

San Francisco Bay Nutrient Management Strategy

Open Bay Biogeochemical Modeling Progress Report for the Model Advisory Group

San Francisco Bay Numerical Modeling Update, FY2023-FY2026

Jordyn Moscoso¹

Pradeep Mugunthan²

Rachel Allen¹

Farid Karimpour¹

Daniel Killam¹

Allie King¹

David Senn¹

¹ SFEI

² Spheros Environmental



September, 2026

SFEI Contribution #1341

Open Bay Biogeochemical Modeling Progress Report for the Model Advisory Group: San Francisco Bay Numerical Modeling Update, FY2023-FY2026

Jordyn E. Moscoso, Pradeep Mugunthan, Rachel M. Allen, Farid Karimpour, Daniel
Killam, Allie King, and David Senn

San Francisco Estuary Institute (SFEI), Richmond, CA, USA

September 24, 2026

Acknowledgements

This project was supported by funding from the San Francisco Bay Nutrient Management Strategy and the United States Environmental Protection Agency (Agreement W9-97T22601, Tracking Bay Health).

This project has been funded wholly or in part by the United States Environmental Protection Agency under assistance agreement W9-97T22601 to the San Francisco Estuary Institute. The contents of this document do not necessarily reflect the views and policies of the Environmental Protection Agency, nor does the Environmental Protection Agency endorse trade names or recommend the use of commercial products mentioned in this document, as well as any images, video, text, or other content created by generative artificial intelligence tools, nor does any such content necessarily reflect the views and policies of the Environmental Protection Agency.

Contents

1	Introduction: Nutrient Management Strategy Modeling	1
1.1	Background	1
1.2	NMS Hydrodynamic and Biogeochemical Modeling History	1
1.3	Base Model Configuration	1
1.3.1	SFB BGCM	1
1.3.2	Nutrient Loads and Boundary Conditions	2
1.4	Modeling Peer Review	3
1.4.1	2023 MAG Recommendations	3
2	Hydrodynamic Model	5
2.1	Model domain and relationship to the BGC model	5
2.2	Inputs to the BGC model and uncoupled model structure	6
2.3	Updates to the hydrodynamic model since the 2023 update	6
2.4	Model Performance Summary	7
3	Biogeochemical Model Improvements	8
3.1	Updates to Light Extinction Coefficients	8
3.2	Updates to the Ocean Boundary Conditions	8
3.3	Water Column and Sediment Initial Conditions	8
3.4	Top-Down Control	9
3.5	Improvements to Sediment-Water Column Cycling of Organic Matter and Nutrients	10
3.5.1	Exploration of a Full Sediment Flux Model	10
3.5.2	Incorporation of New Denitrification Formulation in Sediment	11
3.5.3	Updates to Critical Shear Stress in Particle Settling and Resuspension	12
3.6	Improvements to Dissolved Oxygen	13
4	Model Calibration Approach	16
4.1	Setup and Focus	16
4.1.1	Improving Primary Production	16
4.1.2	Denitrification Drivers	16
4.1.3	Exploring DO Modifications	16
4.2	Data Sets for Evaluating Model Performance	19
4.2.1	High Frequency Mooring and Channel Data	19
4.2.2	Mapping Cruise Data	19
4.2.3	Remote Sensing Data	20
4.2.4	Sediment Flux Measurements	20
4.3	Approach for Evaluation of Model Performance	20
4.3.1	Graphical Evaluations	20
4.3.2	Statistical Metrics	21
4.3.3	Selection of Best Parameter Set(s)	21
4.3.4	Control Volume Analysis	22
5	Model Calibration Results and Discussion	24
5.1	Reference Model Comparison to Data	24
5.1.1	Phytoplankton and Primary Production	24
5.1.2	Dissolved Inorganic Nitrogen, Nitrate, and Ammonia	35

5.1.3	Phosphorus and Silicate	44
5.1.4	Dissolved Oxygen	47
5.1.5	Sediment Fluxes	55
5.2	Sensitivity Analysis	60
5.2.1	Fraction of Organic Matter Remineralized by Oxygen (frmindod1)	61
5.2.2	Half Saturation for Radiation in Phytoplankton (RadSat)	64
5.2.3	Maximum Phytoplankton Growth Rate (PPMaxDiat and PPMaxGreen)	67
5.2.4	Perturbations to Top-Down Control (Mort0 and QMort)	70
5.2.5	Sensitivity to Reaeration Coefficient	73
5.3	Summary and Discussion	78
5.3.1	Summary of Model Performance	78
5.3.2	Discussion	78
A	Supplemental Information for the Hydrodynamic Model	84
B	Supplemental Information for the Biogeochemical Model Base State	86
C	Supplemental Results for Biogeochemical Model Validation	89
D	Supplemental Results for Sensitivity Analysis	105

Acronyms and Abbreviations

Acronym	Definition
AOU	Apparent Oxygen Utilization
BGC	Biogeochemical
CC	Correlation Coefficient
CDF	Cumulative Distribution Function
CDFW	California Department of Fish and Wildlife
CV	Control Volume
DEB	Dynamic Energy Budget
DELWAQ or DWAQ	Delft-3D Water Quality Model (D-Water Quality)
DELWAQG	Full Sediment Flux Module within DELWAQ
DFM	D-Flow Flexible Mesh
DIN	Dissolved Inorganic Nitrogen
DO	Dissolved Oxygen
DOM	Dissolved Organic Matter
DW	Dry Weight
DWR	California Department of Water Resources
FNU	Formazin Nephelometric Unit
GPP	Gross Primary Productivity
HAB	Harmful Algal Bloom
HF	High Frequency
LSB	Lower South Bay
MAE	Mean Absolute Error
MAG	Model Advisory Group
NDBC	National Data Buoy Center
NMS	Nutrient Management Strategy
OLCI	Ocean and Land Color Instrument
PAR	Photosynthetically Active Radiation
POC / PON / POP	Particulate Organic Carbon / Nitrogen / Phosphorus
POM	Particulate Organic Matter
POTW	Publicly Owned Treatment Works
R ²	Coefficient of Determination
RMP	Regional Monitoring Program
RMSE	Root Mean Square Error
SAL	Salinity
SFB	San Francisco Bay
SFB-BGCM	San Francisco Bay Biogeochemical Model
SFEI	San Francisco Estuary Institute
SOD	Sediment Oxygen Demand
SS	Skill Score
STD	Standard Deviation
SUNA	SeaBird Scientific SUNA V2 Nitrate Sensor
UCSC	University of California, Santa Cruz

continued on next page

Acronyms and Abbreviations – continued from previous page

Acronym	Definition
USGS	United States Geological Survey
UTC	Coordinated Universal Time
WY	Water Year

1 Introduction: Nutrient Management Strategy Modeling

1.1 Background

The San Francisco Bay (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 (BGC) 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 5-year modeling work plans (2019-2024, and 2024 - 2028; SFEI 2019, SFEI, 2025).

This report serves as a progress and status update on NMS numerical modeling work, focusing on work carried out during portions of fiscal year (FY) 2023 (Feb – June 2023), and FYs 2024, 2025, and 2026.

1.2 NMS Hydrodynamic and Biogeochemical Modeling History

NMS work on the development of hydrodynamic and biogeochemical models began in 2016. Existing summaries describe the model development as of September 2020 (King et al., 2020) and January 2023 (King et al., 2023). In brief, by 2020, the hydrodynamic model was calibrated and validated for water years (WY) 2003, 2006, 2013, and 2017, including model input from a semi-empirical space-time varying wind field, which largely relied on observations. The SFB Biogeochemical Model (SFB-BGCM) was updated to version v1.0, including semi-empirical space-time varying wind, temperature, and light attenuation inputs, along with more mechanistically-realistic sediment diagenesis pathways. Calibration of the biogeochemical model focused on WYs 2013 and 2017. An aggregated-grid model was developed to enable faster run completion, also focused on Water Years (WY) 2013 and 2017, so that sensitivity analyses of biogeochemical model processes and parameter values could be tested. By 2023, the hydrodynamic model was calibrated and validated for WYs 2014-2016 and 2018, and further detailed validation of stratification and vertical mixing were performed. The hydrodynamic model outputs were tidally filtered and spatially averaged to enable the aggregated-grid biogeochemical model to run WYs 2013-2018, where grazing and sediment flux parameters were calibrated relying on water column and benthic observations and sensitivity tests were performed to inform uncertainties. The full biogeochemical model, updated to v3, was then run using the calibrated parameters for WYs 2013, 2017, and 2018.

In the following sections, we provide a high level overview of the hydrodynamic and biogeochemical model development as of FY2023 (hereafter referred as v2023), but refer the reader back to past reports for additional details (King et al., 2020, King et al., 2023). Throughout this report we refer to the current model presented in this report as v2026, which includes the updates made since the v2023 model.

1.3 Base Model Configuration

1.3.1 SFB BGCM

The previous version of the SFB-BGCM presented in King et al., 2023 tracks nitrogen, phosphorus, silicate, oxygen, along with dissolved and particulate organic matter (table 1, v2023), phytoplankton and zooplankton; many, but not all, of the same variables are simulated in the version presented in this report. Both the aggregated-grid and full-resolution simulations included primary production

Table 1: Variables simulated in DWAQ for each of the versions of the model, v2026 (this report), v2023 (King et al., 2023), and v2020 (King et al., 2020).

Modeled Variables	v2026	v2023	v2020
Ammonia (NH ₄)	x	x	x
Nitrate (NO ₃)	x	x	x
Phosphorus (PO ₄)	x	x	x
Silicate (Si)	x	x	x
Oxygen	x	x	x
Diatoms	x	x	x
Greens	x	x	
Benthic Diatoms			
Zooplankton		x	x
Labile POM	x	x	x
Refractory POM		x	x
Opal	x	x	x
Labile Sediment Detritus	x	x	x
Other Organic (Legacy) Sediment Detritus		x	x

Table 2: Processes simulated in DWAQ for each version of the model: v2026 (this report), v2023 (King et al., 2023), and v2020 (King et al., 2020), as detailed within section (§) of *D-Water Quality Processes Library Description* 2026.

