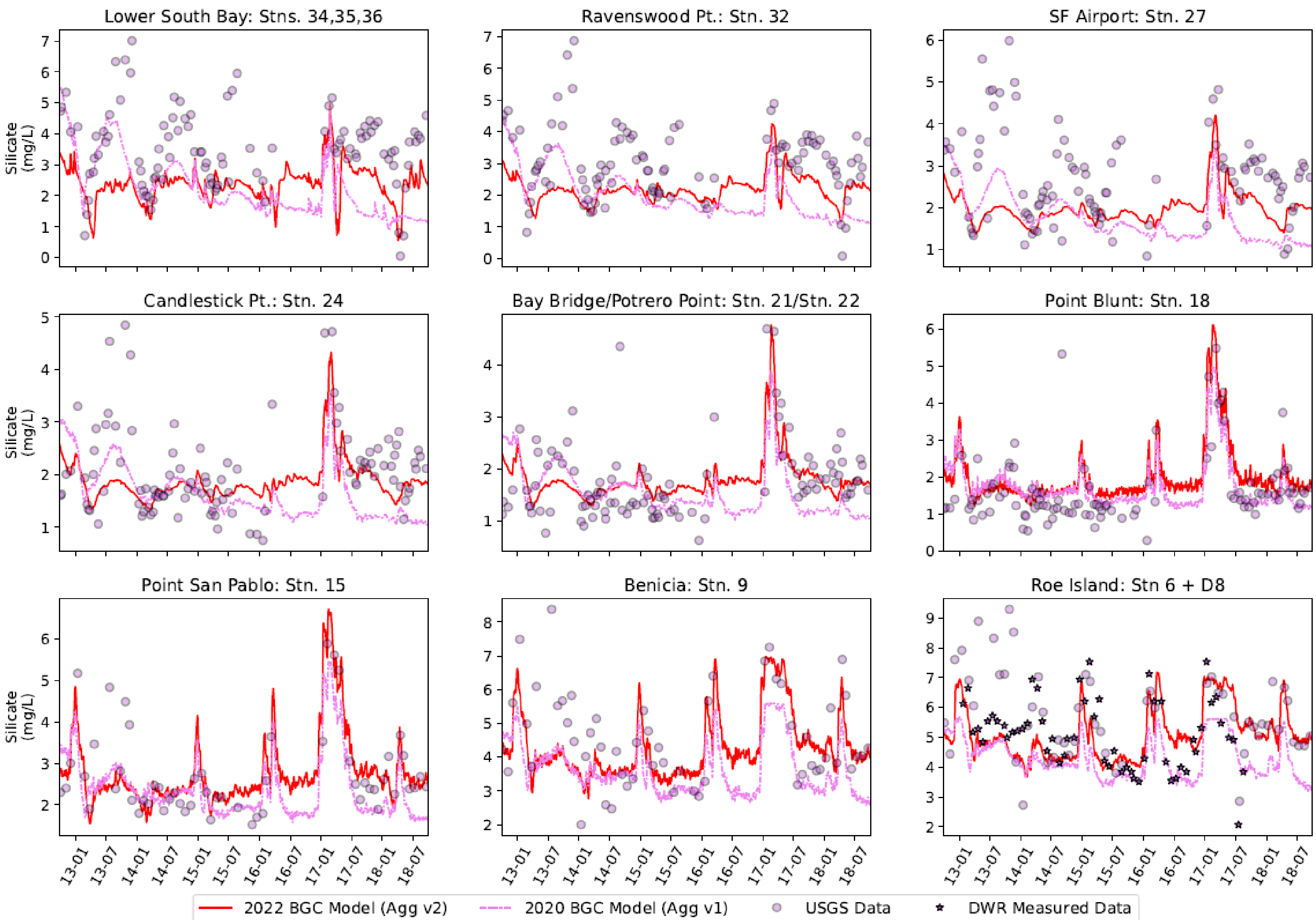


Figure 3.9. Predicted and observed silica concentrations at select USGS stations along the channel for the aggregated grid model WY2013-WY2018 simulation

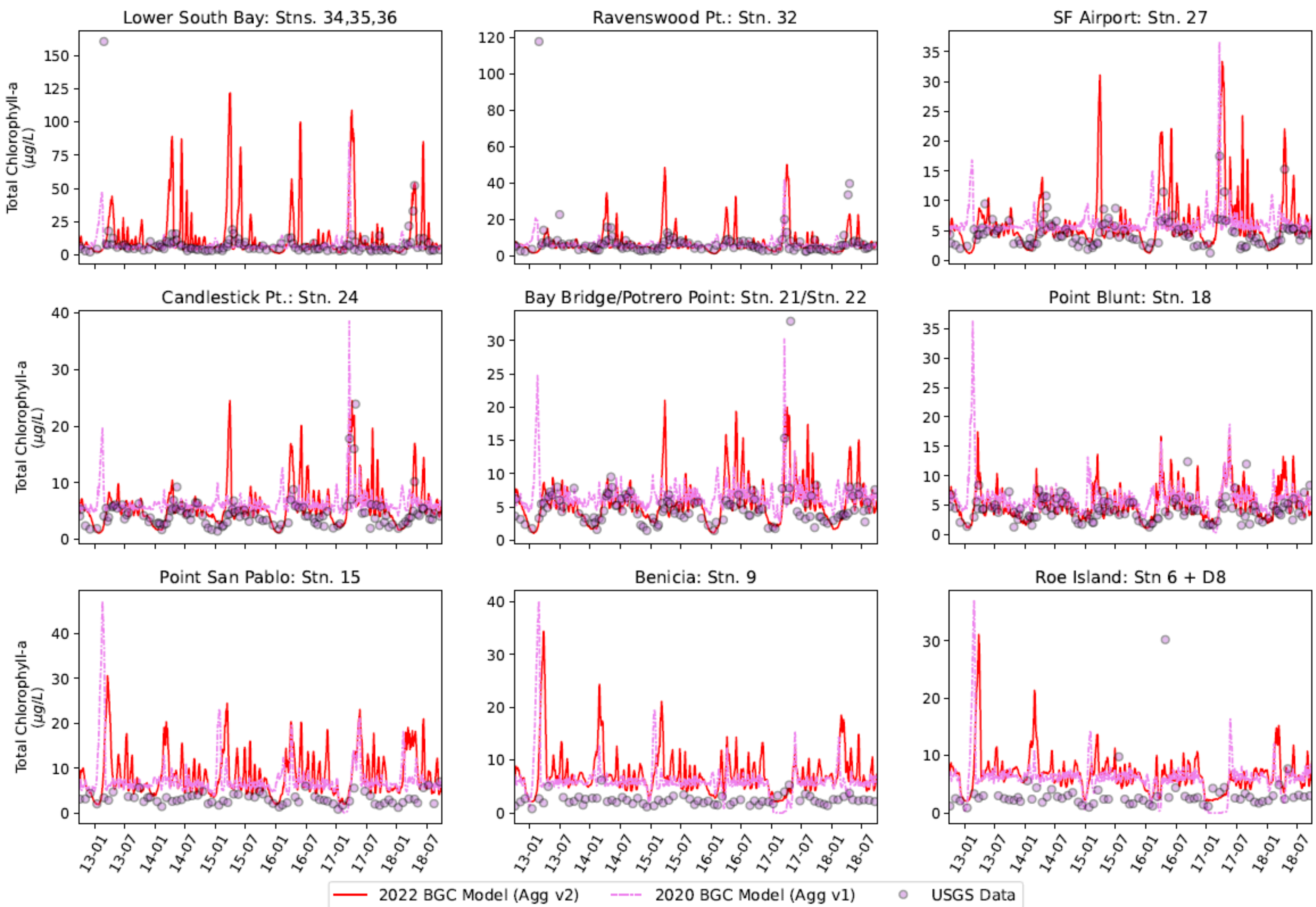
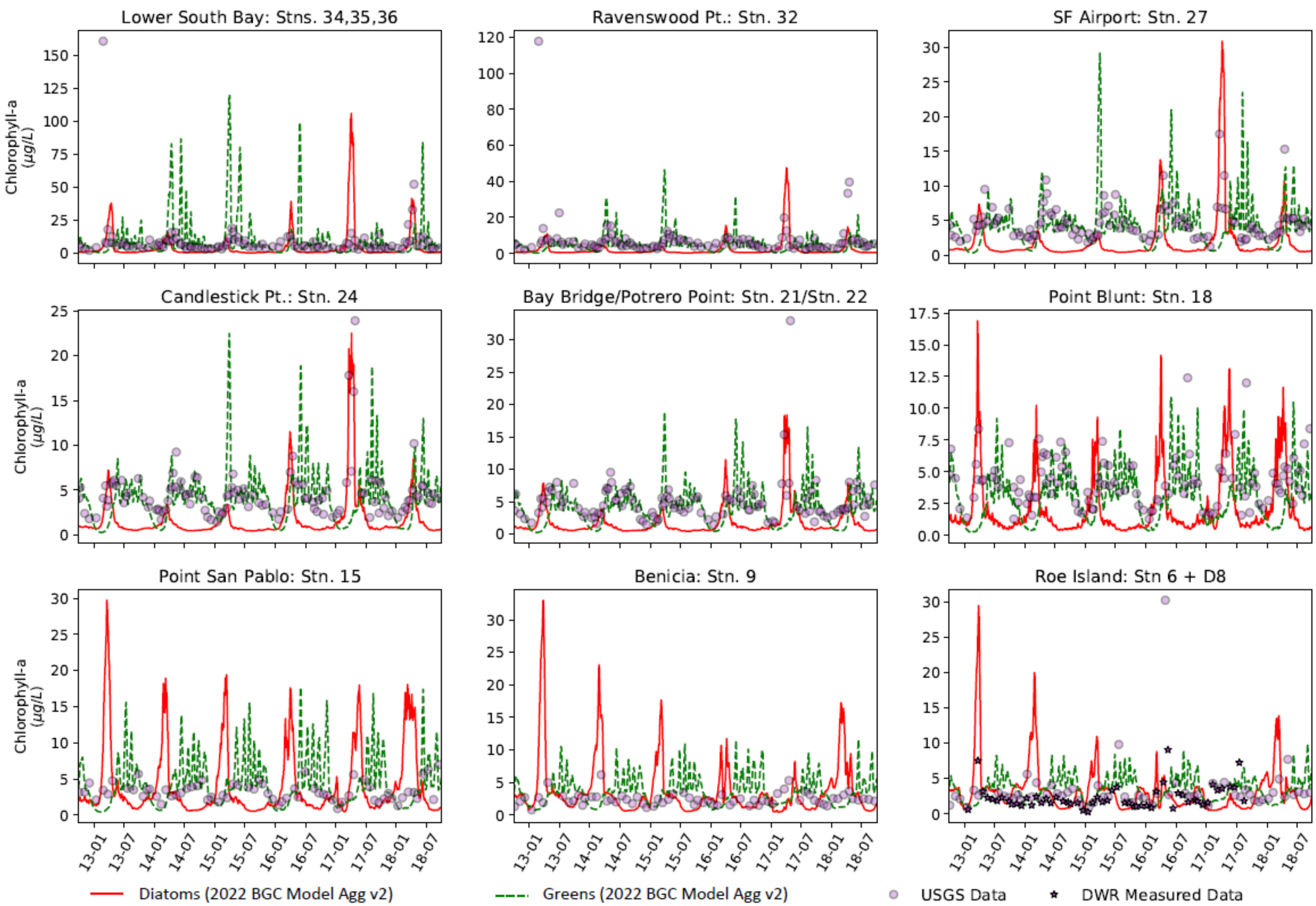
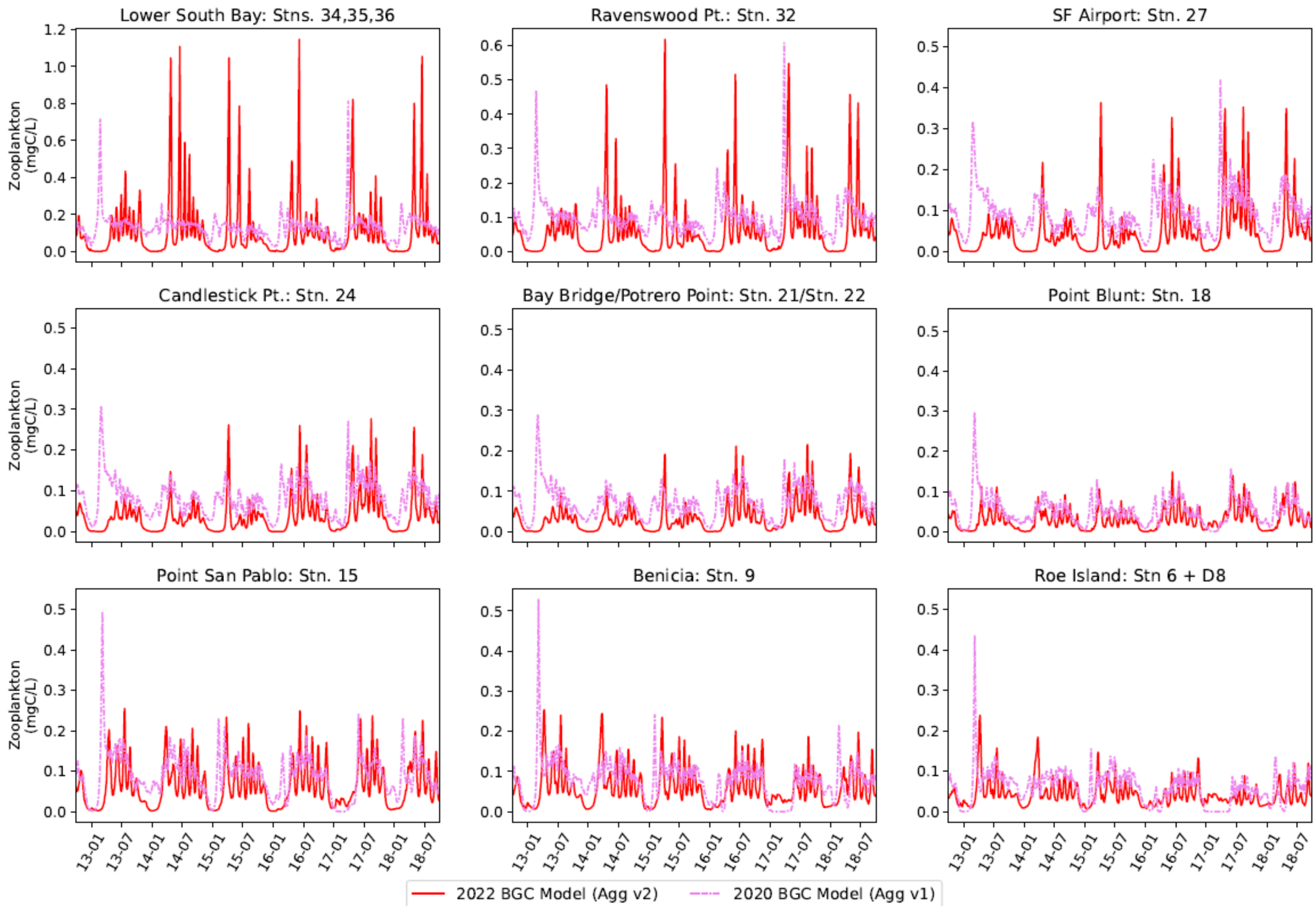


Figure 3.10. Predicted and observed total chlorophyll-a concentrations at select USGS stations along the channel for the aggregated grid model WY2013-WY2018 simulation



Note: The total chlorophyll-a predicted by the model is the sum of the two algal groups simulated; data shows total chlorophyll-a since estimating the fraction of individual algal groups was not possible.

Figure 3.11. Predicted chlorophyll-a for each algal group compared to observed total chlorophyll-a concentrations at select USGS stations along the channel for the aggregated grid model WY2013-WY2018 simulation



Note: Zooplankton data are only available from WY2018 and are presented in Figure 3.19.