Process	v2026	v2023	v2020	§
Phytoplankton Growth	x	x	x	4.10
Phytoplankton Linear Mortality	x	x	x	4.10
Phytoplankton Quadratic Mortality	x			this report: 3.4
Dynamic Energy Budget (DEB) Grazer		x	x	7.2
Organic Matter/Phyto Sedimentation	x	x	x	8.3, 13.4
Organic Matter Resuspension		x	x	13.4
Nitrification	x	x	x	3.1
Pelagic Denitrification	x	x	x	3.3
Sediment Denitrification	x			3.3
Remineralization	x	x	x	8.1, 8.4
Autolysis	x	x	x	4.10

and top-down control by zooplankton, organic matter cycling in both water column and sediments, and nutrient loss processes including denitrification and burial (table 2, v2023). Some of these processes were modified in the v2026 model presented in this report, including the removal of the DEB Grazer and the use of quadratic mortality for top-down control, and the implementation of coupled sediment denitrification. Changes to the v2026 model are detailed in section 3.

1.3.2 Nutrient Loads and Boundary Conditions

Hydrodynamic model inputs and boundary conditions are summarized in sections 2 and 2.2, and detailed in King et al. (2026). The BGC model includes nutrient inputs from each Publicly Owned Treatment Works (POTW) and refinery, which were sourced from a publicly available repository

(<https://bacwa.org/document-category/nutrient-annual-reports/>), with missing data filled with an estimate based on inter-annual trends and seasonal climatology (King et al., 2023, Appendix C). Loads entering from the Delta were estimated using daily discharge computed from the California Department of Water Resources DAYFLOW model and nutrient concentrations from long-term monitoring (King et al., 2023, Appendix C). Nutrient concentrations at the ocean boundary are detailed further in section 3.2.

1.4 Modeling Peer Review

The development of the SFB-BGCM has been periodically reviewed by a Model Advisory Group (MAG) that includes academic and policy experts in hydrodynamic modeling, biogeochemical modeling, field methods, and regulatory policy and implementation. The first series of modeling meetings were held in March 2022 to introduce the MAG to the science and management questions, and provide an overview of model development activities and initial model validation results. A second meeting was held in the Fall of 2022 to update the MAG on the emerging harmful algal bloom (HAB), and potential additional uses to the model to inform management decisions. An in-person MAG workshop was held in February 2023 to address two specific charge questions, including: (1) evaluate the status of model validation and identify areas of potential model improvements; and (2) evaluate the appropriateness of the model for prioritized applications.

1.4.1 2023 MAG Recommendations

The MAG found that, overall, the model generally reproduced the seasonal and spatial patterns of nutrient concentration, chlorophyll-a and dissolved oxygen (DO, Sutula et al., 2023). The MAG provided the following key recommendations:

1. Improve representation of temperature within the hydrodynamic model and verify the freshwater and ocean forcing.
2. Improve understanding of the various pools of nitrogen, and processes contributing to nitrogen cycling and losses, especially the role of denitrification in sediments.
3. Evaluate primary production through different metrics, including seasonal gross primary production (GPP) and alternative methods of representing phytoplankton biomass (including through empirical carbon:chlorophyll-a models).
4. Consider alternative metrics to validate DO predictions including: comparing model predictions to surface and bottom DO measurements separately, where feasible; use apparent oxygen utilization (AOU), which tracks deviations from saturation, to isolate the effects of biogeochemical processes on DO; consider sub-daily timescales rather than daily averaging; sensitivity of DO to biogeochemical rates of organic matter decomposition; and, developing oxygen mass budgets regionally to better understand the biogeochemical sources and sinks relative to transport and mixing processes.
5. Improve sediment-water column coupling of nutrient and organic matter exchanges, extend model validation to years with sediment flux measurements (WY2021), assess sediment oxygen demand (SOD) relative to total water column oxygen consumption, and include seasonal comparisons of nutrients and SOD fluxes.
6. Improve representation of light limitation in shoals, and conduct sensitivity analyses to different approaches to derive light extinction

7. Continue sensitivity analyses on the importance of top-down controls on primary production.
8. Develop a conceptual, system-wide synthesis and overall mass balance, and identify which of the system processes are simulated with confidence and those that are uncertain.

The MAG recommendations provided the basis for further model improvements discussed throughout the report. A table showing the location where we address each of the MAG recommendations are shown in Table 3

Table 3: Table of the locations of each of the sections or reports that address 2023 MAG Recommendations with brief description of the current status of each recommendation.

MAG Recommendation	Status	Location
Improve temperature in the hydrodynamic model	Completed	See King et al., 2026.
Improve understanding of nitrogen, especially the role of denitrification	Ongoing	Section 3.5 (specifically, 3.5.1 and 4.1.2), Section 5.1.2, Section 5.2.1.
Evaluate primary productivity and alternative methods of representing phytoplankton biomass	Ongoing	Section 4.1.1, Section 5.1.1, Section 5.2.4.
Improve methods of representing phytoplankton (carbon:chl-a models)	Not Started	
Alternative methods for DO validation	Completed	Metrics have been incorporated into model validation, see Section 5.1.4
Measure sensitivity of DO to biogeochemical rates of organic matter deposition, transport, and mixing	Ongoing	Section 4.1.3, Section 5.1.4, Section 5.2.5.
Improve sediment-water column coupling	Ongoing	Section 4.1.2, Section 5.1.5, Section 5.2.1.
Extend model validation to years with sediment flux measurements	Completed	This report
Improve representation of light limitation and conduct sensitivity analysis	Completed	Section 3.1, Section 5.2.2.
Continue sensitivity analysis on the importance of top-down control	Ongoing	v2026 uses quadratic mortality “grazing”. Section 3.4, Section 5.2.4.
Develop a conceptual, system-wide synthesis and overall mass balance	Ongoing	Report updates forthcoming.
Identify which system processes are simulated with confidence and those that are uncertain	Ongoing	Section 5.3.1.

2 Hydrodynamic Model

The BGC model is driven by output from the SFEI's three-dimensional hydrodynamic model, documented in full by King et al. (2026). The hydrodynamic model is built on the open-source D-Flow Flexible Mesh (DFM) solver, part of the Deltares Delft3D-FM suite that also includes the D-Water Quality (DWAQ) engine used for the BGC model. Because both models share this framework, the hydrodynamic model output (water levels, velocities, vertical mixing, salinity, and temperature) is used directly as input to the BGC model (King et al., 2026, Section 2.1). This section provides only a brief summary. Refer to King et al. (2026) for the complete model configuration, calibration, and performance evaluation.

2.1 Model domain and relationship to the BGC model

The model domain (Figure 1) extends from the lower Sacramento and San Joaquin Rivers and the Sacramento-San Joaquin Delta in the east, roughly 10 km upstream of their confluence, south to the tidal reaches of the Guadalupe River and Coyote Creek in Lower South Bay (LSB), and west into the coastal ocean approximately 40 km west of Half Moon Bay at the southern boundary and 20 km west of Point Reyes at the northern boundary.

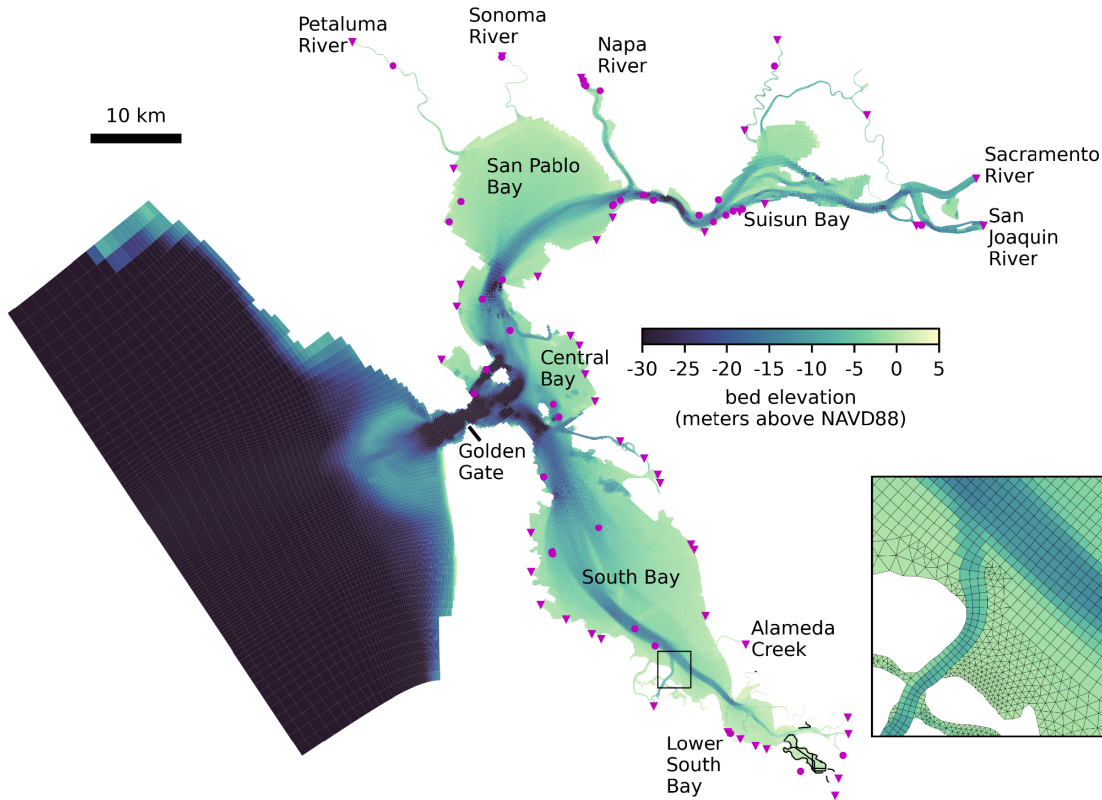


Figure 1: Hydrodynamic model domain, unstructured grid, and bathymetry, showing the locations of point-source discharges (circles) and stormwater inputs (triangles). From King et al. (2026, Figure 1).

The horizontal resolution ranges from about 20 m in the LSB sloughs to 2 km in the coastal ocean, with an intermediate resolution of roughly 250 m in the South Bay and 350 to 500 m in the

North Bay. Ten terrain-following sigma layers resolve the vertical structure of velocity, salinity, and temperature, allowing baroclinic circulation. The BGC model runs on a spatially aggregated version of this grid; the hydrodynamic output is tidally filtered and spatially averaged, and the aggregated-grid model was developed to enable faster run completion (Section 1.2). The aggregated-grid configuration uses tidally averaged hydrodynamics and externally calculated horizontal dispersion coefficients (King et al., 2023).

2.2 Inputs to the BGC model and uncoupled model structure

The hydrodynamic and BGC models are run in an uncoupled (offline) configuration (King et al., 2026, Section 2.1). The hydrodynamic model is executed first, one water year at a time, and its results are post-processed and supplied as fixed inputs to the DWAQ-based BGC model. The coupling is one-way: the hydrodynamic model provides the physical environment to the BGC model, including advective and diffusive transport, vertical mixing, water depth, and temperature, and the BGC model exerts no feedback on the hydrodynamics. Boundary forcing for the hydrodynamic model, including ocean, Delta, small-tributary, and point-source (POTW and refinery) inflows together with wind and meteorological forcing, is specified at the model boundaries as described in King et al. (2026), Section 2.3). Temperature is carried into the BGC model as a driver of temperature-dependent biogeochemical rates, including phytoplankton growth and the associated nutrient uptake, organic-matter mineralization, nitrification, and denitrification (King et al., 2020).

2.3 Updates to the hydrodynamic model since the 2023 update

Four changes to the hydrodynamic model since the 2023 update (King et al., 2023) are relevant to the BGC model:

1. The hydrodynamic model now dynamically simulates the water temperature through a new submodule of heat-exchange and air-sea heat exchange (King et al., 2026, Sections 2.1 and 2.3.5), developed in response to the 2023 MAG recommendation to improve temperature in the hydrodynamic model. The 2023 configuration did not model temperature; it instead supplied the BGC model with a spatially and temporally varying temperature field interpolated from USGS Peterson Cruise and California Department of Fish and Wildlife (CDFW) observations (King et al., 2023).
2. The simulation was extended to a continuous eleven-year period spanning water years 2013 through 2023 (King et al., 2026, Section 2.1), compared with the shorter set of individual water years simulated previously (King et al., 2020, 2023).
3. The model grid was refined, including straightening the coastal-ocean boundary to eliminate a spurious jet, and the model setup scripts were publicly released (King et al., 2026).
4. The performance of the hydrodynamic model is now documented in a dedicated standalone report evaluating water level, velocity, salinity, and temperature over the 11-year period (King et al., 2026).

Temperature performance was evaluated on 14 moored sensors located at USGS and SFEI mooring stations, using 15-minute observations throughout the 11-year simulation period (King et al., 2026, Section 3.3; Figure A1).

Air-sea heat exchange in the model uses DFM’s composite temperature model, with relative humidity, air temperature, and cloudiness specified across the domain; air temperature is drawn

only from the National Data Buoy Center (NDBC) stations closest to the water, which was found to greatly improve the model’s predictions of water temperature (King et al., 2026, Section 2.3.5). Predictions within 1 °C of observations are generally considered good for a hydrodynamic model, and the model bias is below 1 °C at every site, with performance accurate at 12 of the 14 sensors and acceptable at the remaining two (King et al., 2026, Section 3.3; Figure A2).

Biases are slightly negative (cooler) at most sites and show a small to moderate positive bias at the two sites nearest the ocean, Alcatraz and the Richmond-San Rafael Bridge, and across all sites the model predicts slightly more variance in temperature than is observed. The model reproduces the seasonal and interannual temperature cycle as well as diurnal variability (King et al., 2026, Sections 3.3 and 3.4). Because biogeochemical rates are temperature-dependent in the model (King et al., 2020), this update improves a physical driver of the representation of primary production and nutrient cycling of the BGC model.

2.4 Model Performance Summary

Over the eleven-year simulation period (water years 2013 through 2023), the model reproduces the observed physics of the Bay well across all four evaluated variables. Measured against the performance standards used in other San Francisco Bay modeling studies, King et al. (2026) report that the model predicts water level accurately at six of seven sensor sites (acceptable at the seventh), depth-averaged velocity accurately at 35 of 41 sites (acceptable at four, poor at two), salinity accurately at 11 of 14 sites (acceptable at the remaining three), and temperature accurately at 12 of 14 sites (acceptable at the remaining two). Overall, model performance is comparable to or better than that of other existing hydrodynamic models of the Bay. Station-by-station comparisons, performance metrics, and supporting figures are provided in King et al. (2026).

3 Biogeochemical Model Improvements

The FY2023 – FY2026 biogeochemical model development activities focused on various improvements based on MAG recommendations and targeted improvements to certain model state variables based on new lines of evidence.

3.1 Updates to Light Extinction Coefficients

The previous configurations of the biogeochemical model used spatially variable light attenuation derived from statistical analysis, described in King et al., 2020, Appendix 1C. This approach predominantly used USGS Channel Cruise data, and limited high frequency data. The lack of spatial coverage in monitoring stations limited the ability to capture potential channel to shoal lateral gradients in turbidity and hence light extinction coefficients.

To address this, we combined remotely sensed turbidity data, τ_{fnnu} (described below in Section 4.2.3), to derive the light extinction coefficients. The remotely sensed data was extrapolated to fill any missing data using inverse distance weighting, then the light extinction coefficient, K_d , for the combined data, was calculated following Richardson et al., 2023,

$$K_d = 0.11\tau_{\text{fnnu}} + 0.52. \quad (1)$$

3.2 Updates to the Ocean Boundary Conditions

The ocean boundary conditions are set along the western boundary of the model domain. In the 2023 BGC model, the ocean boundary conditions are constant in both space and time. Based on the MAG recommendations to reassess the ocean boundary conditions, we analyzed the coastal ocean data reported in Sandoval-Belmar et al., 2026. The data used in this study includes all available observations from 1980 to 2020 and is used to create a climatological and depth-averaged value. The parameters evaluated from this dataset and the corresponding ocean boundary conditions derived are listed in Table 4.

Table 4: Table of the constant values used in the ocean boundary conditions

Parameter	μM	mg/l
NO_3	18	0.253
NH_4	2	0.028
PO_4	1.47	0.046
SiO_2	22	0.616
O_2	186.94	5.98

The 2023 MAG Report recommends sensitivity analysis to the ocean boundary conditions including the addition of space and time variable ocean boundary conditions, while we have updated the constant boundary conditions and conducted sensitivity analysis, we have yet to implement space and time varying boundary conditions. This will be a priority in upcoming model development.

3.3 Water Column and Sediment Initial Conditions

For the aggregated grid model (v2026) are derived from an 11 year aggregated grid model run. The parameters used in the 11 year run are the same as those used in the reference version of the model (v2026). The model is initialized on Aug 1, 2012, and run until July 1, 2022. This is done because this model is not currently calibrated to simulate the 2022 harmful algal bloom. The initial conditions for this model are constant representative values that are described in King et al., 2020.

Examples of modeled quantities including chlorophyll and nitrogen over the ten year run are shown in Figures A7 and A8, respectively. This choice in model initial conditions allows for a long

spin-up of model state from constant initial conditions Bay-wide. Previous versions of this model (v2020, v2023) include active refractory particulate organic matter, refractory sediment detritus, and legacy sediment material. However, these values are difficult to constrain, so v2026 does not include them in the modeled variables. This choice was made to reduce the amount of uncertainty in the modeled state. As a consequence previous initial conditions as reported in King et al., 2020 and King et al., 2023 for the refractory and legacy sediment material are not included in this model configuration.

The initial conditions for the aggregated grid run are taken from a two week average model state around the initial start date, Jan 1, 2020. This includes all modeled variables except chlorophyll concentration. We find that overall, the modeled chlorophyll concentration is too high compared to the data (see Figure A7). Given the availability of the remote sensing chlorophyll data (see Section 4.2.3), we derive a two week average surface chlorophyll concentration from the remote sensing. The chlorophyll concentration is then converted to total phytoplankton biomass, and the biomass is evenly distributed between the diatom and the greens. This choice matches modeled diatom ratios during the season outside of the spring bloom and summertime sustained productivity (see September 2020, Figure A9). The reference aggregated grid run is integrated from Jan 1, 2020 to July 1, 2022, we refer to this model state as Aggregated (v2026).

The initial conditions from the full resolution version of the model are derived from the Aggregated (v2026) model state on Aug 1, 2020. We use the model state from the aggregated grid run to minimize the number of internal oscillations during the full resolution model spin-up for all modeled quantities except chlorophyll. For each of the other quantities, the data from Aggregated (v2026) is averaged over two weeks, and linearly interpolated and extrapolated using nearest neighbors onto the full resolution grid. Because the aggregated grid is much more coarse than the full resolution grid, there is necessarily a spin-up period of approximately three months. For chlorophyll, we initialize using the remote sensing data which is interpolated and extrapolated onto the model grid. The model is integrated from Aug 1, 2020 to Oct 1, 2021 and is referred to in this report as Full Resolution (v2026).

3.4 Top-Down Control

The 2023 Model used dynamic energy-budget (DEB) based top-down control, with one generic zooplankton representing both pelagic and benthic grazing pressures on phytoplankton. Previous sensitivity analyses findings indicated that top-down control was necessary to keep modeled phytoplankton biomass to be within the range of measurements. However, the DEB model is heavily parameterized with little information available to constrain any of the numerous coefficients used within the model. The previous sensitivity analyses showed that zooplankton rapidly responded to increases in biomass or not at all even with minor perturbations in parameters, resulting in large amplitude phytoplankton-zooplankton oscillations. While rapid changes in phytoplankton levels are common in the SF Bay, it is unclear whether or not large predator-prey oscillations are consistent with observations. Mesocosm experiments by Franks et al., 1986 indicate such oscillations are not realistic. Furthermore, while zooplankton and benthic grazer biomass data and grazing rates are available within San Pablo and Suisun Bays (Kimmerer et al., 2014, Zierdt Smith et al., 2023), information in much of South Bay and LSB is limited, particularly for recent years when spatially extensive nutrients and turbidity data are available to better constrain and validate the model.