Figure 3.12. Predicted zooplankton concentrations at select USGS stations along the channel for the aggregated grid model WY2013-WY2018 simulation

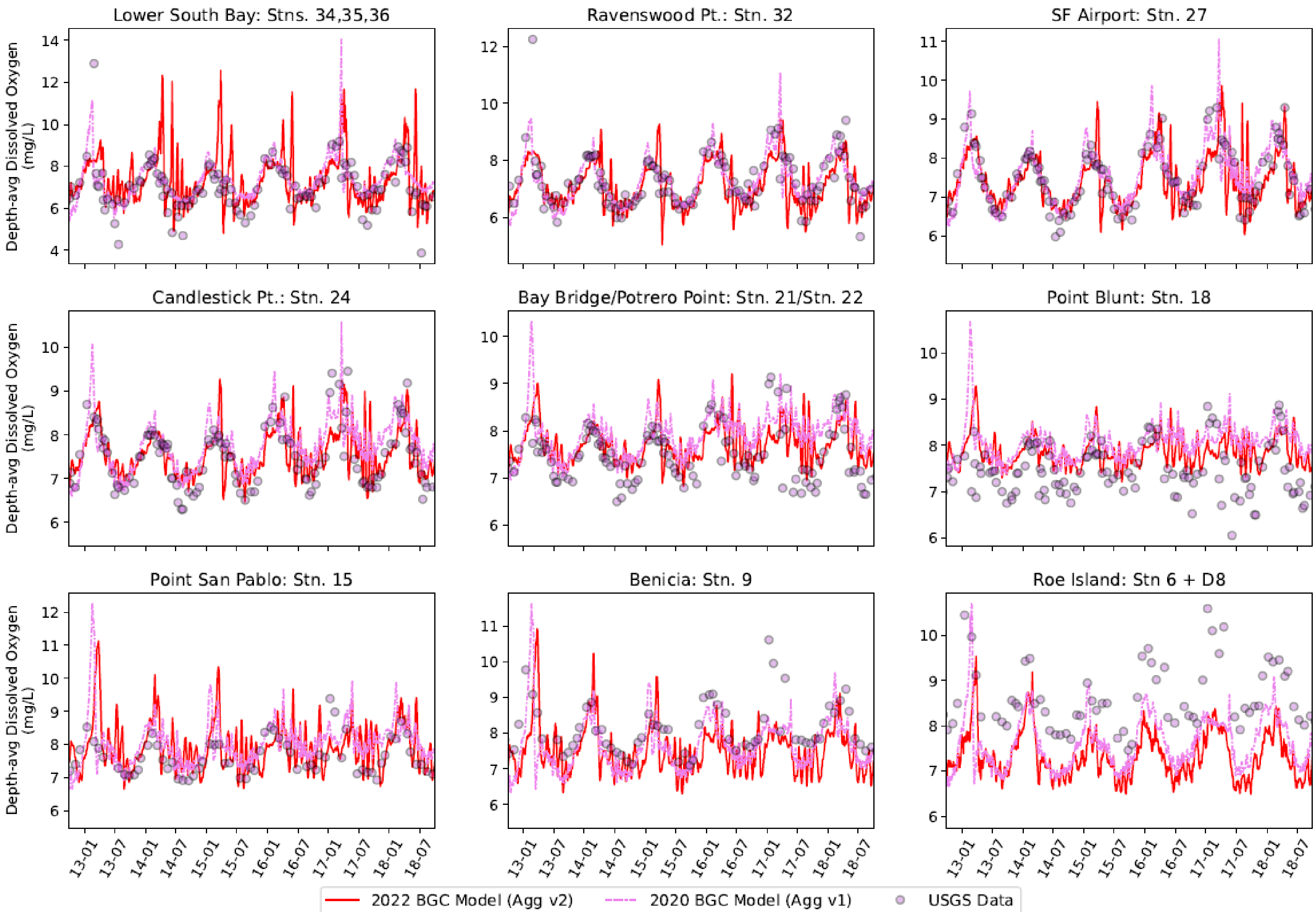
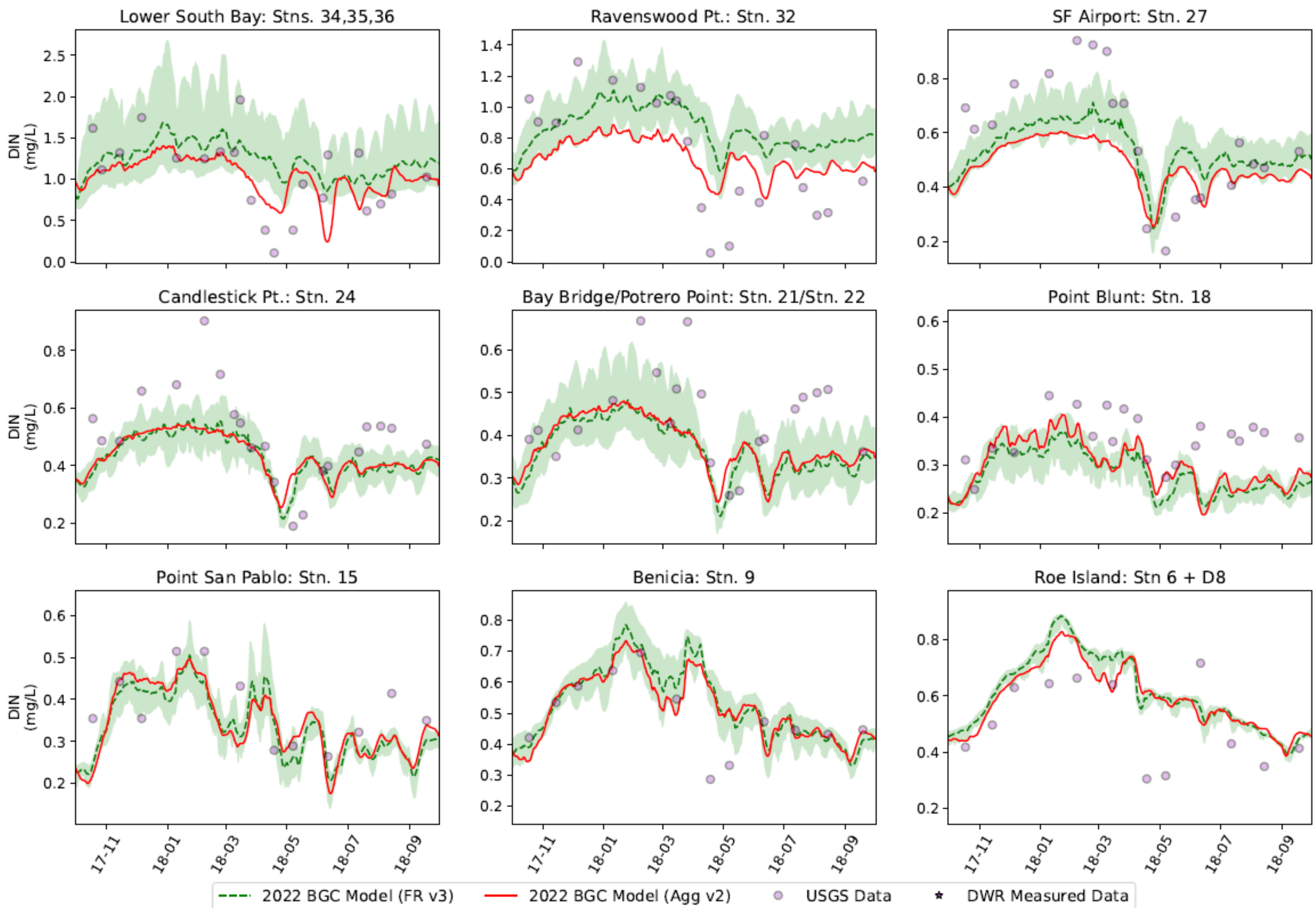
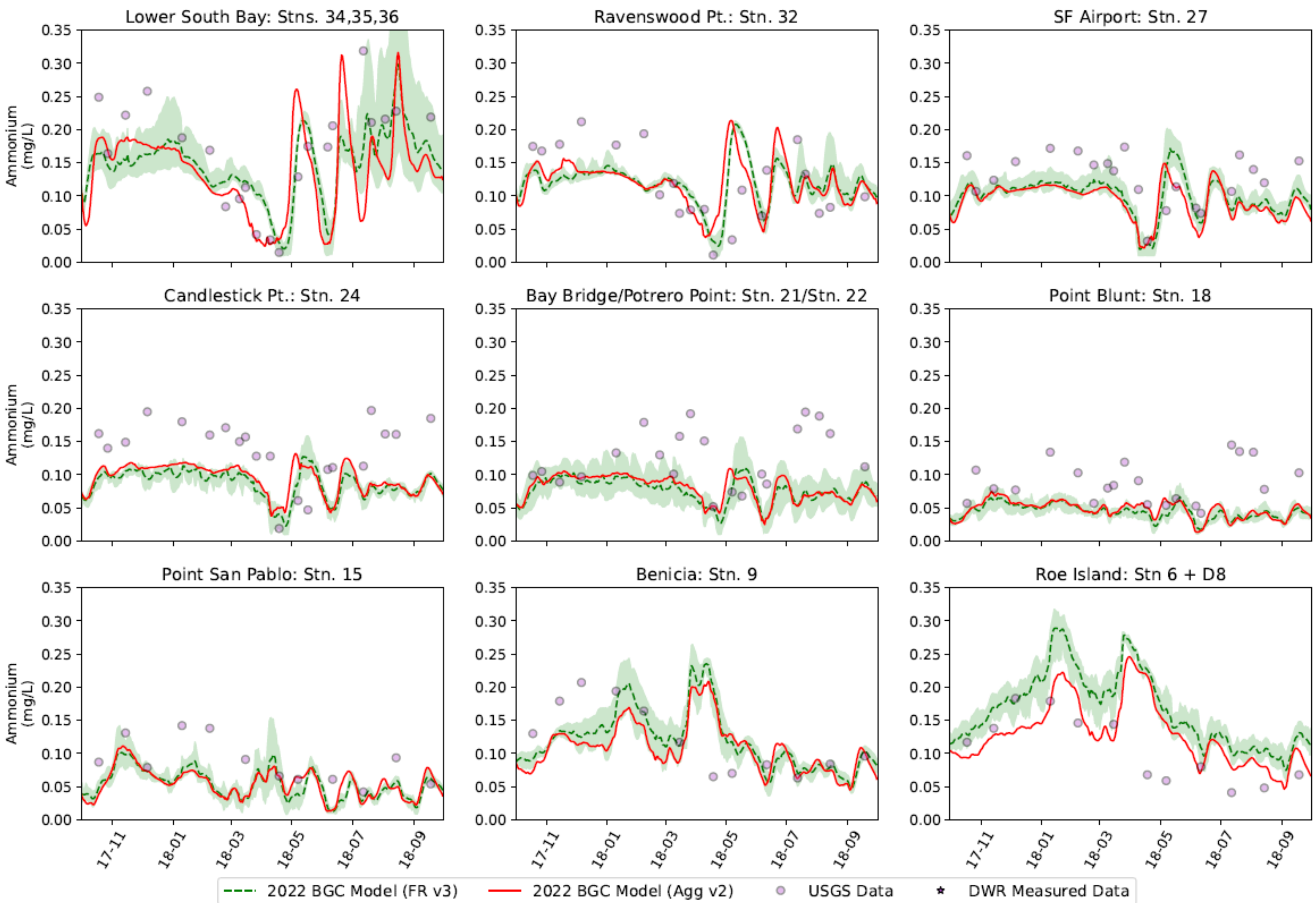


Figure 3.13. Predicted and observed DO concentrations at select USGS stations along the channel for the aggregated grid model WY2013-WY2018 simulation



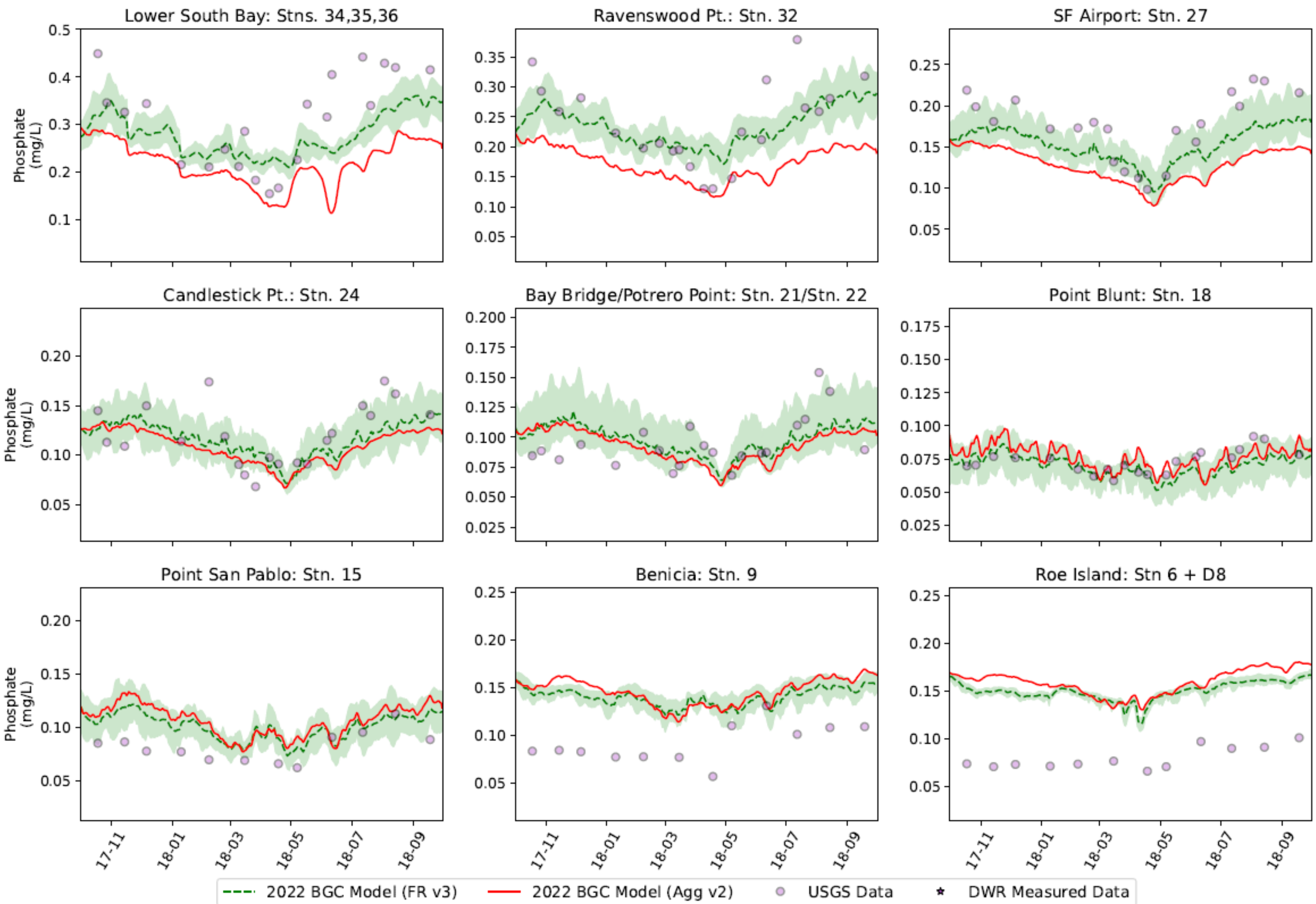
Note: The shaded area for the full resolution model simulations shows the diurnal range, and the corresponding line shows the daily average. The aggregated grid model uses diurnally averaged tides from the full resolution hydrodynamic model.

Figure 3.14. Predicted and observed DIN concentrations at select USGS stations along the channel for the full-resolution WY2018 simulation



Note: The shaded area for the full resolution model simulations shows the diurnal range, and the corresponding line shows the daily average. The aggregated grid model uses diurnally averaged tides from the full resolution hydrodynamic model.