To overcome these oscillatory artifacts from the DEB module, we implemented quadratic phytoplankton mortality to simplify top-down control. Previous work by Murray et al., 1999 found that with quadratic phytoplankton mortality, the steady-state phytoplankton biomass increases with the square root of nutrient loading until the phytoplankton growth is nutrient-saturated at high

loads. While this decision is one way to control phytoplankton biomass and reduce the number of degrees of freedom for model calibration other work has indicated that zooplankton quadratic mortality may be more relevant (Franks, 2002). For now, we implement a quadratic mortality term on phytoplankton, and will explore further calibration of the DEB model in future efforts, as necessary.

The quadratic mortality term is defined as,

$$M_{\text{mort}} = k_{\text{mort}} \alpha_{\text{mort}}^{T-20} P^2, \quad (2)$$

where P is the phytoplankton biomass, and k_{mort} and α_{mort} are user-defined parameters in the allometric temperature function in phytoplankton biomass. The various fluxes of mortality include autolysis, and the transfer to the labile and refractory detrital pools, and follow the original formulations implemented within DWAQ for first-order phytoplankton mortality. These are computed as fractions of carbon returned to the system by autolysis, $F_{\text{det,aut}}$, fraction of carbon diverted to the labile pool, $F_{\text{det,lab}}$, and refractory pool, $F_{\text{det,refrac}}$. These are determined by user-defined parameters, r_{refract} and r_{aut} , the fraction of all detritus diverted to the refractory and autolysis pools respectively. Note that the fraction diverted to the labile pool is then determined by, $r_{\text{lab}} = 1 - r_{\text{refract}} - r_{\text{aut}}$. The transfer of phytoplankton mortality to the detrital carbon is defined by

$$f_{\text{lab,C}} = \frac{1 - r_{\text{refract}} - r_{\text{aut}}}{1 - r_{\text{aut}}}, \quad (3a)$$

$$F_{\text{detC,lab}} = f_{\text{lab,C}} M_{\text{mort}}, \quad (3b)$$

$$F_{\text{detC,refract}} = (1 - f_{\text{lab,C}}) M_{\text{mort}}. \quad (3c)$$

Then the transfer of nitrogen from phytoplankton mortality are given by,

$$F_{\text{detN, aut}} = r_{\text{aut}} \alpha_{\text{C:N}} M_{\text{mort}}, \quad (4a)$$

$$F_{\text{detN, lab}} = r_{\text{lab}} \alpha_{\text{C:N}} M_{\text{mort}}, \quad (4b)$$

$$F_{\text{detN, refract}} = r_{\text{refract}} \alpha_{\text{C:N}} M_{\text{mort}}. \quad (4c)$$

The transfer of detrital phosphorus and silicate follow Equations 4a-4c using the corresponding stoichiometric conversions for C:P and C:Si, respectively.

3.5 Improvements to Sediment-Water Column Cycling of Organic Matter and Nutrients

One of the main recommendations from the MAG is to reevaluate the sediment-water column exchanges of nutrients and organic matter. We evaluated the following improvements within the water quality model.

3.5.1 Exploration of a Full Sediment Flux Model

We evaluated the use of a more comprehensive sediment flux module (DELWAQG) instead of the simplified S1/S2 sediment flux module used in the 2023 model. For context, the S1/S2 module has two sediment layers, does not have any sediment porewater, and provides a bulk simulation of diagenesis of particulate organic matter, and releases nutrients to, and removes oxygen and nitrate from, the bottom layer of the water column. The full sediment flux module (DELWAQG) provides the capability to simulate more than 2 sediment layers, includes sediment porewater, and a full

coupling of all water column state variables through (1-D) advective and diffusive exchange at the sediment-water interface. In addition, it has more rigorous formulations for representing sediment redox conditions following the electrochemical series for electron acceptors (oxygen consumption, denitrification, manganese reduction, iron reduction, sulfate reduction, and methanogenesis). The DELWAQG exploration did not provide any significant improvement in the performance of the model. Core issues with limited SOD and denitrification were not resolved due to high sediment porewater DO and limited organic matter deposition. Additionally, there were some limitations in the DELWAQG's coupling of some state variables between the water column and sediments - the pelagic and benthic phytoplankton coupling to sediments was incomplete. Deltares addressed the pelagic phytoplankton coupling, but the benthic coupling remained unresolved as of 2025. However, benthic phytoplankton coupling is not an issue for the base model which did not utilize microphytobenthos. Nonetheless, even after Deltares corrected the pelagic coupling, model performance did not improve notably. This suggested that the primary issues originated in the water column rather than within the sediments. Due the lack of improvements in the overall model performance even with a more rigorous sediment flux module, and recognizing the more extensive parameterization needed for DELWAQG that cannot be adequately constrained with existing information (for example, we do not have sediment porewater data), we pivoted to other approaches discussed below.

3.5.2 Incorporation of New Denitrification Formulation in Sediment

DWAQ provides the ability for customized processes defined by the user. We utilized this capability to incorporate a better denitrification formulation than the one used in the S1/S2 module. Within the S1/S2 module, denitrification is not coupled to organic matter concentration within the sediments, rather it is simulated as a first-order removal of nitrate directly from the bottom layer of the water column under the assumption that sediment organic matter pool is sufficiently high, and anoxic conditions exist in deeper sediments to sustain denitrification so long as there is nitrate available. To overcome this limitation, we coupled denitrification with remineralization of organic matter. We used a fraction of aerobic remineralization as a user input to allow both DO and nitrate to fuel remineralization. The fraction is adjusted internally, if necessary, based on the bottom water column DO and nitrate concentrations as follows:

$$F_{\text{ox}} = \begin{cases} 0 & \text{if } O_2 < O_2^{\min} \\ 1 & \text{if } NO_3 < NO_3^{\min} \text{ and } O_2 > O_2^{\min} , \\ f_{\text{ox}} & \text{user defined, otherwise} \end{cases} \quad (5)$$

and

$$F_{\text{nit}} = \begin{cases} 0 & \text{if } NO_3 < NO_3^{\min} \\ 1 & \text{if } O_2 < O_2^{\min} \text{ and } NO_3 > NO_3^{\min} , \\ f_{\text{nit}} & \text{user defined, otherwise} \end{cases} \quad (6)$$

where, $O_2^{\min} = 0.01\text{g/m}^3$ or 0.31 mmol/m^3 and $NO_3^{\min} = 0.001\text{g/m}^3$ or 0.0714 mmol/m^3 .

The remineralization of organic carbon is given by

$$R_{\text{remin,C}} = \gamma(T)_{\text{rm}} r_{\text{rm,C}} C_{\text{sed}}, \quad (7)$$

where,

$$\gamma(T)_{\text{rm}} = \alpha_{\text{rm}}^{T_{\text{bot}} - 20}, \quad (8)$$

is the Arrhenius temperature function for denitrification, C_{sed} is the organic carbon concentration in the sediments, T_{bot} is the temperature in the bottom grid cell, $r_{\text{rm,C}}$ is the user defined remineralization rate for carbon, and α_{rm} is the coefficient for the temperature function. The general formulation of denitrification is given by,

$$R_{\text{denit}} = \alpha_{\text{N:C}} F_{\text{nit}} R_{\text{remin,C}}, \quad (9)$$

where, $\alpha_{\text{N:C}}$ is the nitrogen to carbon ratio for denitrification, 0.933 gN/gC, or 0.8 mol N/mol C (Sarmiento, 2013). The SOD can be calculated by

$$\text{SOD} = \alpha_{\text{O}_2:\text{C}} F_{\text{ox}} R_{\text{remin,C}}. \quad (10)$$

$\alpha_{\text{O}_2:\text{C}}$ is the oxygen to carbon ratio, 2.67 gO₂/gC or 1 mol O₂/mol C. Similarly, the remineralization of nitrogen and phosphorus are defined as,

$$R_{\text{remin,N}} = \gamma(T)_{\text{rm}} r_{\text{rm,N}} N_{\text{sed}} \quad (11)$$

and,

$$R_{\text{remin,P}} = \gamma(T)_{\text{rm}} r_{\text{rm,P}} P_{\text{sed}}. \quad (12)$$

3.5.3 Updates to Critical Shear Stress in Particle Settling and Resuspension

Particulate organic matter (detritus) settles from the water column onto the bed, becoming part of the sediment detritus pools. Once in the sediment, the detritus can remain where it is permanently, be resuspended into the water column, or be mineralized into inorganic nutrients (diagenesis) which can then diffuse back into the bottom layer of the water column. Critical shear stresses can be used to limit both settling and resuspension of particulate organic matter/detritus. The volumetric settling rate ($R_{\text{set},i}$) for substance i is reduced when the bed shear stress τ increases to near the critical shear stress for settling ($\tau_{\text{cr},i}$), following equation 13. the settling rate, given primarily by the settling velocity s_i , the concentration C_i , and the water depth H , is at its maximum. To prevent negative settling fluxes, $\left(1 - \frac{\tau}{\tau_{\text{cr},i}}\right)$ is supposed to be strictly non-negative. When there is no bed shear stress,

$$R_{\text{set},i} = \left(1 - \frac{\tau}{\tau_{\text{cr},i}}\right) \times \frac{s_i \times C_i}{H} \quad (13)$$

Resuspension of detritus is limited by a critical shear stress in a similar fashion, though here the bed shear stress τ must exceed the critical shear stress for resuspension $\tau_{\text{cr},j}$ in order for resuspension to occur, where $j = 1$ or 2 denotes the sediment layer (S_j). In simple form, the resuspension flux $F_{\text{res},j}$ is given by equation 14, where $F_{\text{res}0}$ is the zero-order resuspension flux, r is the resuspension rate, C_{dm} is the amount of dry matter in the sediment, and A is the surface area of the sediment-water interface. Similar to settling, to prevent negative resuspension fluxes, $\frac{\tau}{\tau_{\text{cr},j}} - 1$ is supposed to be strictly non-negative, even if the bed shear stress drops below the critical shear stress ($\tau_{\text{cr},j}$).

$$F_{\text{res},j} = \left(\frac{\tau}{\tau_{\text{cr},j}} - 1\right) \times \left(F_{\text{res}0} + r \times \frac{C_{\text{dm},j}}{A}\right) \quad (14)$$

The BGC model does not simulate full sediment transport, and therefore, these formulations based on critical shear stress to limit settling and resuspension were not constrained based on observed bed deposition or erosion rates as would be done in a typical sediment transport model.

Such attempts to simulate bed sediments for particulate organic matter (detritus) alone would not be meaningful without also simulating inorganic suspended sediments, and the full suite of water column processes that affect sediment transport (such as flocculation, wave action, etc.). Thus, we chose to set the critical shear stress for resuspension $\tau_{cr,j}$ to a sufficiently large value such that organic matter and phytoplankton that reach the sediment bed would stay in the bed. The net result is that sediment-water column exchanges are primarily driven through sedimentation, and remineralization. This approach is consistent with other estuarine BGC models which also do not include resuspension of particulate organic matter (Cerco et al., 2017).

3.6 Improvements to Dissolved Oxygen

Many of the improvements to primary production and sediment-water column exchanges also have a direct bearing on DO. In this section, we explore parameters that operate directly on atmospheric exchange of DO and the DO saturation in water.

In DELWAQ the oxygen flux across the air-water interface is calculated as a first-order process proportional to the difference between the saturation concentration and the simulated dissolved oxygen concentration in the water column, as shown below:

$$R_{\text{rear}} = k_{l,\text{rear}} \times \frac{(\text{DO} - \max(\text{DO}_{\text{sat}}, 0.0))}{H} \quad (15)$$

$$k_{l,\text{rear}} = k_{l,\text{rear}20} \times t_c^{(T-20)} \quad (16)$$

$$k_{l,\text{rear}20} = \left(\frac{a \times v^b}{H^c} \right) + (d \times W^2) \quad (17)$$

$$\text{DO}_{\text{sat}} = f(T, \text{Sal}) \quad (18)$$

$$f_{\text{sat}} = 100 \times \frac{\max(\text{DO}, 0.0)}{\text{DO}_{\text{sat}}} \quad (19)$$

with:

- a, b, c, d : coefficients with different values for each reaeration option
- DO : actual dissolved oxygen concentration [$\text{gO}_2 \text{ m}^{-3}$]
- DO_{sat} : saturation dissolved oxygen concentration [$\text{gO}_2 \text{ m}^{-3}$]
- f_{sat} : percentage of saturation [%]
- H : depth of the water column [m]
- $k_{l,\text{rear}}$: reaeration transfer coefficient in water [m d^{-1}]
- $k_{l,\text{rear}20}$: reaeration transfer coefficient at reference temperature 20°C [m d^{-1}]
- t_c : temperature coefficient of the transfer coefficient [–]
- R_{rear} : reaeration rate [$\text{gO}_2 \text{ m}^{-3} \text{ d}^{-1}$]
- SAL : salinity [kg m^{-3}]
- T : temperature [$^\circ\text{C}$]
- v : flow velocity [m s^{-1}]
- W : wind speed at 10 m height [m s^{-1}]

387 The two key parameters that control the air-water oxygen exchanges are the values of the
 388 coefficients a , b , c and d that control the calculation of reaeration coefficient ($k_{l, \text{rear}}$), and the
 389 formulations used for calculation of DO saturation. $k_{l, \text{rear}}$ may be calculated using one of ten
 390 formulations based on water speed, water depth, wind speed, and temperature. DELWAQs different
 391 reaeration formulations are optimized for performance across various water bodies including lakes,
 392 rivers, stream, estuaries and open ocean. Of these, there is only one formulation (option 7) that
 393 has been optimized for the calculation of reaeration in estuaries. We used this formulation in our
 394 runs because it is catered to estuaries. The formulation is based on O'Connor et al., 1956 and
 395 Banks et al., 1977 with coefficients below:

$$a = 3.863, \quad b = 0.5, \quad c = 0.5, \quad d = 0.065 k_{l, \text{rear}0}$$

396 rear0 is a user specified modifier to modulate the effect of wind speed on air-water exchanges.
 397 The default values in DELWAQ is 1.0 remained unmodified from the v2023 model.

398 The saturation oxygen concentration can be calculated from water temperature and salinity
 399 using one of two available formulations:

400 FORMULATION 1: (OPTION `SWSAT0xy = 1` IN DELWAQ): This formulation calculates the dis-
 401 solved oxygen saturation concentration (DO_{sat}) in $[\text{g O}_2 \text{ m}^{-3}]$ using a third-order polynomial in
 402 temperature and a first-order correction for chloride concentration. The chloride concentration
 403 (C_{cl}) is in $[\text{g Cl m}^{-3}]$.

$$\text{DO}_{\text{sat}} = (a - bT + cT^2 - dT^3) \left(1 - \frac{C_{cl}}{m}\right). \quad (20)$$

404 FORMULATION 2: (OPTION `SWSAT0xy = 2` IN DELWAQ): This formulation represents the tem-
 405 perature dependency of oxygen solubility as a function of temperature function (T_f) and salinity.
 406 Unlike the previous option, this option directly depends on salinity and not chloride.

$$\text{DO}_{\text{sat}} = \exp \left(a + \frac{b}{T_f} + c \ln(T_f) + dT_f + SAL (m + nT_f + oT_f^2) \right) \frac{32000}{22400}. \quad (21)$$

407 where the temperature function (T_f) is defined as:

$$T_f = \frac{T + 273}{100}. \quad (22)$$

408 In this equation there is a factor $\frac{32000}{22400}$ which is the ratio of the molecular weight of oxygen
 409 ($32000 \text{ mg mol}^{-1}$) to the molar volume of ideal gas ($22400 \text{ mL mol}^{-1}$). This ratio means:

$$1 \text{ mL O}_2 = 1.4286 \text{ mg O}_2 \quad (23)$$

410 Table 5 defines the parameters, and Table 6 presents the coefficient values used by DELWAQ.

Table 5: Definitions of variables and empirical parameters used in the dissolved oxygen saturation formulations.

Parameter	Definition	Units
DO_{sat}	Saturation dissolved oxygen concentration	$\text{gO}_2 \text{ m}^{-3}$ (or $\text{mgO}_2 \text{ L}^{-1}$)
T	Water temperature	$^{\circ}\text{C}$
T_f	Temperature function	Dimensionless
SAL	Salinity	kg m^{-3} (or g l^{-1})
C_{cl}	Chloride concentration	g Cl m^{-3}
a, b, c, d	Empirical temperature coefficients	Dimensionless
m, n, o	Empirical salinity coefficients	Dimensionless

Table 6: Coefficients used in the two formulations for the dissolved oxygen saturation concentration.

SWSatOxy	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>m</i>	<i>n</i>	<i>o</i>
1	14.652	0.41022	0.089392	0.042685	10^5	–	–
2	–173.4292	249.6339	143.3483	–21.8492	–0.033096	0.014259	–0.0017

411 Formulation 1 is the default in DELWAQ, and was used in v2023 model. Because chloride is not
412 explicitly simulated in the model, C_{cl} defaulted to zero, which effectively eliminated any correction
413 for salinity. Because the model simulates both temperature and salinity, we switched to using
414 Formulation 2 in the v2026 model. This provided significant improvements to DO within Suisun
415 Bay, which is more sensitive to salinity effects on DO due to fluctuations in freshwater inflow from
416 the Delta.

4 Model Calibration Approach

4.1 Setup and Focus

The model calibration approach built on lessons learned from past parameter explorations and sensitivity analyses in the model, and incorporated MAG recommendations on specific improvements. We explored key processes by testing a series of hypotheses described below to progressively refine model performance for specific processes/state variables. The process improvements were discussed in Section 3, and the hypotheses explored are outlined below. We note that even though the hypotheses are discussed as individual categories, because of the interconnected nature of the biogeochemical processes, these were explored in tandem to some extent, and lessons learned from one informed improvements in other categories. We discuss these inter-relationships throughout. For each hypotheses the parameter space for one or a group of related parameters was explored through a latin hypercube sampling (LHS) approach within reasonable parameter ranges. These explorations predominantly occurred through the aggregated grid model.

4.1.1 Improving Primary Production

The objective of exploring primary production improvements was two-fold: (1) explain the chlorophyll-a and GPP measurements; (2) explain the seasonal draw down and recovery patterns of nutrients in the water column. A secondary objective of these explorations was to provide sufficient flux of labile organic matter to the sediments to sustain sediment nitrogen removal through denitrification.

Primary production improvements entailed exploring the parameter spaces for diatoms and green algae growth, mortality and settling parameters. Explorations for mortality included both first-order mortality, and the top-down control introduced through quadratic mortality (Table 7). We note that all of these parameters also directly affect nutrient uptake and cycling within the water column, and fluxes of biomass/labile organic matter to the sediments. Improvements to light extinction based on satellite data (Section 3.1) also potentially contributed to these improvements, but were not a part of the parameter exploration process.

4.1.2 Denitrification Drivers

Denitrification was explored through two avenues. We first explored a range of zero-order denitrification rates, adjusted for temperature, to assess whether the observed dissolved inorganic nitrogen (DIN) drawdown patterns can be reproduced. The second-set of simulations explored the coupled denitrification process parameters, primarily the fraction of organic matter remineralized through denitrification, and the first-order diagenesis rate (which also affected SOD and DO).

4.1.3 Exploring DO Modifications

Dissolved oxygen was explored through modifications to the formulation and parameterization of DO saturation dependence on salinity and temperature, reaeration coefficient and piston velocity, oxygen consumption in the sediment, oxygen production due to phytoplankton and benthic algae, and DO experiments on reaeration and critical shear stress.

Table 7: Primary Production and Phytoplankton Parameters Explored for GPP and Water Column Nutrient Cycling Improvements.