Figure 3.15. Predicted and observed ammonium concentrations at select USGS stations along the channel for the full-resolution WY2018 simulation

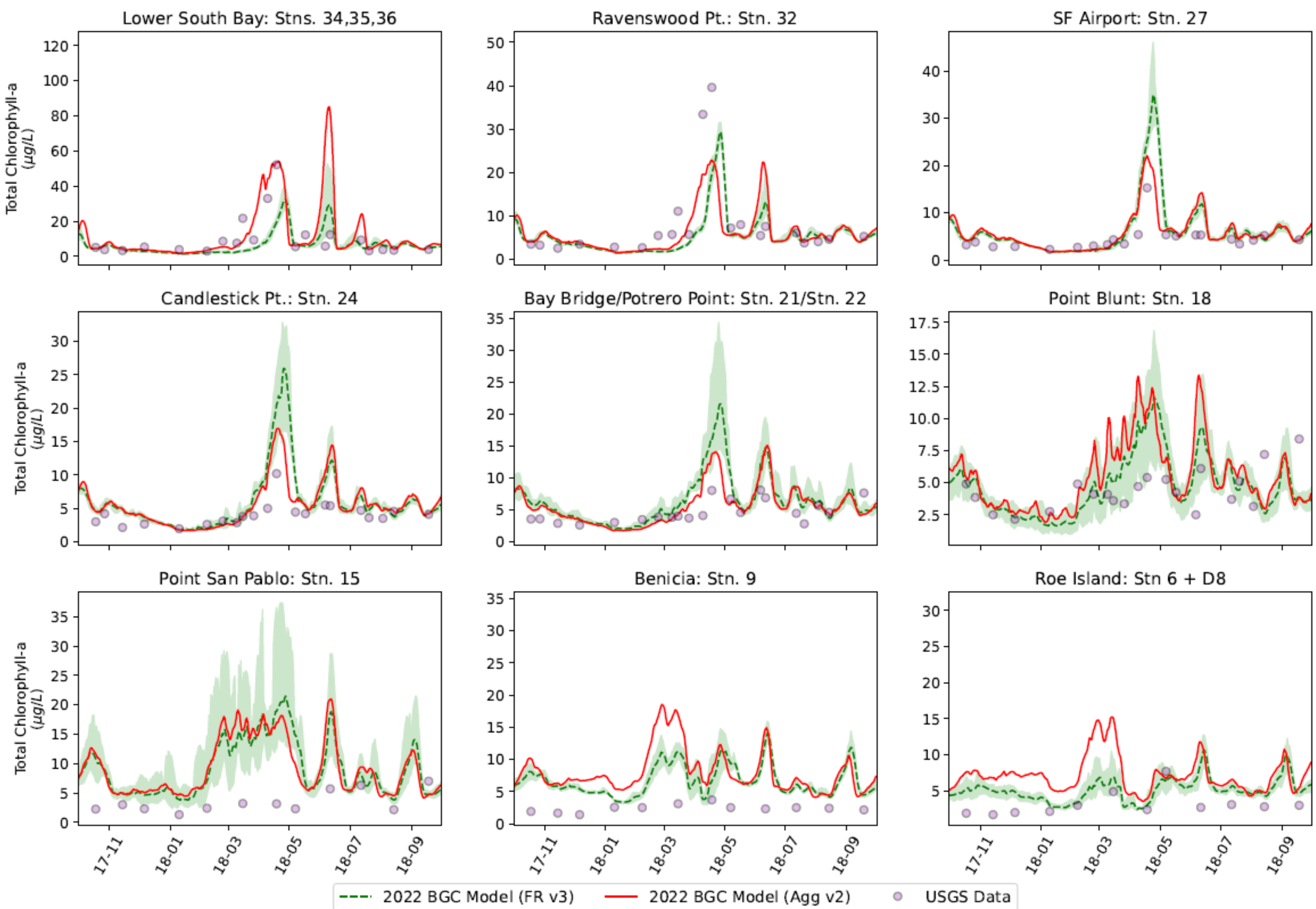


Note: The shaded area for the full resolution model simulations shows the diurnal range, and the corresponding line shows the daily average. The aggregated grid model uses diurnally averaged tides from the full resolution hydrodynamic model.

Figure 3.16. Predicted and observed dissolved orthophosphate concentrations at select USGS stations along the channel for the full-resolution WY2018 simulation

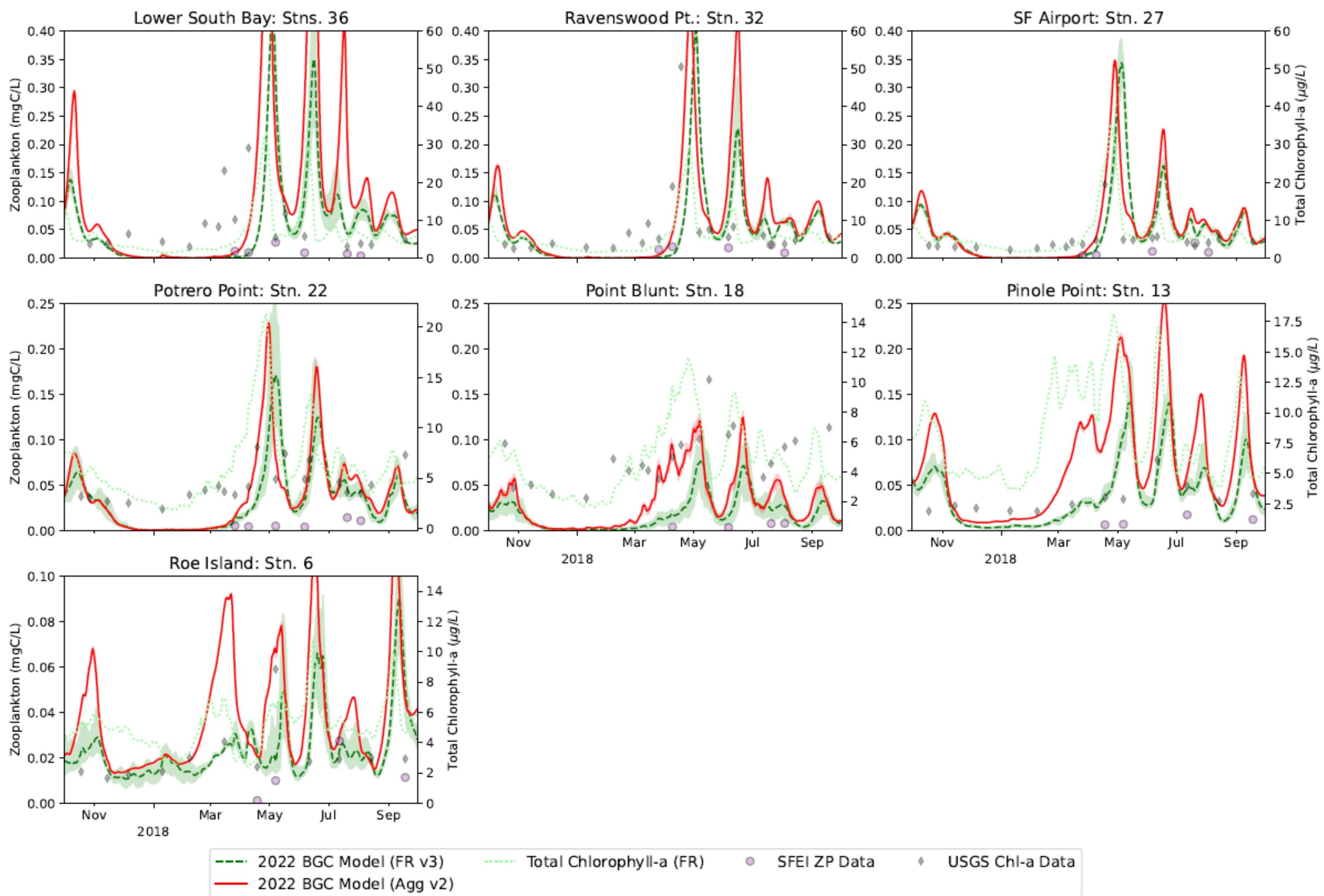
Note: The shaded area for the full resolution model simulations shows the diurnal range, and the corresponding line shows the daily average. The aggregated grid model uses diurnally averaged tides from the full resolution hydrodynamic model.

Figure 3.17. Predicted and observed dissolved silicate (as silicon) concentrations at select USGS stations along the channel for the full-resolution WY2018 simulation



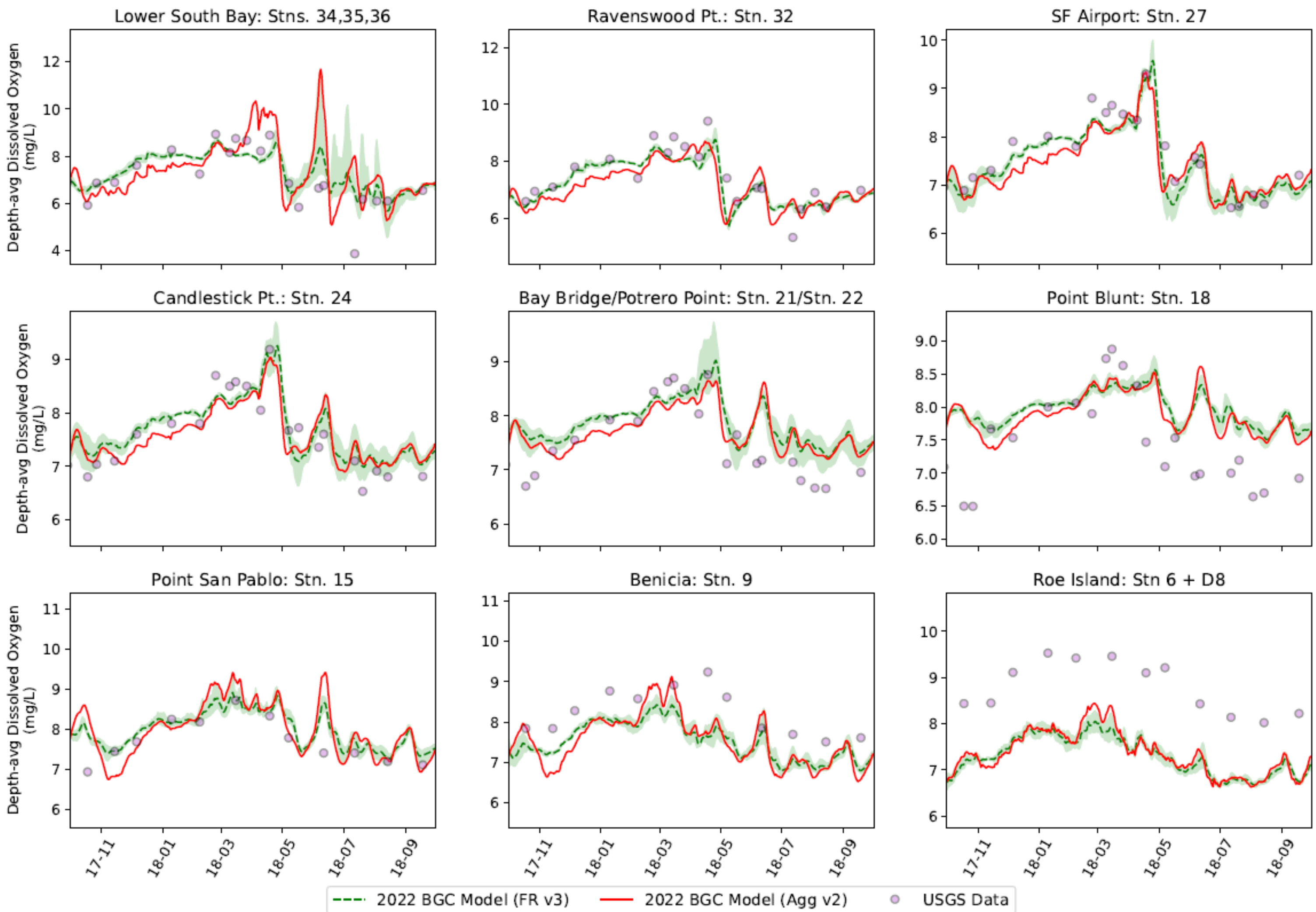
Note: The shaded area for the full resolution model simulations shows the diurnal range, and the corresponding line shows the daily average. The aggregated grid model uses diurnally averaged tides from the full resolution hydrodynamic model.

Figure 3.18. Predicted and observed chlorophyll-a concentrations at select USGS stations along the channel for the full-resolution WY2018 simulation



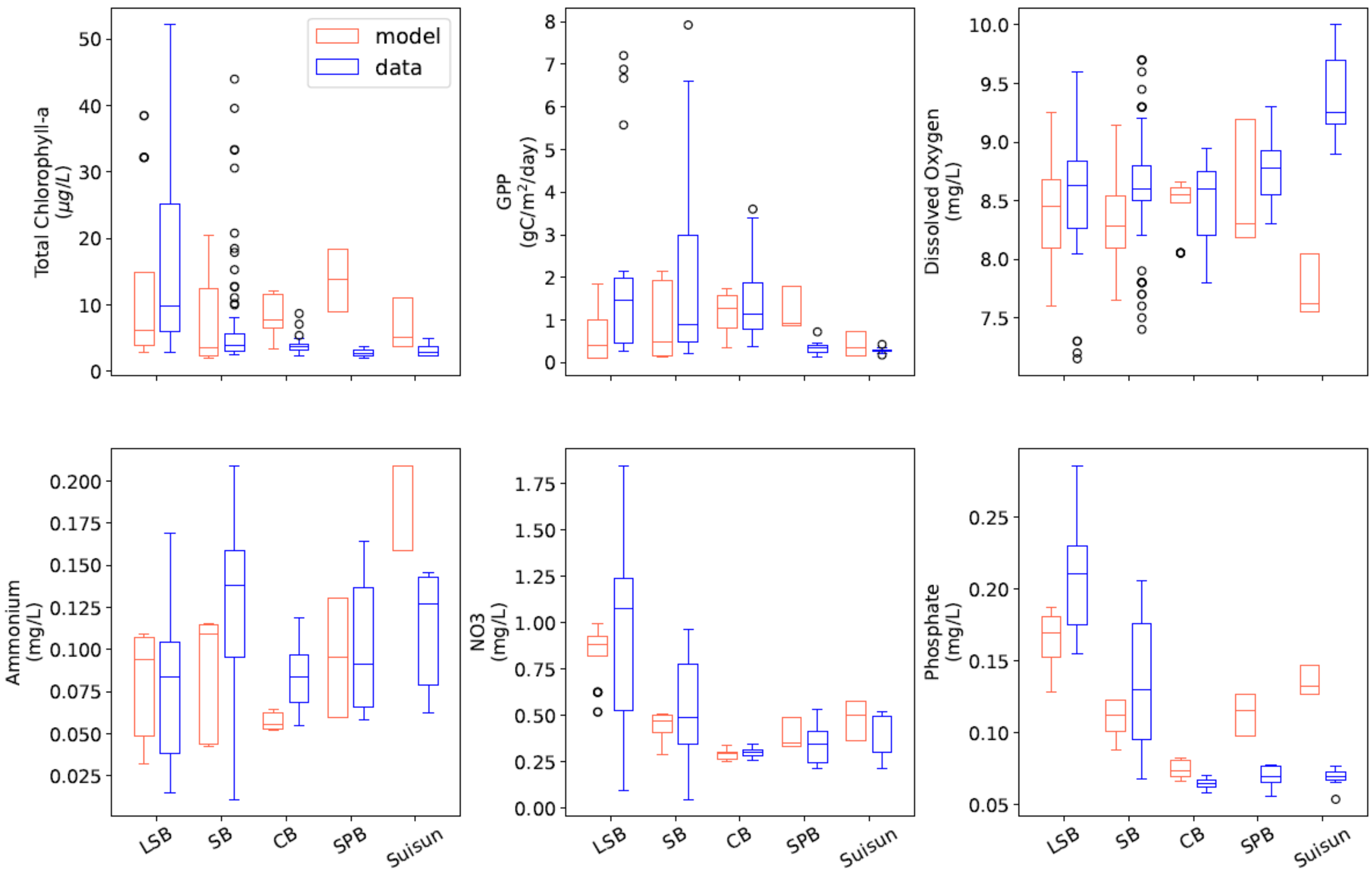
Note: The shaded area for the full resolution model simulation shows the diurnal range, and the corresponding line shows the daily average. The aggregated grid model uses diurnally averaged tides from the full resolution hydrodynamic model. The stations shown here reflect locations where zooplankton data were collected and are slightly different from other WY2018 figures.