Process	Parameter Name	Description	Calibrated Value	Range Explored
Light Extinction	ExtVIBak	Background extinction coefficient	Remote sensing Kd	GAM Kd from King et al., 2023
Phytoplankton Growth	RadSatDiat	Half saturation radiation for diatoms	40.5 W/m ²	[20.25, 60.75]
	RadSatGree	Half saturation radiation for green	17 W/m ²	[8.5, 25.5]
	PPMaxDiat	Maximum growth rate for diatoms	2.4 1/d	[1.8, 3.0]
	PPMaxGreen	Maximum growth rate for diatoms	3.4 1/d	[2.55, 5]
Phytoplankton Mortality	Mort0Diat	Linear mortality diatoms	0.044 1/d	[0.033, 0.055]
	Mort0Green	Linear mortality green	0.04 1/d	[0.03, 0.05]
	kqmrtdiat	Quadratic mortality diatoms	0.173 m ³ /mg C/d	[0.13, 0.216]
	kmrtgrn	Quadratic mortality green	0.61 m ³ /mg C/d	[0.463, 0.763]

Table 8: Nutrient Parameters Explored for Water Column Nutrient Cycling Improvements

Process	Parameter Name	Description	Calibrated Value	Range Explored
Sediment Denitrification	frmindod1	Fraction mineralized by oxygen	0.9	[0.8, 0.95]
	rsdet1c	Sediment remineralization rate	0.146 1/d	[0.11, 0.1825]
Water Column Remineralization	ku_dFdc20	Upper limit mineralization	0.16 1/d	[0.12 ,0.2]
	kl_dFdc20	Lower limit mineralization	0.12 1/d	[0.09, 0.15]
Organic Matter Settling	V0SedDiat	Settling rate diatoms	0.21 m/d	[0.16, 0.27]
	V0SedGreen	Settling rate green	0.075 m/d	[0.06, 0.09]
	V0SedPOM1	Settling rate for all labile OM	0.5 m/d	[0.25, 0.75]
Ocean Boundary Conditions	NO ₃ ^{ocean}	Nitrate concentration at ocean boundary	18 mmol/m ³	[13.5, 22.5]
	NH ₄ ^{ocean}	Ammonia concentration at ocean boundary	2 mmol/m ³	[1.75, 2.5]
	PO ₄ ^{ocean}	Phosphorus concentration at ocean boundary	1.47 mmol/m ³	[1.10, 1.84]
	Si ^{ocean}	Silicate concentration at ocean boundary	22 mmol/m ³	[16.5, 27.5]
	O ₂ ^{ocean}	Oxygen concentration at ocean boundary	186.94 mmol/m ³	[140.2, 233.7]

4.2 Data Sets for Evaluating Model Performance

4.2.1 High Frequency Mooring and Channel Data

In WY2021, high frequency measurements of chlorophyll, dissolved oxygen, and other parameters collected with an EXO2 are available at San Leandro Marina, SFEI Shoal Buoy, Hayward, San Mateo Bridge, and Dumbarton Bridge (Table 9, Chelsky et al., 2026; Noland et al., 2025). In addition, nitrate estimates, computed from EXO2 and SUNA observations, exist at the Dumbarton Bridge and SFEI Shoal Buoy.

Table 9: A subset of station information published in Noland et al., 2025 used to validate the model from the NMS Moored Sensor Program stations.

Station	ID	Height above bed (m)	Depth (m)	Installed	Habitat type	Sensor
San Leandro Marina	SLM	1.3	0.5–4.5	Nov 2020	Shoal	EXO, LL, BL
SFEI Shoal Buoy	SHL	–	1.3	Jan 2020	Shoal	EXO, SUNA
Hayward	HAY	1	0.75–4.0	Oct 2020	Shoal	EXO
San Mateo Bridge	SMB	18.5	2.5–5.5	Jul 2014	Subsurface channel	EXO
Dumbarton Bridge	DMB	12.6	1–4	Jul 2013	Subsurface tidal channel	EXO, SUNA

Notes:

EXO – YSI EXO2 multiparameter water quality sonde

SUNA – SeaBird Scientific SUNA V2 nitrate sensor

LL – Solinst Levellogger 5 Junior depth sensor

We use the USGS data along the channel cruise locations (Schrage et al., 2017). This data is collected approximately bi-weekly to monthly and is available for surface nutrients (nitrate, ammonia, silicate, phosphorus), and depth profiles of chlorophyll and DO. In addition to this, we use discrete measurements of chlorophyll, nutrients and DO from the Interagency Ecological Program (Martinez et al., 2021) for model-data comparisons in Suisun Bay.

4.2.2 Mapping Cruise Data

High-frequency mapping data was collected in 2021 on March 23rd, May 5th, June 23rd, August 4th, September 16th and October 20th from a flow-through system on the USGS Research Vessel Aiken, with YSI EXO2 and Seabird SUNA V2 sondes. Each cruise proceeded on a zig-zag path across the South Bay from Emeryville to Dumbarton Bridge and returning to Emeryville, from approximately 06:00 to 14:00 local standard time. The data collected includes chlorophyll-a, dissolved oxygen, turbidity, nitrate+nitrite, along with other parameters at 1-second resolution, equating to one data point approximately every 8-9 meters. Associated discrete measurements of a full suite of nutrient parameters are collected at approximately 15 named stations, and up to 90 discrete samples of ammonium are collected every 2 km of travel. See Killam et al., 2023 for more details.

4.2.3 Remote Sensing Data

Remote sensed chlorophyll-a in $\mu\text{g/L}$ is generated from the Ocean and Land Color Instrument (OLCI) aboard the Sentinel-3 satellite, using a scattering line height algorithm described in Kudela et al., 2024, following an assessment from Wynne et al., 2022. Currently, the remote sensing calibration for chlorophyll has some major limitations, specifically the chlorophyll algorithm shows uncertainty in estimating concentrations larger than $10 \mu\text{g/L}$. Despite this uncertainty, it provides a useful indicator of the general spatial gradients in chlorophyll. So, we use this as one of multiple lines of evidence in evaluating modeled lower trophic level productivity. Remote sensed turbidity in formazin nephelometric unit (FNU) is generated from OLCI data via the algorithm of Nechad et al., 2010, with accessory algorithms such as for atmospheric correction, cloud masking validated based on comparison to USGS moored turbidity sonde records and the high-frequency mapping data (Killam et al., 2025). Unlike chlorophyll, we are more confident in the algorithm that determines the turbidity field, which is used as inputs for the light extinction coefficient, K_d . Remotely sensed data are available at approximately 1-2 day intervals over the entire bay, and depend on atmospheric conditions for complete coverage.

4.2.4 Sediment Flux Measurements

2021 presents a unique opportunity to compare modeled fluxes directly to sediment flux measurements from LSB and South Bay. In 2020 and 2021, sediment cores were collected in LSB and South Bay at channel and transect locations (Figure A10). This study quantified SOD, denitrification and nutrient fluxes based on laboratory incubations of sediment cores collected seasonally. The details of the methods and results of the sediment flux measurements are published in Czapla et al., 2025.

4.3 Approach for Evaluation of Model Performance

Estuarine circulation is characterized by strong tidal flow, rough topography, sharp gradients, and rapidly changing conditions (MacCready et al., 2010). This makes the comparison to model outputs to data through traditional goodness of fit measures challenging. Therefore, we employ several graphical/visual and statistical methods of evaluation across different dataset to assess model performance.

4.3.1 Graphical Evaluations

Model performance was assessed visually through time series plots of model and observations at the USGS Peterson Cruise and DWR discrete data at 12 locations spanning stations across all sub-embayments. Except for chlorophyll-a and DO, 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. In locations where depth profiles of chlorophyll-a and DO are collected, profile comparisons are conducted along the channel. Additionally, for DO, where comparisons are also made as AOU, which represents the deviation from saturation.

A comparison of the surface chlorophyll-a concentration from the OLCI remote sensing product and the model data output were used to assess performance over space. The model surface chlorophyll is output at 00:00 UTC (16:00 local standard time). The Sentinel-3 satellite operates on a 27-day cycle allowing for flyovers approximately once every 1-2 days. This makes direct comparison to satellite data difficult as the model output is quite large and flyover times are irregular throughout the year.

The high frequency mapping cruise in-situ data is additionally compared to modeled instantaneous surface chlorophyll, nitrate, and oxygen. The cruise across south bay starts at Emeryville and travels in a zigzag pattern across south bay until the Dumbarton Bridge, and moves northward. The entire length of the cruise is approximately 8 hours, from 6am to 2pm. The high frequency data is then compared to daily averages of surface chlorophyll, nitrogen, and dissolved oxygen using direct visual comparison and statistical analysis.

Key model state variables and derived model outputs (DIN and GPP) 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. DIN is calculated as the sum of ammonium and nitrate in model outputs, and as sum of ammonium, nitrite and nitrate for the measurements (we note that nitrite is generally negligible in the measurements due to the transient nature of nitrite under ambient nitrification, and while the model does not explicitly track nitrite, nitrate effectively accounts for both within the model). Gross primary production compared from model outputs are compared to estimates derived from observed chlorophyll-a concentrations and light extinction at the USGS Peterson Cruise locations. The GPP calculation for the data follows Jassby et al., 2002 and Cloern, 2007, and estimates the total GPP given observed chlorophyll concentration, surface photosynthetically active radiation (PAR), and the light attenuation coefficient. To calculate the observed K_d , we use the same formulation from Equation 1. For the model, GPP is derived from the model-predicted growth and respiration fluxes for phytoplankton. The model provides outputs for net primary production and respiration, from which the GPP is derived.

Scatter plots are used to assess correlation between observed and modeled values, and to infer any systematic bias.

4.3.2 Statistical Metrics

We also computed more traditional statistical goodness-of-fit metrics to assess model performance. The metrics used include mean error (ME), mean absolute error (MAE), root mean square error (RMSE), the coefficient of determination (R^2), skill score (SS), correlation coefficient (CC), bias (B), and standard deviation (STD). The hydrodynamics are tidally averaged and the constituent transport within DWAQ are not subjected to diurnal tides within the aggregated grid model (see King et al., 2020 for a detailed discussion on how the effects of tidal dispersion on constituent transport was incorporated even though diurnal tides are not included in the aggregated grid model). Therefore, the daily average is used for the aggregated grid model error statistics. The hourly averages are used for the full resolution model simulations to match the model output frequency at these locations by time when calculating error statistics.

4.3.3 Selection of Best Parameter Set(s)

The graphical and statistical approaches described above are used on a parameter space exploration through LHS. The LHS approach for the initial phase of model calibration is performed using the aggregated grid model. LHS creates a random combination of parameter sets within realistic parameter ranges such that a large multi-dimensional parameter space can be queried using a sparse, but space-filling parameter set (McKay et al., 1979). This allows for efficient exploration of the parameter space and identify parameter combinations that are sensitive. In this case, we refer to parameters listed in Table 7.

To calibrate, we first select parameters from each focal process discussed in Section 4.1 and explore a representative set of parameter combinations relevant to that process (Table 7 and Table 8). For each parameter combination, model outputs are compared to USGS and DWR data in the channel, and to HF mooring data at the SFEI mooring site and at the Dumbarton Bridge based on MAE and SS. For each focal process, the best 5-10% of the model runs in both metrics are chosen for exploration in a second phase. This method is then iterated over each process until the model output, channel, and mooring data is reasonably well calibrated. During this first round of calibration, we additionally identify parameters to which model outputs are sensitive for further refinement.

The second phase of this model calibration then combines the best combinations of sensitive parameters from the individual process-level explorations over groups of processes. For the most sensitive parameters, a second round of combined LHS exploration is conducted over a narrower range, reflected in the top 5-10% of the each sensitive parameter, to further improve the calibration.

Finally, an exploration around the local minima of sensitive parameters is conducted to narrow down a set of parameters that serve as the “reference” model configuration presented in Section 5.1. In this phase of model calibration, we use a combination of both the aggregated grid model and the full resolution model. This resulted in the “calibrated” parameter set indicated in Tables 7 and 8. The results of the aggregated and full resolution model performance are discussed in section 5.

4.3.4 Control Volume Analysis

Sub-embayment level mass balances can provide insights into the fate and transport of nutrients in SFB. These mass budgets were generated over polygons that coincided with the aggregated model grid, and aggregated into larger spatial groups (sub-subembayment to subembayment - Figure 2) to infer key transport and transformation drivers for model state variables and derived parameters (such as DIN and GPP). This provides the information necessary to assess relative importance of the nutrient cycling and lower trophic-level processes and allows the exploration of various hypotheses related to key process drivers of nutrient cycling and primary production within the Bay.

The control volume quantities include a variety of modeled processes. All the internal model fluxes (eg. denitrification rate, nutrient uptake rate, phytoplankton mortality, etc.), the total mass, and the transport between control volumes, are calculated. These values are used both to help inform model calibration efforts and to test hypotheses. For a full description of the control volume setup and analysis methods see King et al., 2020, Appendix 1A.

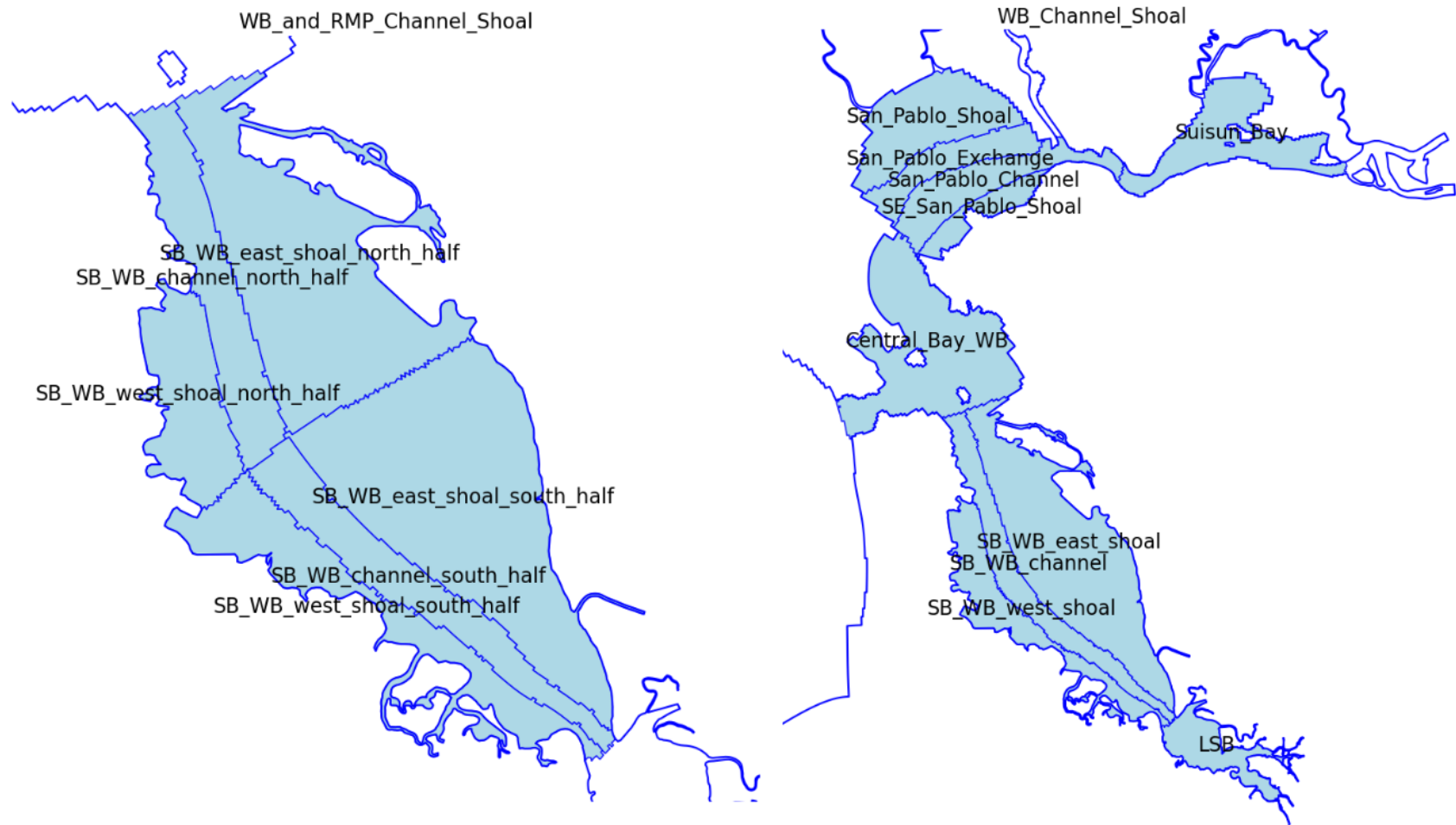


Figure 2: This figure shows an example of a control volume grouping within the Bay. The subembayments here are as defined by the water board in the south bay (left) and the entire bay (right). These control volumes are used as possible regions of analysis in the model.

5 Model Calibration Results and Discussion

5.1 Reference Model Comparison to Data

While we have a large number of simulated variables in the SFB-BGC model, we focus the discussion on phytoplankton (diatoms and greens), nutrients (primarily nitrogen and ammonia, and secondarily phosphorus and silica), and DO. Throughout the discussion, we refer to this version of the “calibrated” parameters in Tables 7 and 8 as v2026. In all of our model comparisons, we show, when available, a comparison to the Full Resolution version of the model and the Aggregated grid model. The full resolution model is run over the entire water year, from 1 August 2020 to 1 October 2021.

5.1.1 Phytoplankton and Primary Production

Comparisons to Channel Data

Two phytoplankton groups are represented in the model: “diatoms” representing the spring bloom group and “greens” representing the summer group; neither references a specific genera or species. Modeled biomass are in units of carbon, but comparisons to data are presented as chlorophyll concentration. The latter is derived using a fixed conversion factor of 44.4 grams of carbon per gram chlorophyll (or equivalently $22.5 \mu\text{g}$ chlorophyll per mg carbon) for both phytoplankton groups represented in the model. Results are presented for total chlorophyll (calculated as the total diatoms and greens biomass at any given location). The diatoms and greens concentration time series and spatial fractions simulated within the model are presented in Appendix A and relevant figures are referenced throughout the discussion.

The total chlorophyll concentration along the deep channel in SF Bay is compared to the USGS Peterson Cruise in Figure 3. The total chlorophyll concentration is the sum of the chlorophyll from the modeled diatom and green concentrations (see Figures A11 and A12 for components). Scatter plots showing modeled to observed chlorophyll are shown for the USGS transect locations in Figure 4. Additionally, we also present sensitivity of the assumed carbon to chlorophyll ratio relative to a typical range of 14-50 that is often used for nutrient-rich but light-limited regions (Sathyendranath et al., 2009; Jakobsen et al., 2016). The modeled total phytoplankton biomass converted to chlorophyll using this range of carbon to chlorophyll ratio are shown in Appendix Figure A13.

The modeled total chlorophyll generally captures the observed seasonal patterns of low biomass in winter, a spring bloom starting in late February to March, peaking in May, with subtle temporal shifts in the peak timing from South to North Bay, and a lower but steady biomass level over mid to late summer. The magnitude of the peaks, including the spatial differences in the magnitudes (higher in LSB and lower values progressively north), are captured reasonably well in the model. Toward the end of the bloom event, the model does not capture the reduction in chlorophyll concentration that is seen along the channel, with the model exhibiting a sustained positive bias in summertime chlorophyll concentration. The model also exhibits generally higher chlorophyll levels in San Pablo and Suisun Bay. We note that the model does not include clams, which may be a key driver in keeping phytoplankton levels low within San Pablo and Suisun Bays. The spatially constant maximum quadratic mortality rate (varies only based on temperature), while successful overall in limiting excessive phytoplankton levels, appears to be limited in its ability to reproduce some of these spatial differences in top-down control.