Figure 3.19. Predicted and observed zooplankton concentrations at select USGS stations along the channel for the full-resolution WY2018 simulation



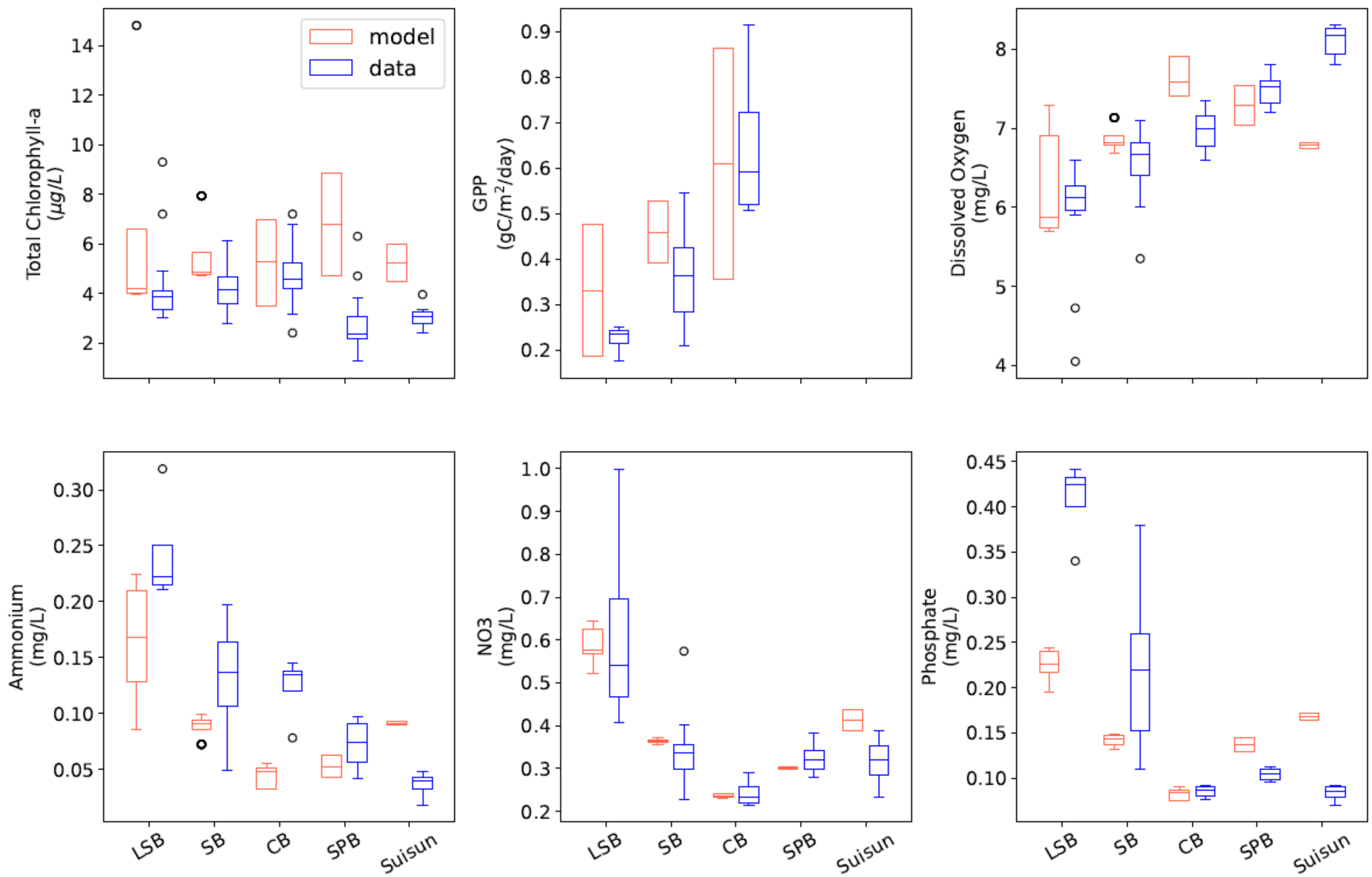
Note: The shaded area for the full resolution model simulations shows the diurnal range, and the corresponding line shows the daily average. The aggregated grid model uses diurnally averaged tides from the full resolution hydrodynamic model.

Figure 3.20. Predicted and observed depth-averaged dissolved oxygen concentrations at select USGS stations along the channel for the full-resolution WY2018 simulation



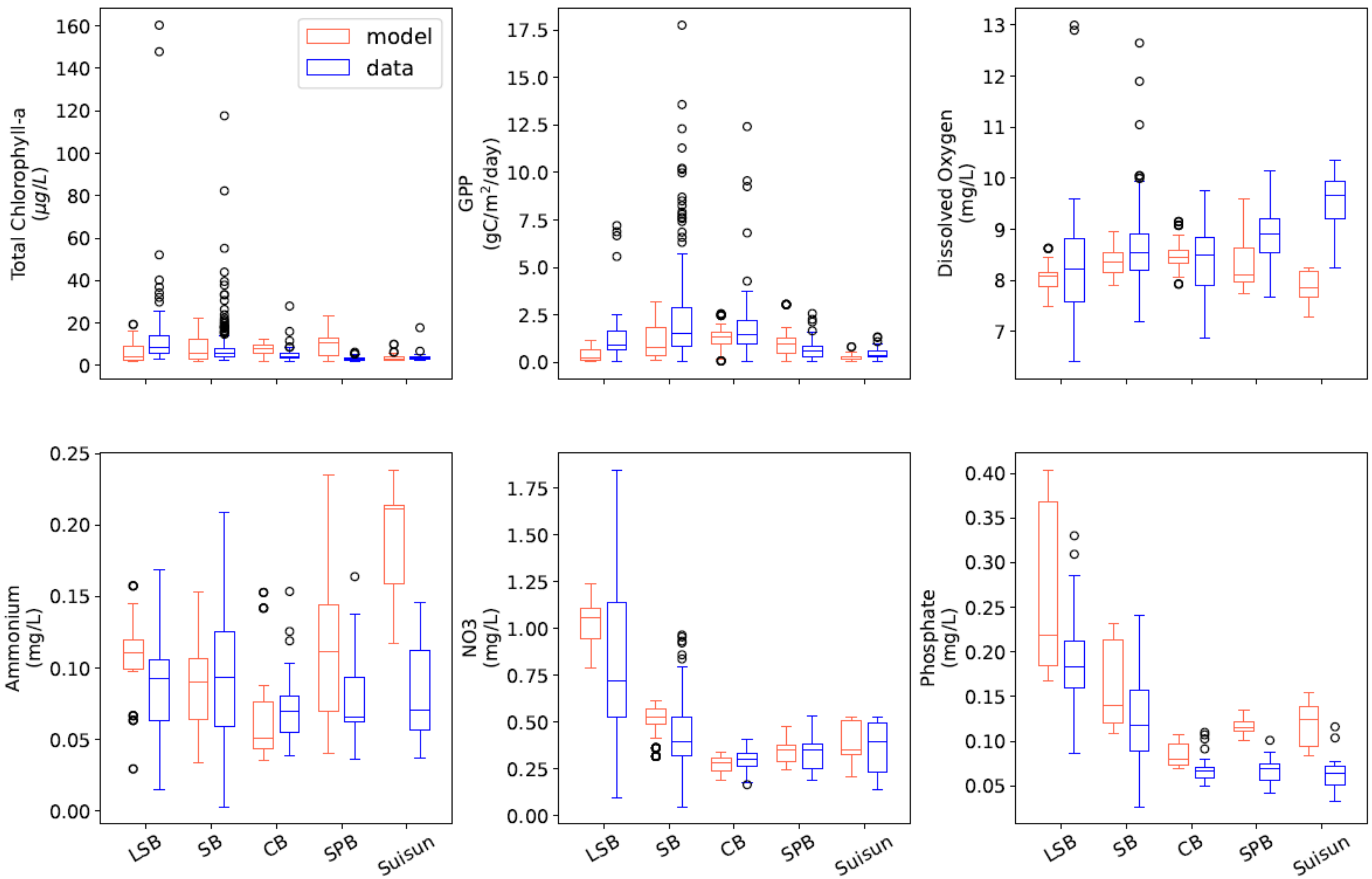
Note: Model outputs are paired with data based on the date/time of measurement over all Peterson Cruise locations within the corresponding subembayment. For LSB a high outlier of 120 µg/L chlorophyll-a observed in Feb 2013 is not shown to preserve the presentation scale for the other locations. GPP estimates are based on the equation from Cloern et al. (2007) and use the measured chlorophyll-a and light extinction coefficients from the Peterson Cruises, and solar insolation used in the model inputs based on meteorological data from CIMIS Union City station.

Figure 3.21. Box and whisker plots of aggregated grid model outputs and data for spring (Feb – Apr) for WY2013-WY2018



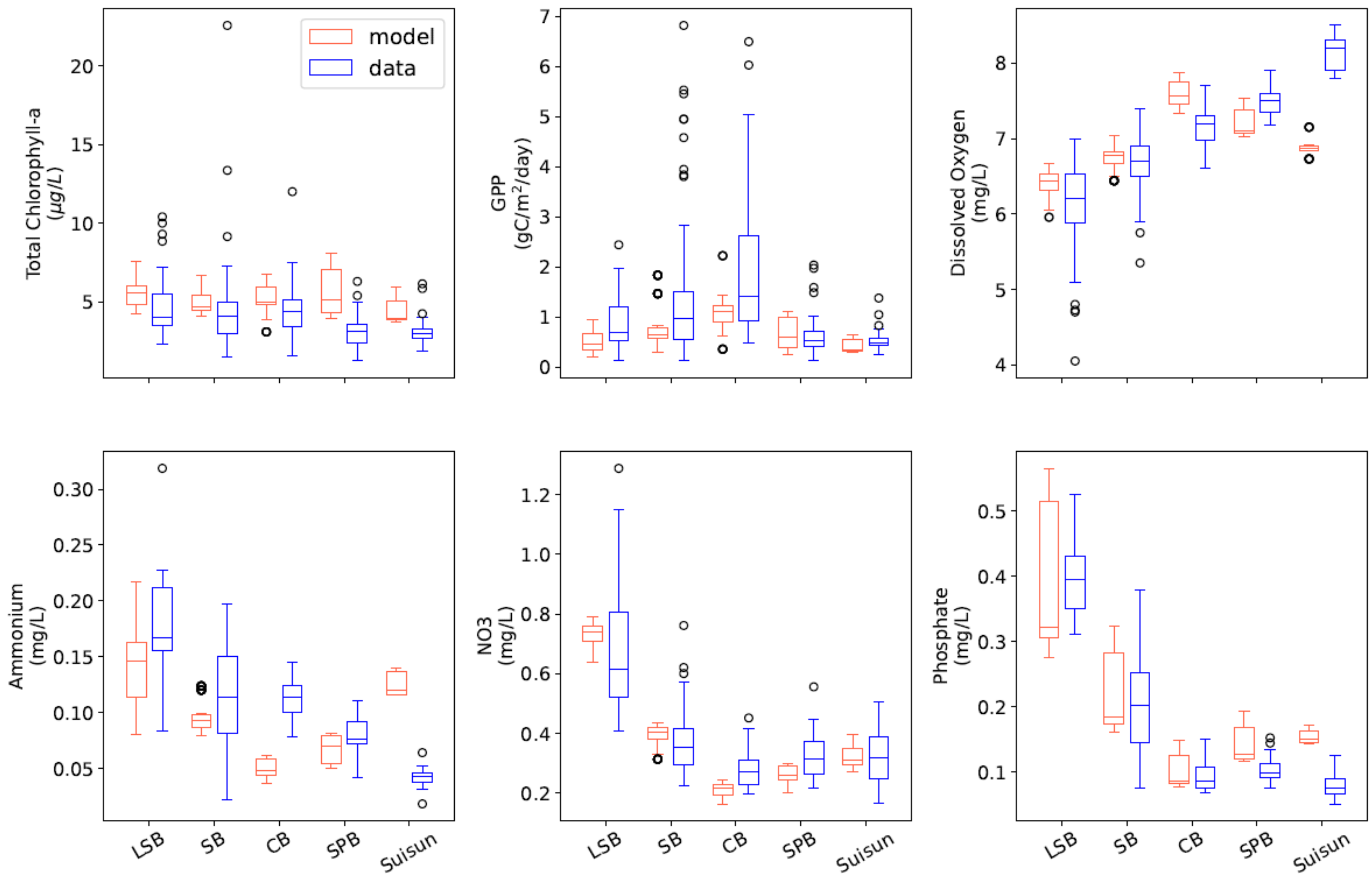
Note: Model outputs are paired with data based on the date/time of measurement over all Peterson Cruise locations within the corresponding subembayment. GPP estimates are based on the equation from Cloern et al. (2007) and use the measured chlorophyll-a and light extinction coefficients from the Peterson Cruises, and solar insolation used in the model inputs based on meteorological data from CIMIS Union City station.

Figure 3.22. Box and whisker plots of aggregated grid model outputs and data for summer (Jul-Aug) for WY2013-WY2018



Note: Model outputs are paired with data based on the date/time of measurement over all Peterson Cruise locations within the corresponding subembayment. For LSB a high outlier of 120 µg/L chlorophyll-a observed in Feb 2013 is not shown to preserve the presentation scale for the other locations. GPP estimates are based on the equation from Cloern et al. (2007) and use the measured chlorophyll-a and light extinction coefficients from the Peterson Cruises, and solar insolation used in the model inputs based on meteorological data from CIMIS Union City station.

Figure 3.23. Box and whisker plots of full resolution model outputs and data for spring (Feb – Apr) for WY2013, WY2017, and WY2018



Note: Model outputs are paired with data based on the date/time of measurement over all Peterson Cruise locations within the corresponding subembayment. Gross Primary Production (GPP) estimates are based on the equation from Cloern et al. (2007) and uses the measured chlorophyll-a and light extinction coefficients from the Peterson Cruises, and solar insolation used in the model inputs based on meteorological data from CIMIS Union City station.

Figure 3.24. Box and whisker plots of full resolution model outputs and data for summer (Jul-Aug) for WY2013, WY2017, and WY2018

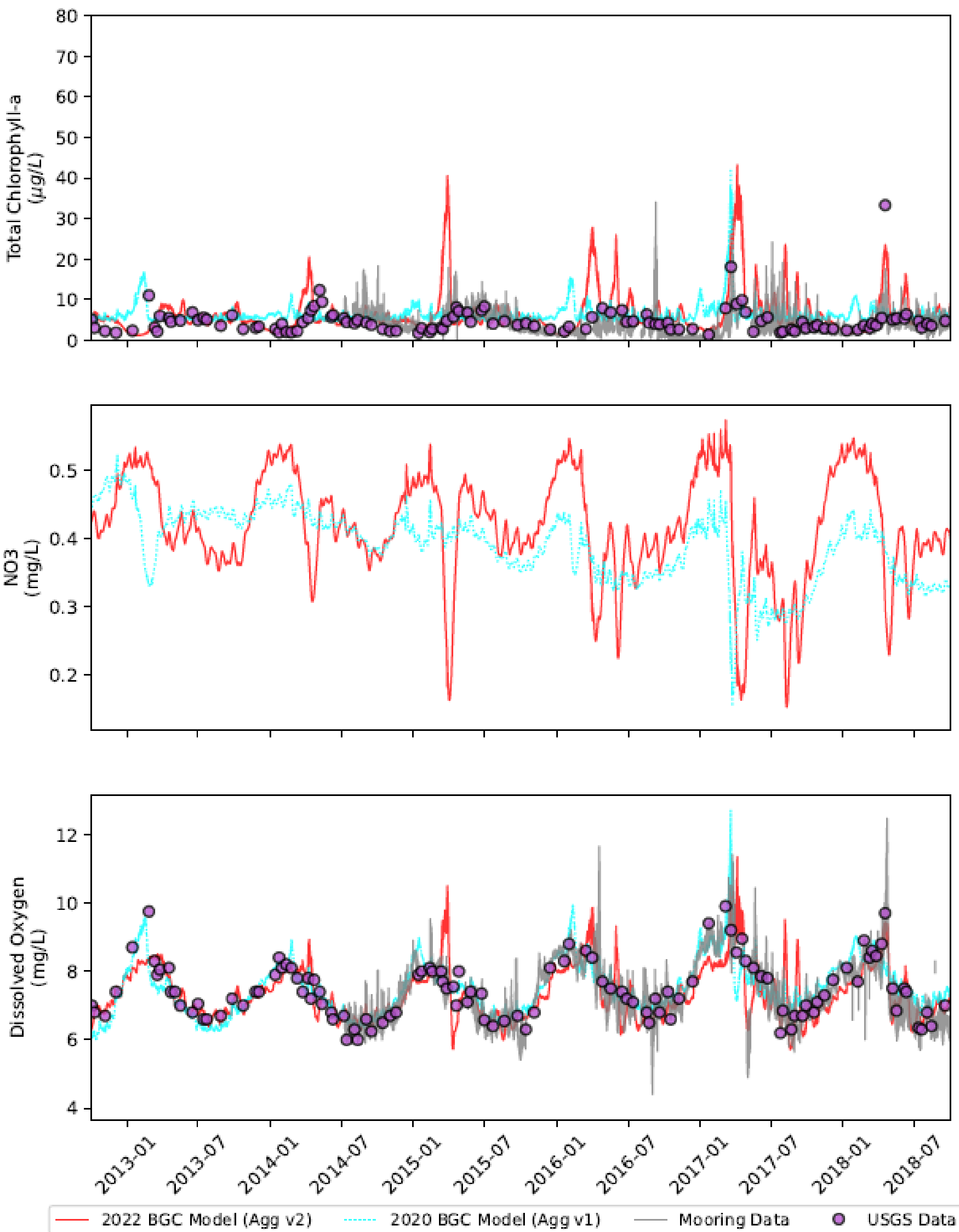


Figure 3.25. Comparison of aggregated grid model outputs with high-frequency data at San Mateo Bridge

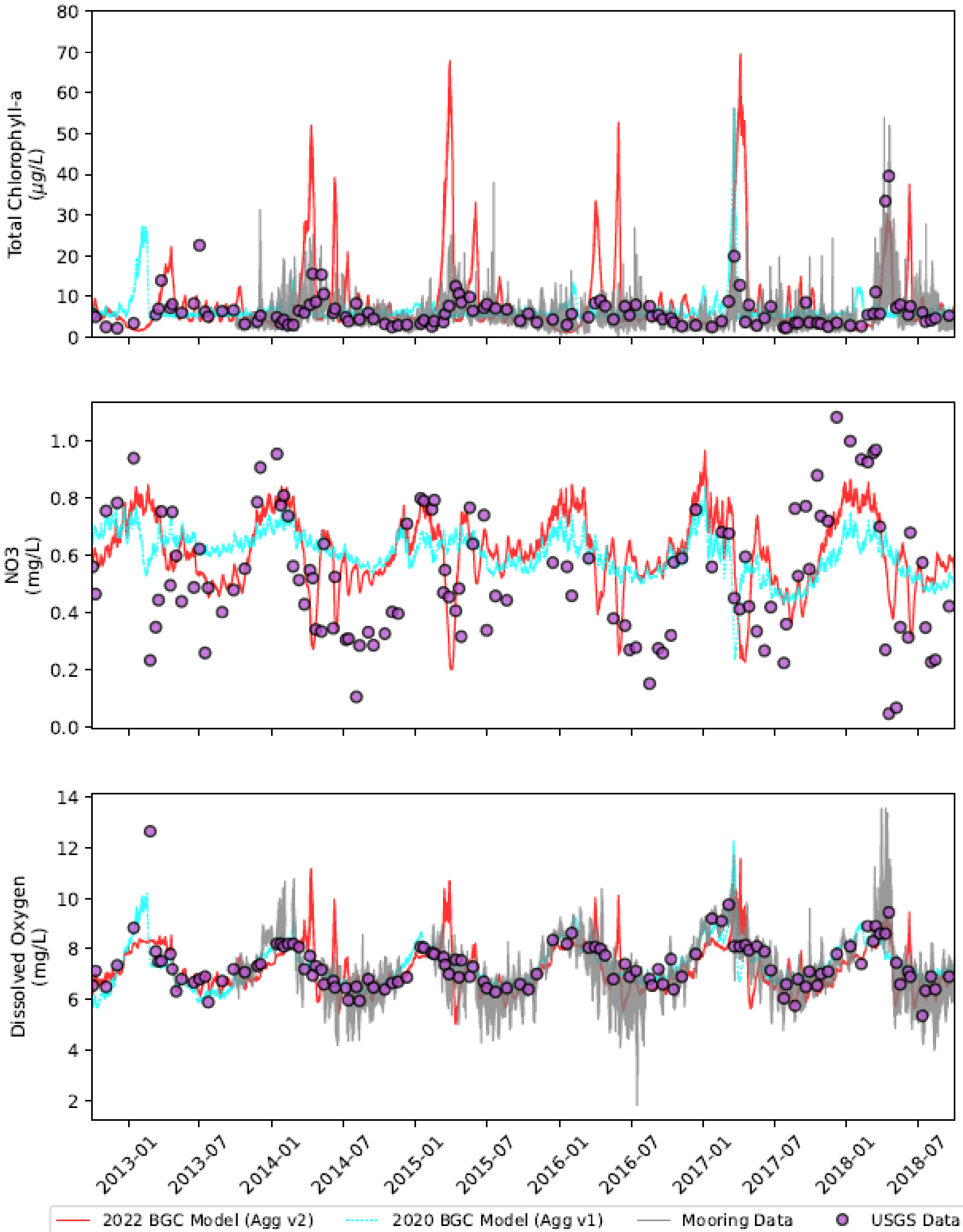


Figure 3.26. Comparison of aggregated grid model outputs with high-frequency data at Dumbarton Bridge

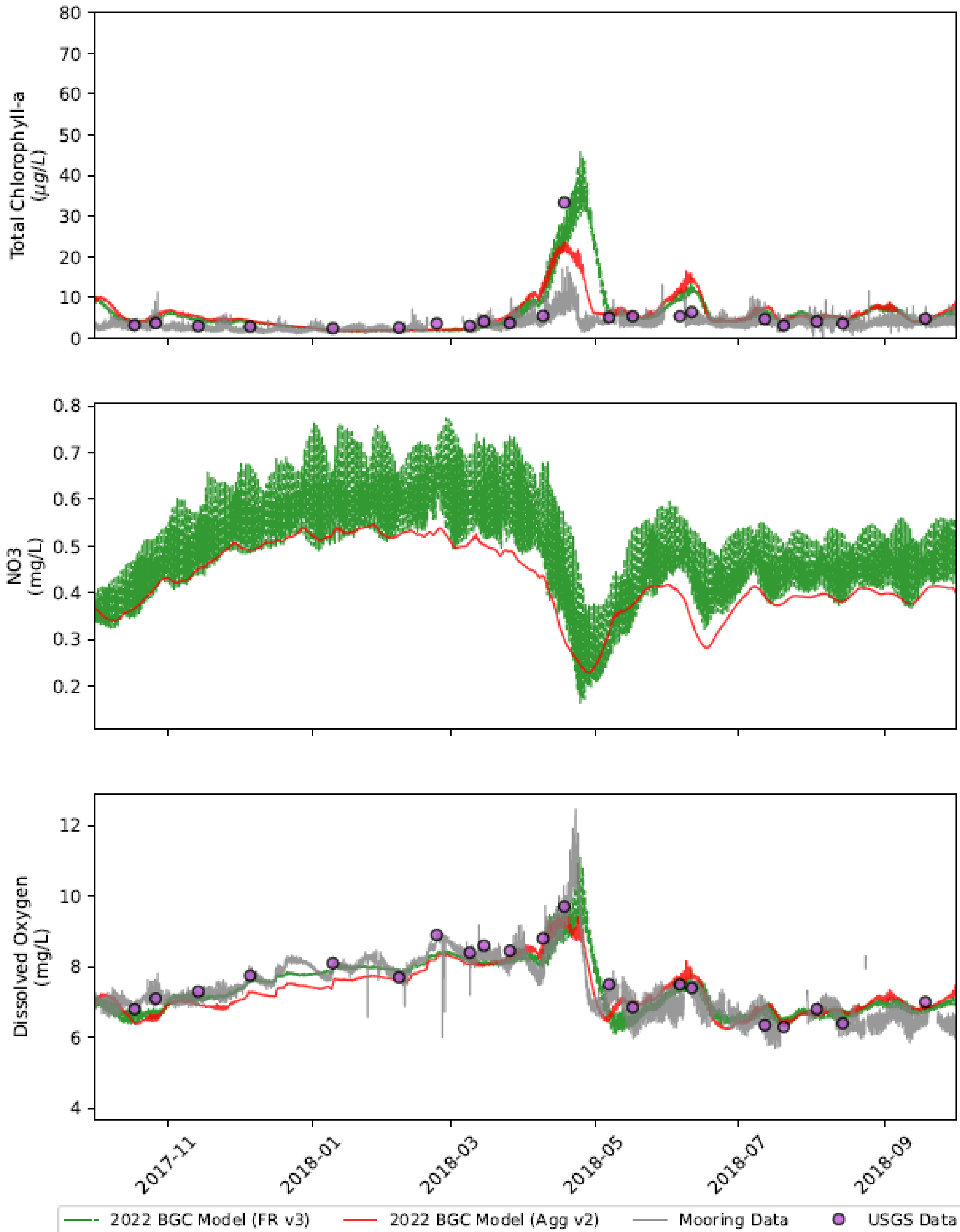


Figure 3.27. Comparison of full resolution model outputs with high-frequency data at San Mateo Bridge for WY2018

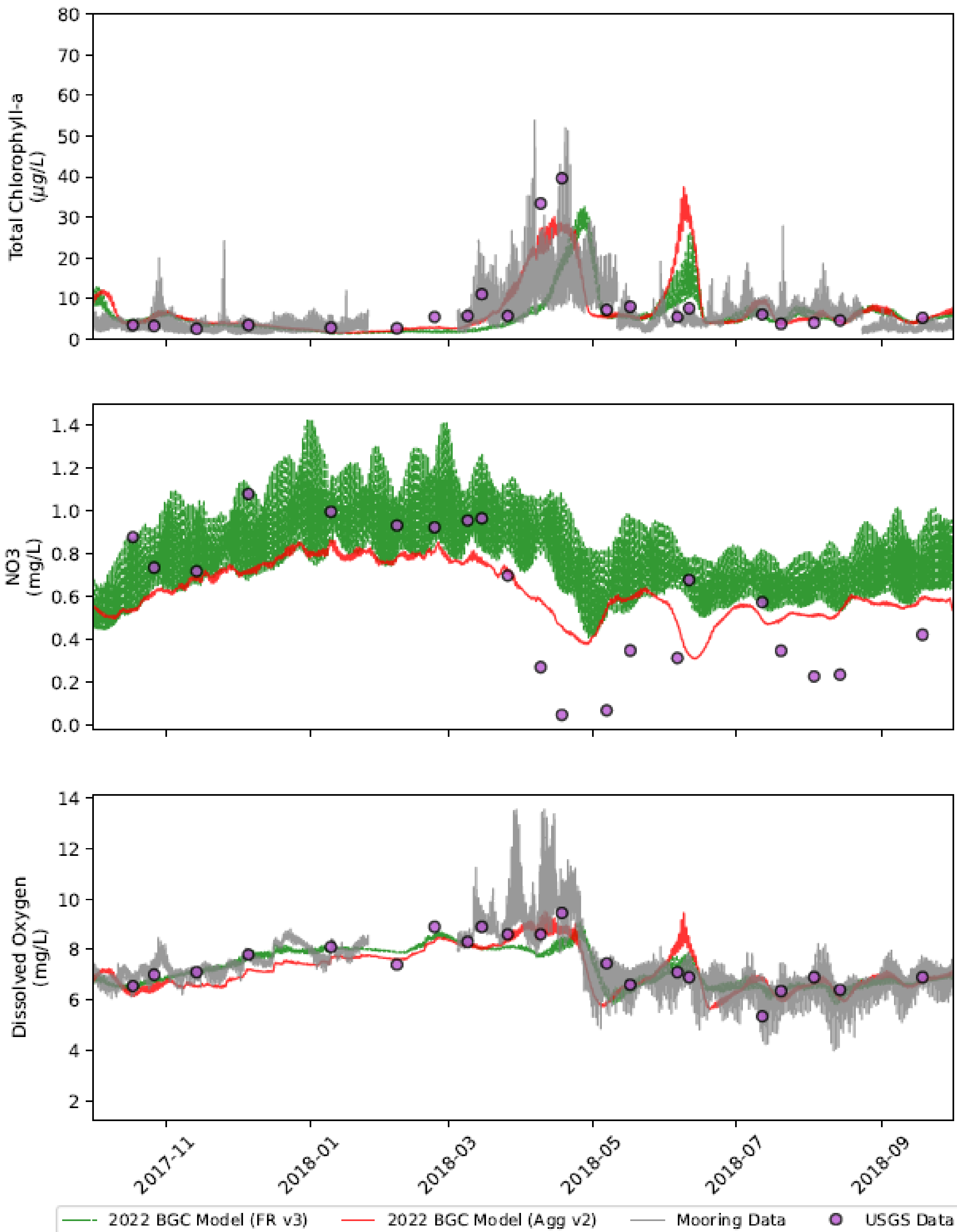


Figure 3.28. Comparison of full resolution model outputs with high-frequency data at Dumbarton Bridge for WY2018

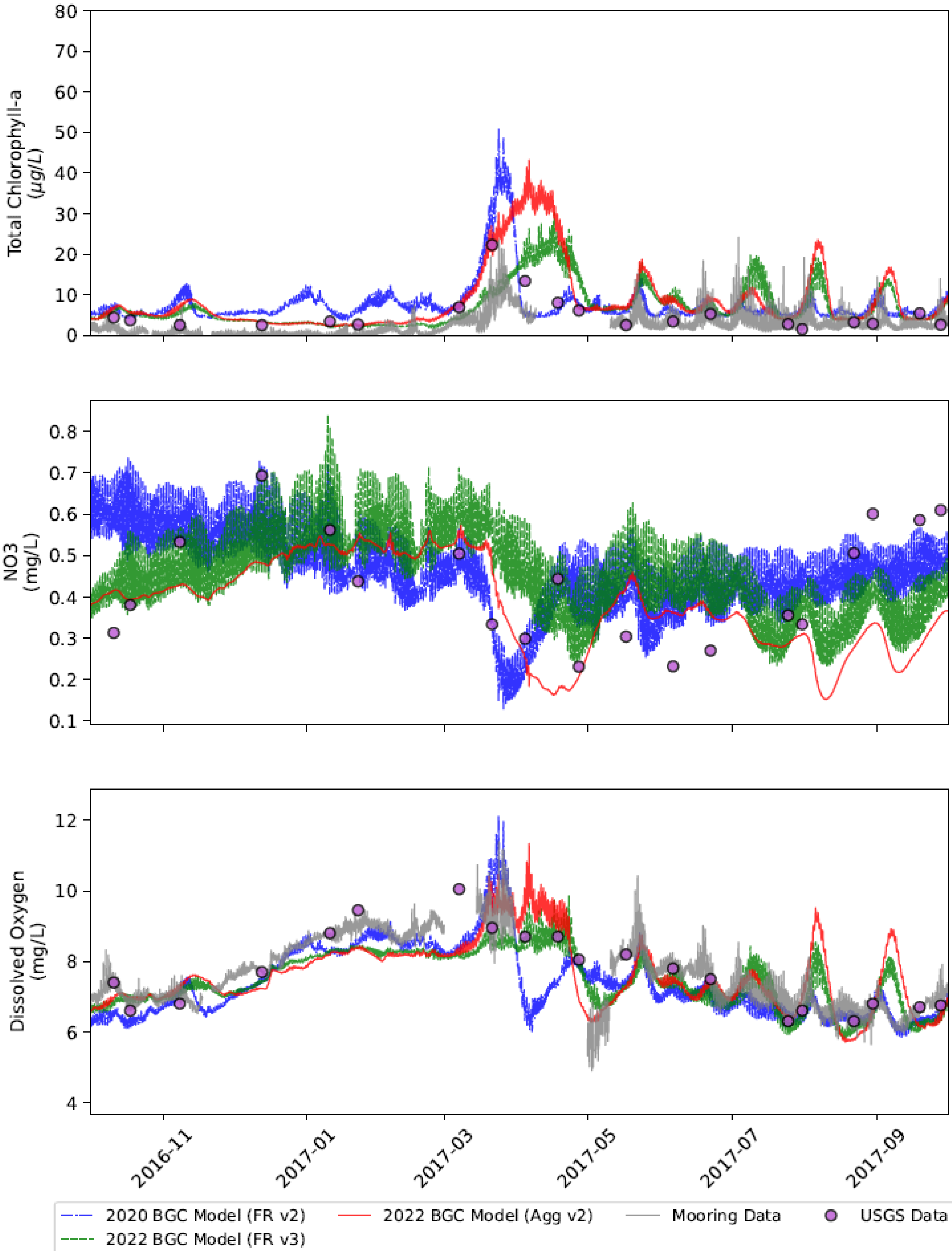


Figure 3.29. Comparison of full resolution model outputs with high-frequency data at San Mateo Bridge for WY2017

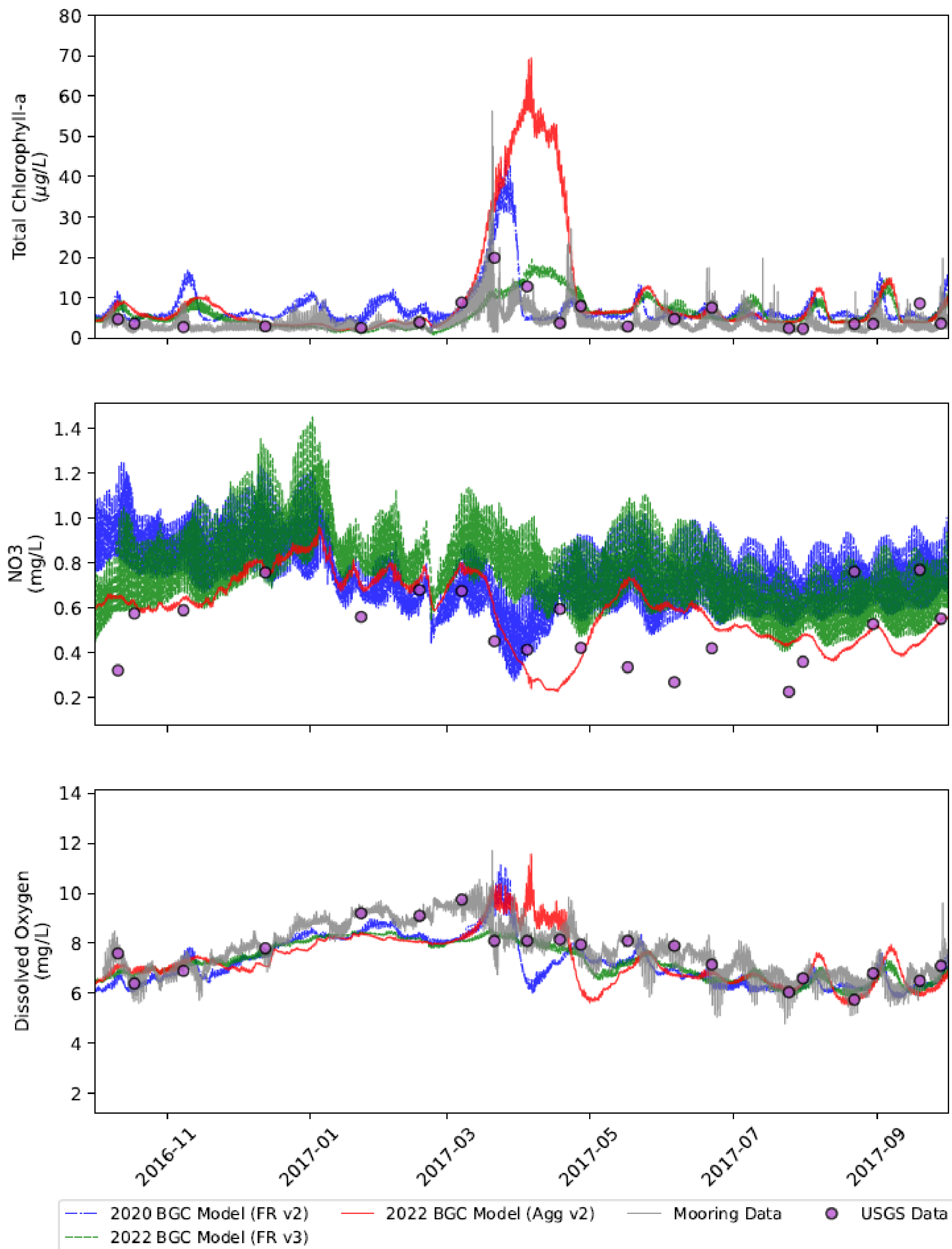


Figure 3.30. Comparison of full resolution model outputs with high-frequency data at Dumbarton Bridge for WY2017

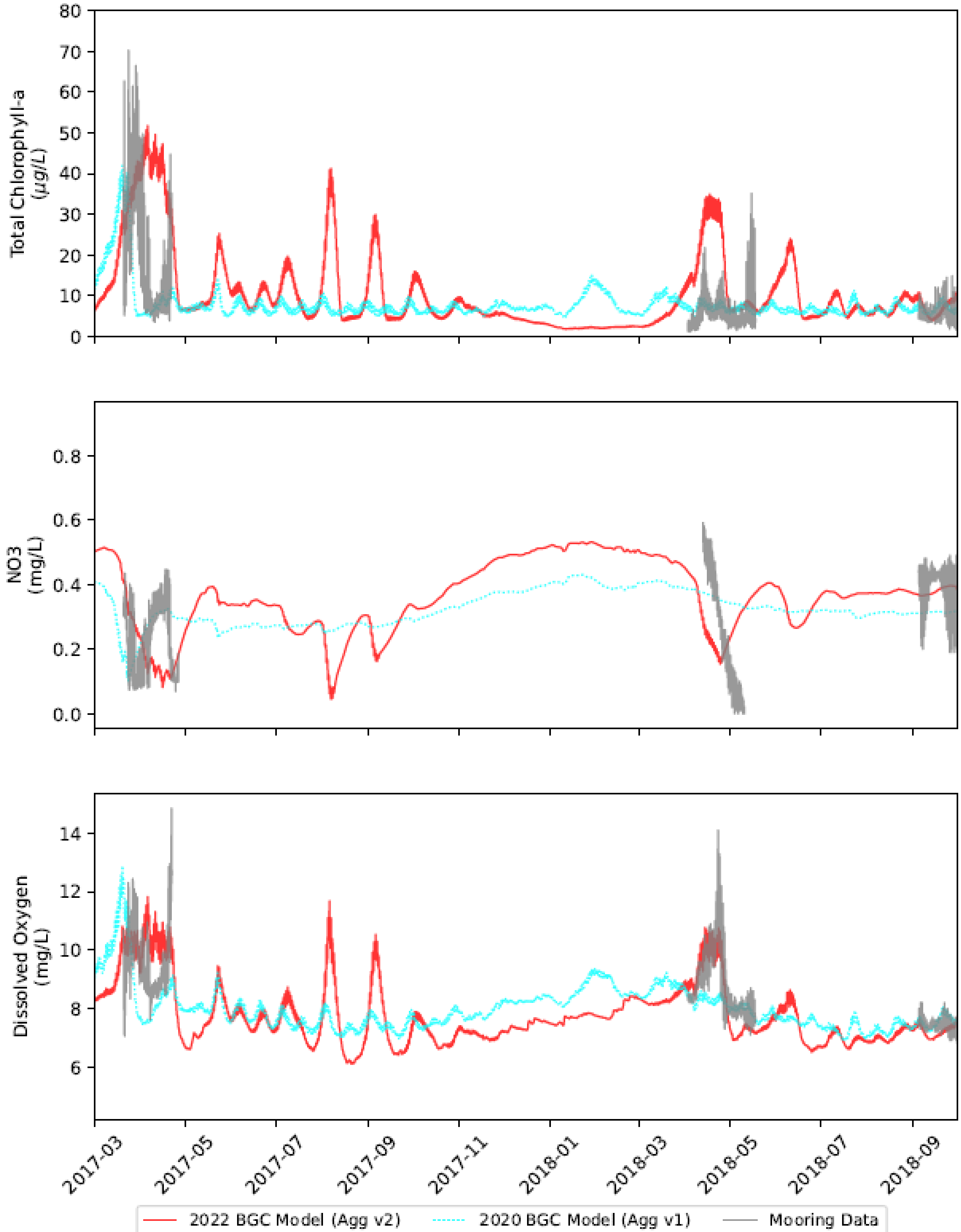


Figure 3.31. Comparison of aggregated grid model outputs with high-frequency data at SFEI Shoal Mooring

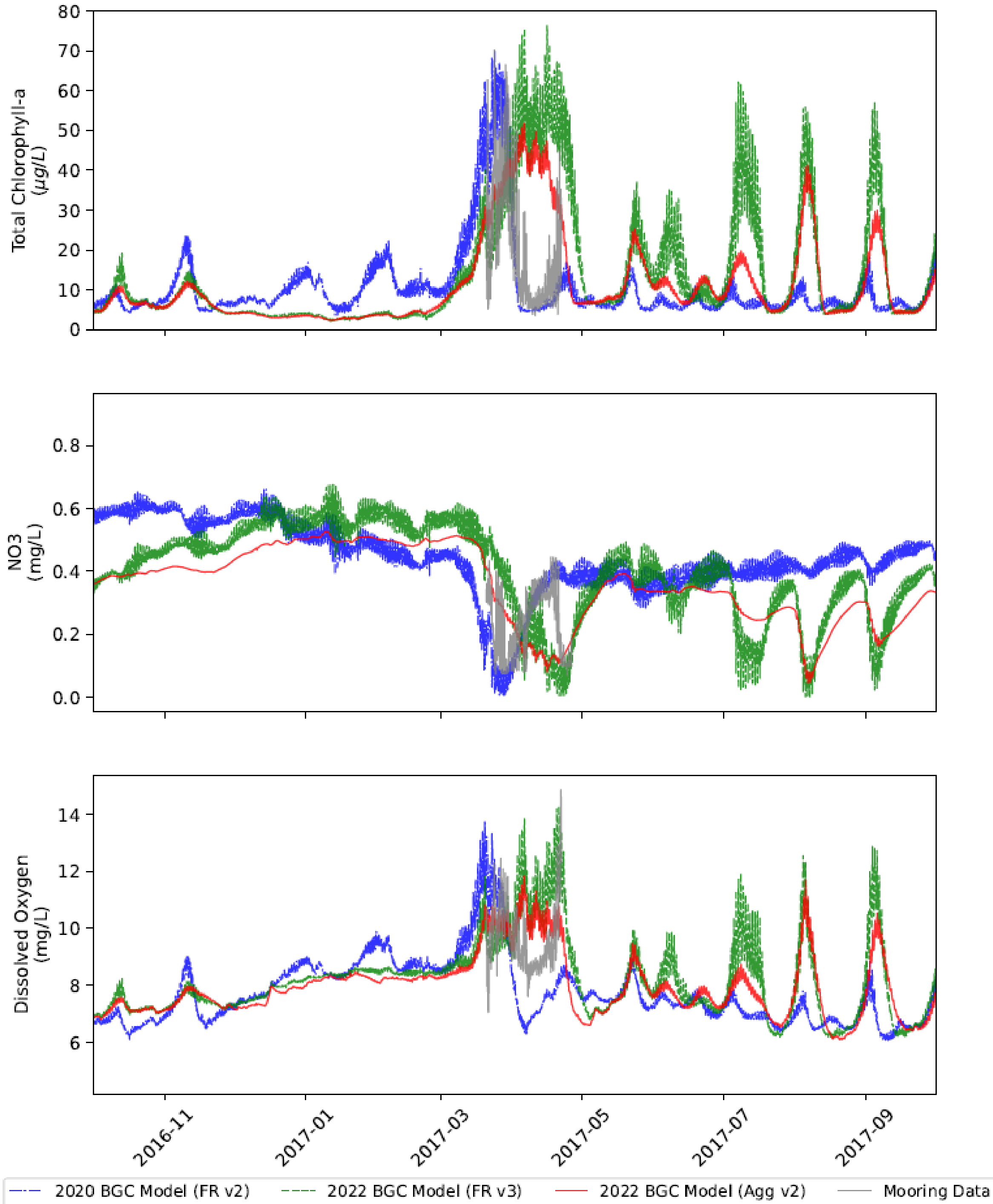


Figure 3.32. Comparison of full resolution model outputs with high-frequency data at SFEI Shoal Mooring for WY2017

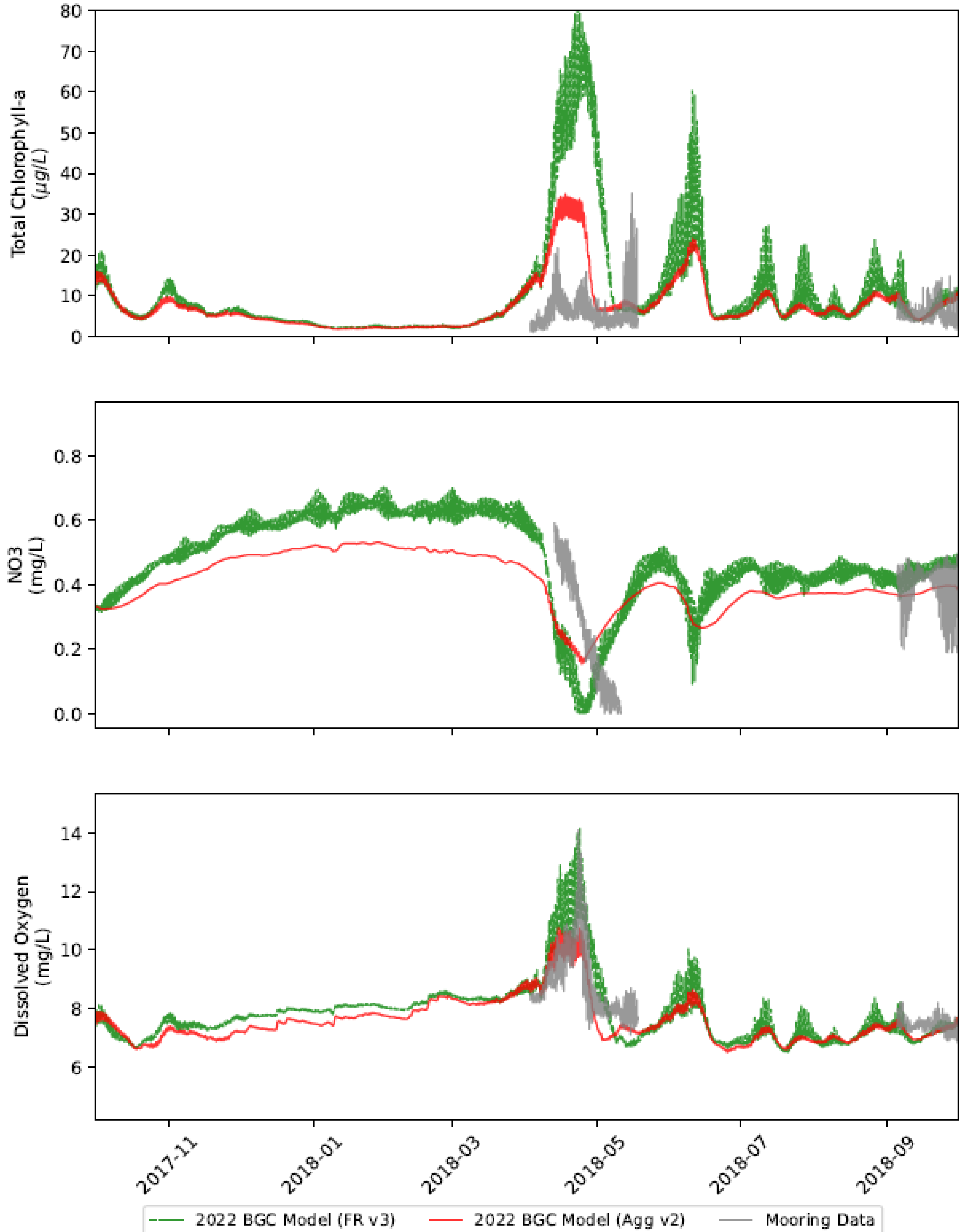


Figure 3.33. Comparison of full resolution model outputs with high-frequency data at SFEI Shoal Mooring for WY2018

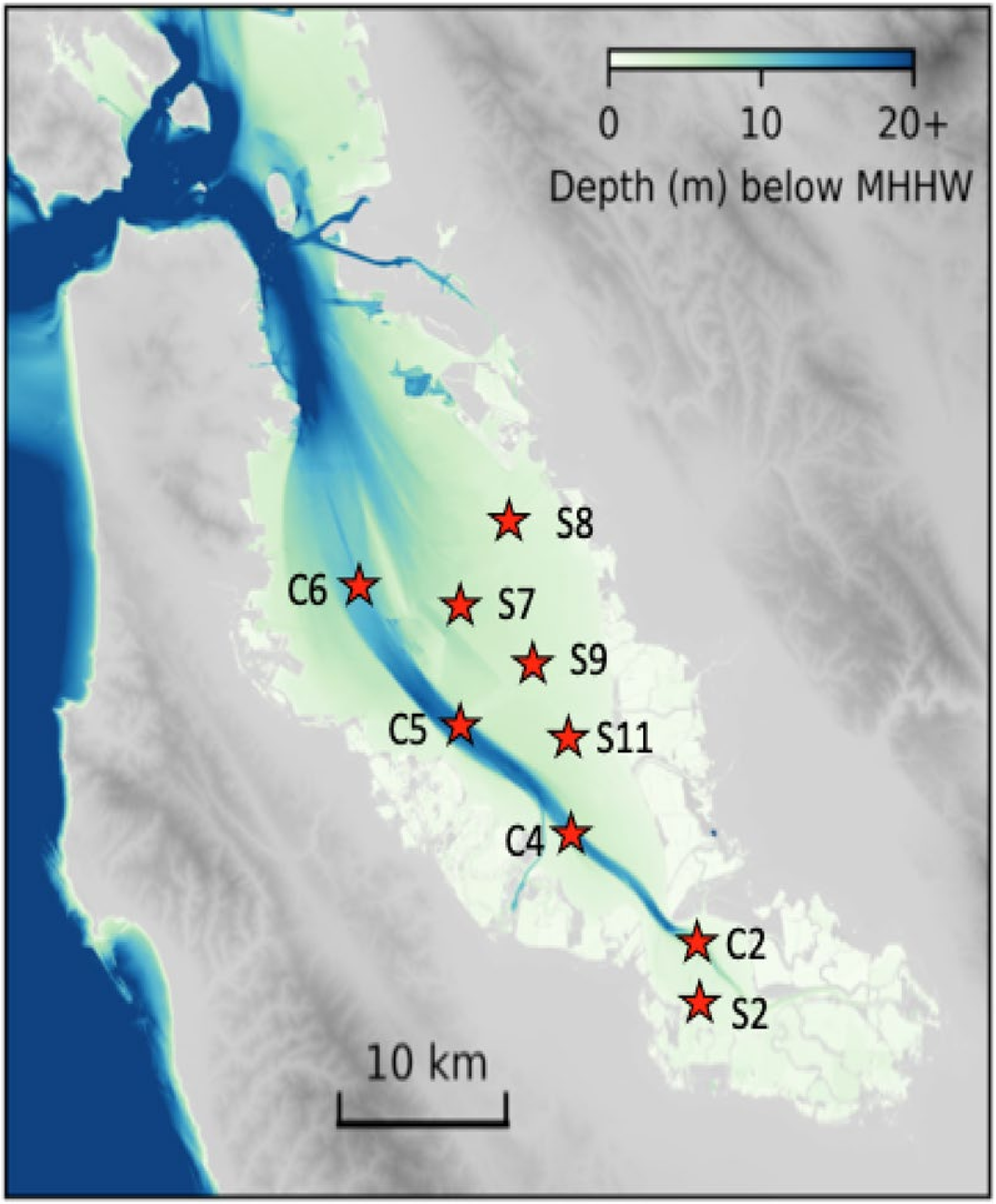
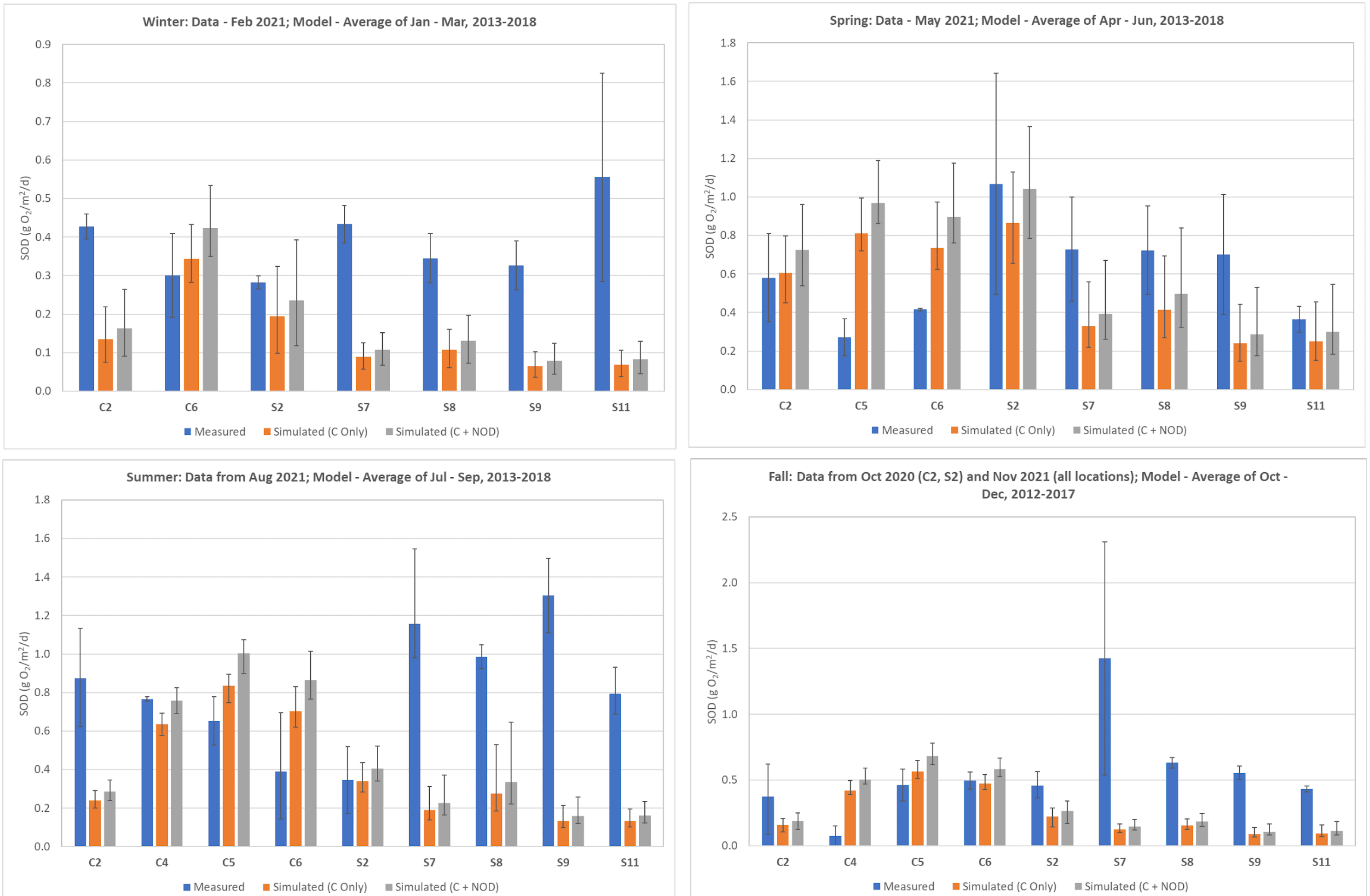


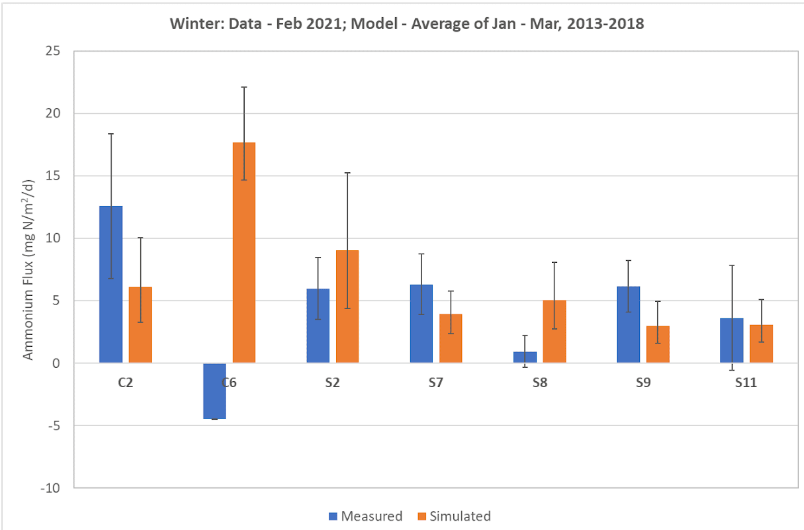
Figure 3.34. Locations of sediment cores collected in Lower South Bay and South Bay for laboratory sediment incubation studies

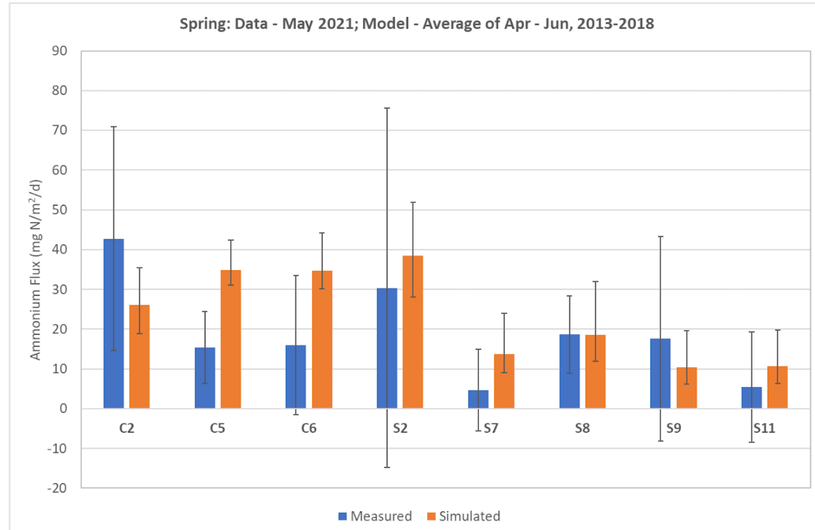


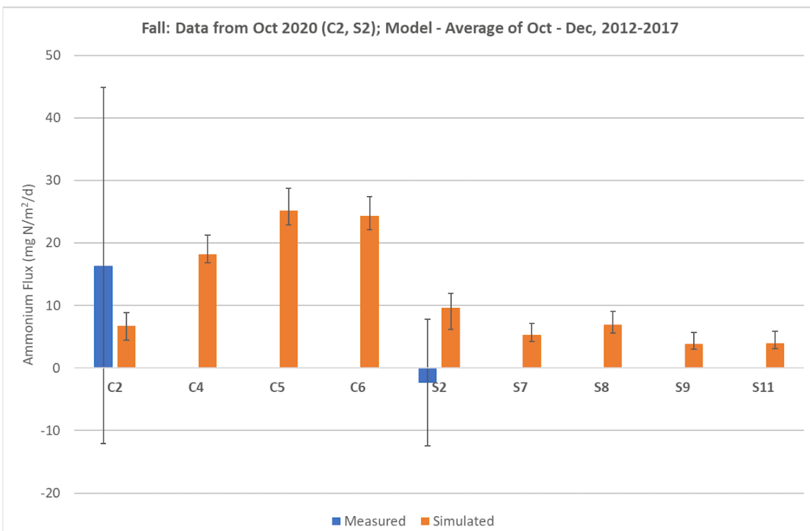
Notes:

1. SOD is shown as a positive quantity even though it represents flux of DO into sediments and, as such, is the opposite of the convention presented for other constituents.
2. Error bars show the range of model predictions over WY2013-WY2018 for the averaging period, and the range of replicates and years (where applicable) for the measurements.
3. Figure 3.34 shows the locations of sediment cores.

Figure 3.35. Model predicted and observed sediment oxygen demand at South Bay and Lower South Bay sites

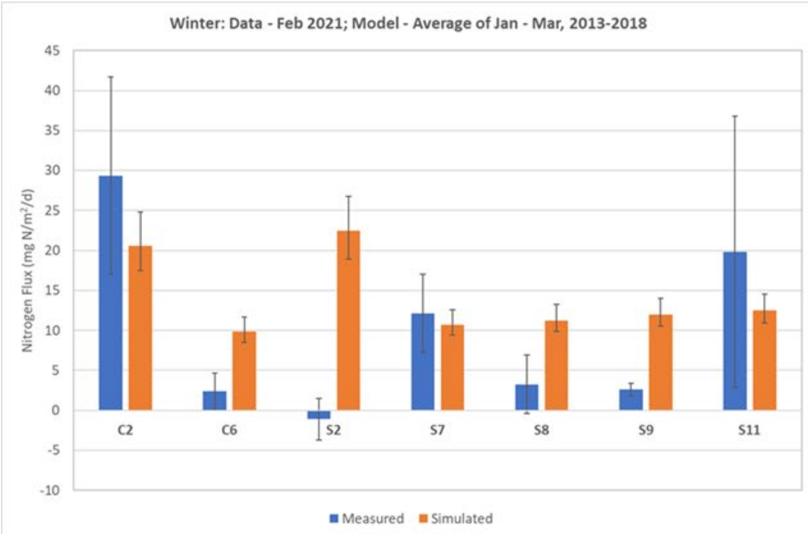


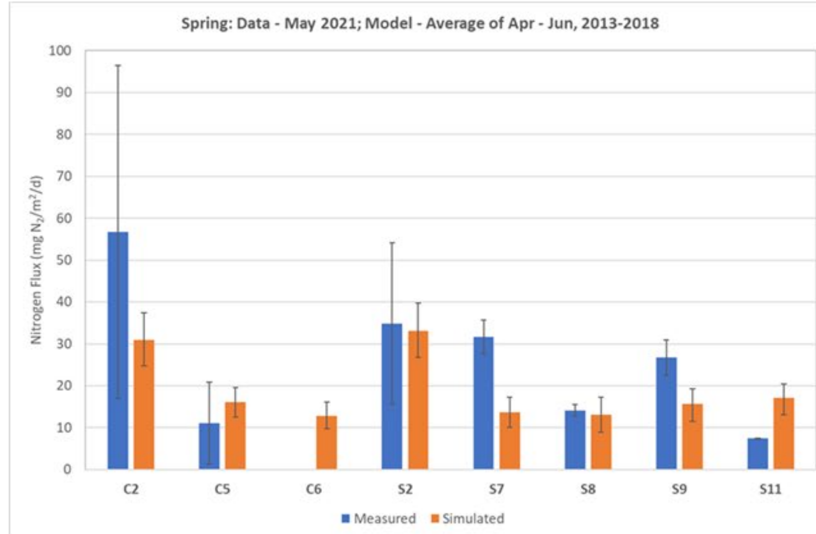


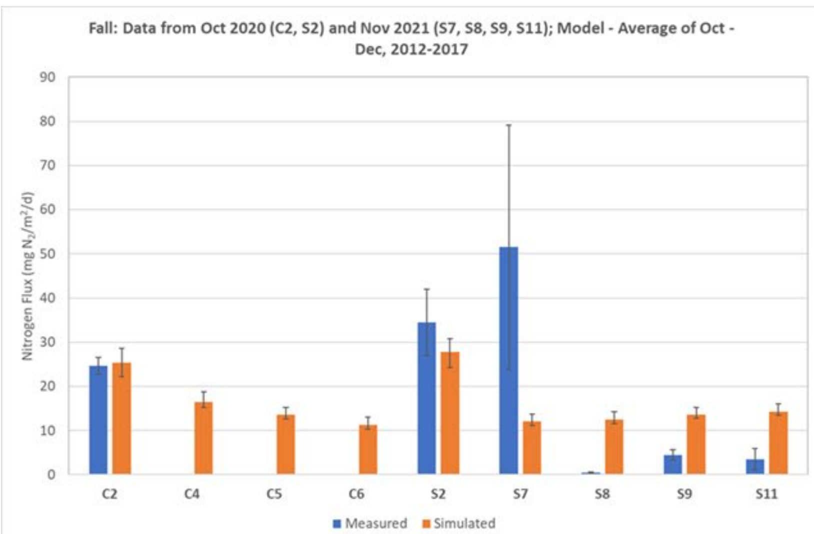


Notes:
Negative flux represents transfer into sediments, and positive flux represents transfer out of sediments.
Error bars show range of model over the averaging period, and range over replicates for data.
Summer nitrogen diagenesis flux data from August 2021 sediment cores were not available at the time this report was finalized.
Figure 3.34 shows the locations of the sediment cores.

Figure 3.36. Model predicted and observed model predicted and observed nitrogen diagenesis at South Bay and Lower South Bay Sites







Notes:

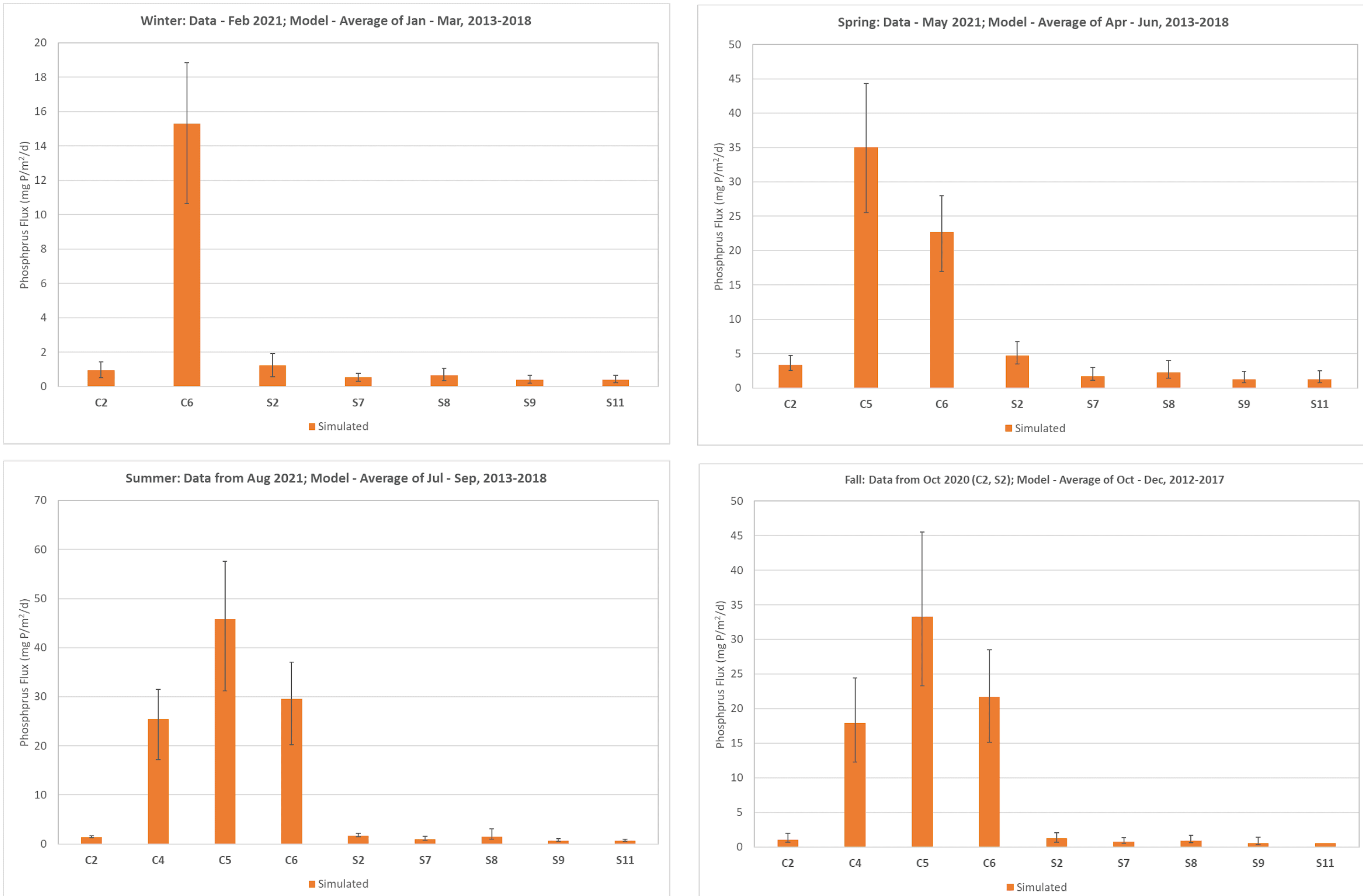
Negative flux represents transfer into sediments, and positive flux represents transfer out of sediments.

Error bars show range of model over the averaging period, and range over replicates and years (where applicable).

Summer nitrogen flux data from August 2021 sediment cores were not available at the time this report was finalized.

Figure 3.34 shows the locations of the sediment cores.

Figure 3.37. Model predicted and observed denitrification fluxes at South Bay and Lower South Bay Sites



Notes:

1. Negative flux represents transfer into sediments, and positive flux represents transfer out of sediments.
2. Error bars show range of model over the averaging period, and range over replicates and years (where applicable).
3. Figure 3.34 shows the locations of the sediment cores. However, due to potential analytical issues for laboratory phosphorus fluxes sediment incubation results have not been included in this comparison.

Figure 3.38. Model predicted phosphorus fluxes at South Bay and Lower South Bay Sites

4. Next Steps

The following model refinement and application activities are planned for the rest of FY2023 and FY2024:

- Model refinement will be continued with including improvements to light field and evaluating model performance to better simulate nutrient cycling in the fall period, and potential refinements to the ocean boundary condition. In addition, model refinements will also incorporate feedback from the Model Advisory Group.
- Simulate hydrodynamics and biogeochemistry for additional water years (WY2019 – WY2022), including recent years with more high-frequency in situ sensor data, and use those years to further refine model calibration.
- Apply model to answer NMS science and management questions on factors that affect nutrient cycling, primary production, and DO. These applications will use subembayment and sub-subembayment level control volume-based mass budgets and process-level fluxes.
- Address long-term and near-term management priorities for the effects on load reductions on the spatial/temporal changes in key water quality parameters (for example, DIN, DO, and chlorophyll-a) and the effects on coastal export of nutrients.
- Develop approaches for evaluating management relevant ways for estimating uncertainty in model predictions relative to science and management questions, and potential load reduction actions. The approach to incorporating uncertainty will reconcile a range of boundary conditions in the evaluation of management scenarios and how key decision variables (for example, DO, chlorophyll-a, primary production response) will be affected over those range of conditions.

5. References

Brady, D.C., Testa, J.M., DiToro, D.M., Boynton, W.R., and Kemp, M.W. 2013. Sediment flux modeling: Calibration and application for coastal systems, Estuarine, Coastal and Shelf Science, 117: 107-124, doi: http://dx.doi.org/10.1016/j.ecss.2012.11.003

Cerco, C.F. and Noel, M.R. 2017. The 2017 Chesapeake Bay Water Quality and Sediment Transport Model, A Report to the U.S. Environmental Protection Agency Chesapeake Bay Program, by U.S. Army Engineer Research and Development Center, Vicksburg, MS. May.

Cloern, J.E., Schraga, T.S., Nejad, E., and Martin, C. 2020. "Nutrient Status of San Francisco Bay and Its Management Implications." *Estuaries and Coasts* 43:1299-1317. doi.org/10.1007/s12237-020-00737-w.

Cloern, J.E., Jassby, A.D., Thompson, J.K., and Hieb, K.A. 2007. A cold phase of the East Pacific triggers new phytoplankton blooms in San Francisco Bay, *Proceedings of the National Academy of Sciences*, 104(47): 18561-18565, doi: www.pnas.org/cgi/doi/10.1073/pnas.0706151104

Cole, T.M., and Wells, S.A. 2015. CE-QUAL-W2: A Two-Dimensional, Laterally Averaged, Hydrodynamic and Water Quality Model, Version 3.72, Department of Civil and Environmental Engineering, Portland State University, Portland, OR. March.

Deltares. 2019a. D-Water Quality Process Library Description, Technical Reference Manual: Delft3D Flexible Mesh Suite, Version 5.01, January.

Deltares. 2019b. D-Water Quality, Process Library Tables, Version 5.0.2, Deltares, Delft, The Netherlands. May.

DiToro, D.M. 2001. Sediment Flux Modeling, Wiley-Interscience, New York.

Foxgrover, A.C., Finlayson, D.P., Jaffe, B.E., and Fregoso, T.A. 2011. Bathymetry and digital elevation models of Coyote Creek and Alviso Slough, South San Francisco Bay, California (ver. 5.0, June 2020): U.S. Geological Survey Open-File Report 2011–1315, 21 p., https://doi.org/10.3133/ofr20111315.

Kimmerer, W.J. and Thompson, J.K. 2014. Phytoplankton Growth Balanced by Clam and Zooplankton Grazing and Net Transport in the Low-Salinity Zone of the San Francisco Estuary, *Estuaries and Coasts*, 37:1202-1218, DOI 10.1007/s12237-013-9753-6.

King, A. 2019. Wind Over San Francisco Bay and the Sacramento-San Joaquin River Delta: Forcing for Hydrodynamic Models, Technical Report, March 25. Available from: https://drive.google.com/drive/folders/1g1D8xXDX0oclaMrCPNH5--h7Wgif7oXF

Lucas, L.V., and Thompson, J.K. 2012. Changing restoration rules: exotic bivalves interact with residence time and depth to control phytoplankton productivity. *Ecosphere*, 3(12): 1 - 25.

Ludwig, F. L. and Sinton, D. 2000. Evaluating an objective wind analysis technique with a long record of routinely collected data. *J. Appl. Meteor.*, 39:335-348.

Ludwig, F. L., Livingston, J. M., and Endlich, R. M. 1991. Use of Mass Conservation and Dividing Streamline Concepts for Efficient Objective Analysis of Winds in Complex Terrain, J. Appl. Meteorol., Vol. 30, pp. 1490–1499.

Martinez, M. and Perry, S. 2021. Interagency Ecological Program: Discrete water quality monitoring in the Sacramento-San Joaquin Bay-Delta, collected by the Environmental Monitoring Program, 1975-2020. ver 4. Environmental Data Initiative. https://doi.org/10.6073/pasta/31f724011cae3d51b2c31c6d144b60b0 (Accessed 2021-09-30).

Martyr-Koller, R.C., Kernkamp, H.W.J., van Dam, A., van der Wegen, M., Lucas, L.V., Knowles, N., Jaffe, B., and Fregoso, T.A. 2017. Application of an unstructured 3D finite volume numerical model to flows and salinity dynamics in the San Francisco Bay-Delta, 2017. Estuarine, Coastal and Shelf Science, Vol. 192, pp 86–107, https://doi.org/10.1016/j.ecss.2017.04.024.

Nuss, E., Zhang, Z., Holleman, R., Chelsky, A., Winchell, T., Wu, J., and Senn, D. 2018. Hydrodynamic and water quality model calibration and application in San Francisco Bay. 36 Technical report, San Francisco Estuary Institute, Richmond, CA. SFEI Contribution No. 913.

Paraska, D.W., Hipsey, M.R., and Salmon, S.U., 2014. Sediment diagenesis models: Review of approaches, challenges and opportunities, *Environmental Modelling and Software*, 61: 297-325, doi: http://dx.doi.org/10.1016/j.envsoft.2014.05.011

Pubben, S. 2017. 3D mixing patterns in San Francisco South Bay, 2017. Masters Thesis, TU Delft.

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

SFEI (San Francisco Estuary Institute). 2014a. External Nutrient Loads to San Francisco Bay. San Francisco Estuary Institute, Richmond, California. Contribution No. 704.

SFEI. 2014b. Scientific Foundation for the SanFrancisco Bay Nutrient Management Strategy. San Francisco Estuary Institute,Richmond, California. Contribution No. 979.

SFEI. 2015. Characterizing and quantifying nutrient sources, sinks and transformations in the Delta: synthesis, modeling, and recommendations for monitoring

SFEI. 2017. San Francisco Bay Interim Model Validation Report. Prepared by San Francisco Estuary Institute, Richmond, CA. December.

SFEI. 2018. Hydrodynamic and Water Quality Model Calibration and Application in San Francisco Bay. San Francisco Estuary Institute, Richmond, CA. Contribution #913. Published online: http://sfbaynutrients.sfei.org/books/reports-and-work-products

SFEI. 2019. 5-year modeling work plan (2019-2024).

SFEI. 2020. San Francisco Bay Numerical Modeling: FY2020 Update: Draft Report. San Francisco Estuary Institute, Richmond, CA. Contribution #1006. Published online: https://drive.google.com/file/d/1JndvEL2CeauGwF-tMinOH3lIyoBovQ-q/view?usp=sharing

SFEI. 2021. Delta-Suisun Biogeochemical Model: Calibration and Validation (WY2016, WY2011). San Francisco Estuary Institute, Richmond, CA. Contribution #1056. Published online: https://drive.google.com/file/d/11xHakErQsKvhI_UoEA0RJarBAb6Hmgmv/view

Testa, J., Brady, D., DiToro, D., Boynton, W., Cornwell, J., and Kemp, W. 2013. Sediment flux modeling: Simulating nitrogen, phosphorus, and silica cycles, Estuarine, Coastal and Shelf Science, 131: 245-263, doi: http://dx.doi.org/10.1016/j.ecss.2013.06.014

Thornton, K.W., and Lessem, A.S. 1978. A temperature algorithm for modifying biological rates, *Transactions of the American Fisheries Society*, 107(2): 284-287.

Wang, R., and Ateljevich, E. 2012. A Continuous Surface Elevation Map for Modeling (Chapter 6). In Methodology for Flow and Salinity Estimates in the Sacramento-San Joaquin Delta and Suisun Mars, 23rd Annual Progress Report to the State Water Resources Control Board. California Department of Water Resources, Bay-Delta Office, Delta Modeling Section.

White, S.W. 2019. No Impediment to the Sediment: A Numerical Model for Sediment Flux and Flocculation in San Francisco Bay [Undergraduate honors thesis]. Stanford University.

Willmott, C.J. 1981: On the Validation of Models, J. Geophys. Res., Vol. 112, C01013.

Zi, T., Mckee, L., Yee, D. and Foley, M. 2021. San Francisco Bay Regional Watershed Modeling Progress Report, Phase 1. SFEI Contribution No. 1038. San Francisco Estuary Institute: Richmond, CA.

Appendix A: Hydrodynamic Model Validation Results

The complete set of validation figures for the full resolution hydrodynamic model, water years WY2013 - WY2018, may be found in the following links:

Appendix A1: water level validation figures

Appendix A2: USGS mooring salinity validation figures

Appendix A3: USGS channel transect salinity validation figures

Velocity validation was conducted for WY2013 only -- see Nuss et al. (2018).

Appendix B: Evaluation of Stratification and Mixed Layer Depth in San Francisco Bay

Appendix C: Publicly Owned Treatment Work Concentration Time Series from WY2013-WY2018

Appendix D: Updated Full Resolution Results for WY2017

Appendix E: Updated Full Resolution Results for WY2013

Appendix F: Scatter Plots Showing Comparisons for Model State Variables and Derived Metrics

Appendix G: Biogeochemical Model Sensitivity Analyses Figures