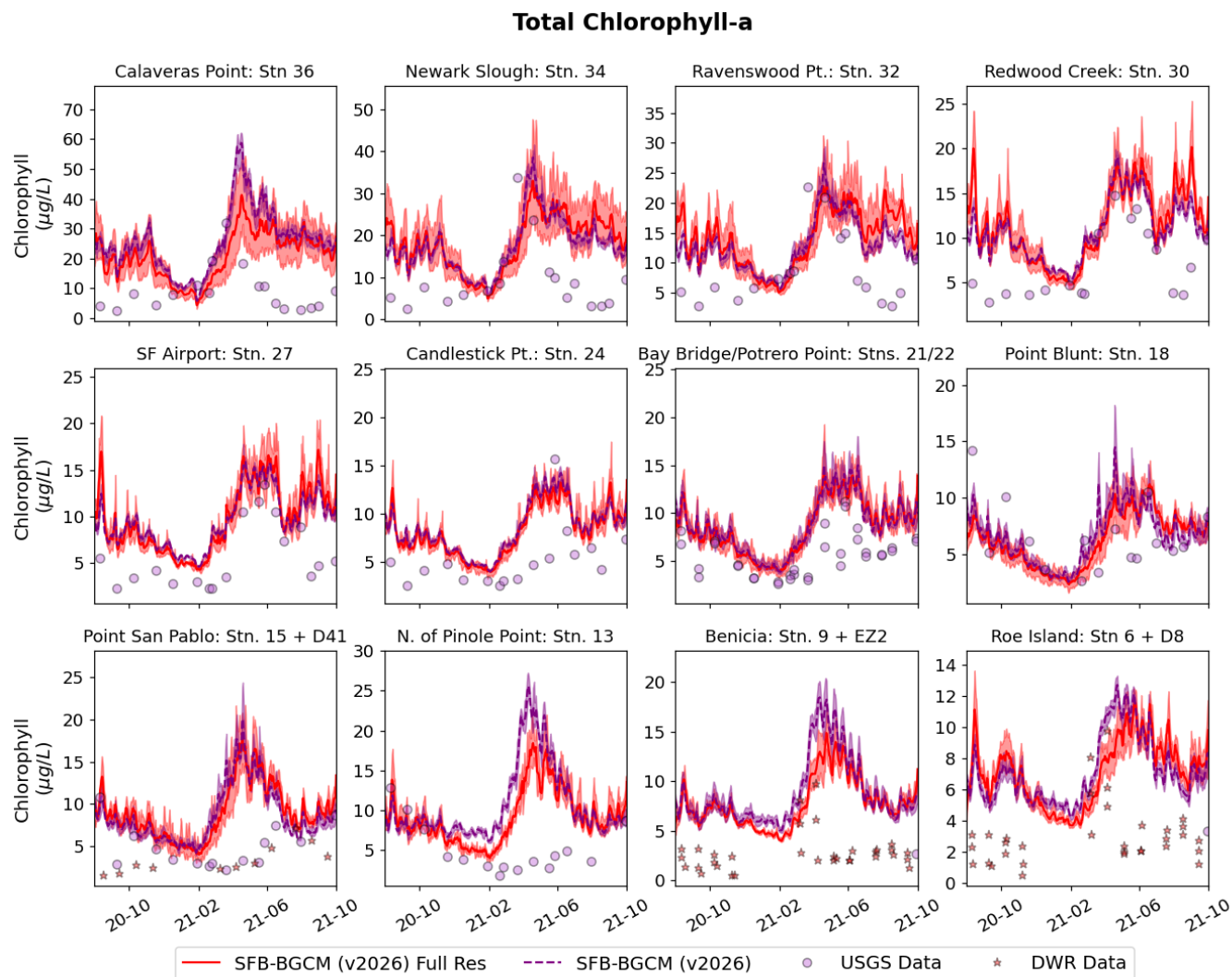


Figure 3: Comparison of daily averaged surface chlorophyll from the full resolution (red) and aggregated grid (purple dashed) model output with observations of surface chlorophyll from the USGS Peterson Cruise Data (circles) and the DWR Environmental Monitoring Program (stars) at 12 long-term monitoring stations, one shown in each subfigure organized from LSB (top left) to Suisun Bay (bottom right). The shaded areas show the daily minimum and maximum values.

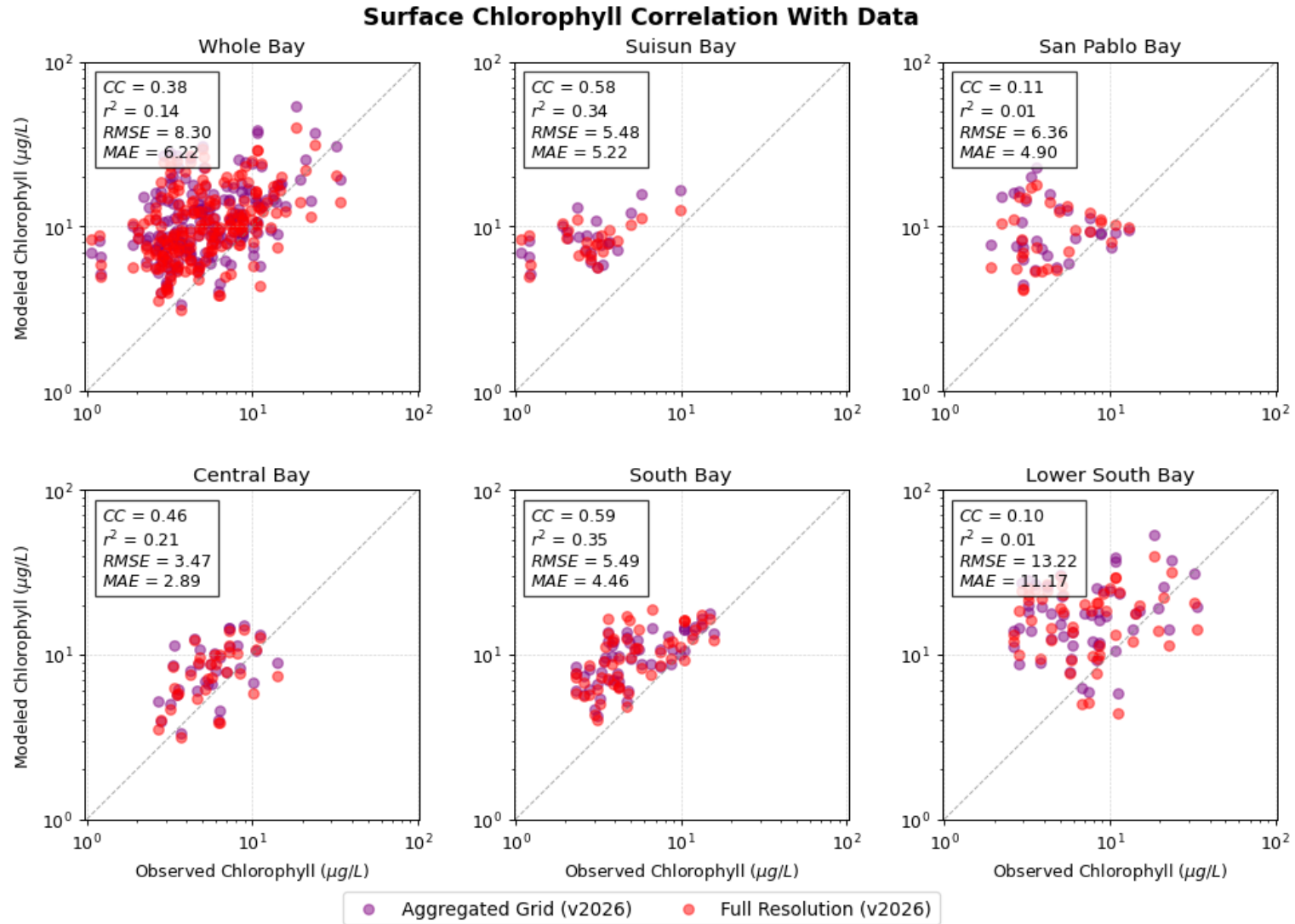


Figure 4: Log-log scatter plots of USGS and DWR observations compared to modeled chlorophyll concentration for the Whole Bay and major subembayments, listed left to right, top to bottom: Suisun Bay, San Pablo Bay, Central Bay, South Bay, and LSB. The modeled chlorophyll for the Full Resolution (red) and Aggregated Grid (purple) are shown along with the 1:1 line (black, dashed). The Pearson Correlation coefficient (CC), Pearson R^2 , root mean square error (RMSE), and mean absolute error (MAE) is shown in the upper left legends in each subfigure for the full resolution model-data comparison only.

The GPP along the channel is shown in Figure 6. In South Bay and Central Bay (stations 32, 30, 27, 24, 22, 18), our modeled GPP agrees well with the observed estimates of GPP during the springtime bloom and the sustained summertime productivity. In LSB (station 36, 34), the springtime bloom is captured, but the model overestimates GPP in the summertime. In San Pablo Bay and Suisun Bays, the model overestimates the GPP.

Comparisons to High Frequency Data

A comparison of the modeled chlorophyll concentration at the high frequency mooring locations is shown in Figure 7. The high frequency measurements exhibit substantial variability at Dumbarton Bridge, San Leandro Marina and the Hayward Mooring, likely reflective of conditions in the outer margins, in the inter-tidal marshes and shallow regions of the Bay. The east shoal SFEI mooring and the San Mateo Bridge, both in sub-tidal/channel areas, show significantly lower variability overall in the chlorophyll concentrations. The model captured the overall seasonal patterns, but did not exhibit the same level of variability at some of these locations. Specifically, chlorophyll concentrations at the Dumbarton Bridge show that the model captures the second spring bloom in April and several fall and summer blooms, but underestimates the peak of the springtime bloom in the model (approximately mid- to late-March). At the San Mateo Bridge, the model is biased high over the entire year by 5-15 $\mu\text{g/L}$. However, the model predictions are generally consistent with the range (5 - 15 $\mu\text{g/L}$) indicated in the USGS Peterson Cruise measurements (at Station 30 and Station 27 - see Figure 3), although the model overestimates the winter data even relative to the USGS Peterson Cruise measurements. Similarly, at the SFEI shoal mooring station the modeled chlorophyll exhibits a high bias ranging again from 5-15 $\mu\text{g/L}$ in winter months but an apparent spring bloom in May captured in both model and USGS Peterson Cruise measurements in May is not reflected in the SFEI shoal mooring station measurements. This suggests that the high frequency measurements at the San Mateo Bridge, which is located approximately mid-way between stations 30 and 27, and the SFEI shoal mooring station which is located to the east of Station 27, may be biased low or may not be reflective of the ambient surface chlorophyll concentrations during these periods where an alternative line of evidence is contrary to these observations. At the San Leandro Marina, the predicted chlorophyll concentration is nearly constant year-round at approximately 20 $\mu\text{g/L}$. In March 2021, the observations show a sustained month-long spring bloom that is not captured in the model. At the Hayward Mooring, there are considerable oscillations in the chlorophyll concentration that appear to be tracking the spring-neap cycles: the modeled chlorophyll concentrations fall within the upper amplitudes of these oscillations, but in general, chlorophyll is over-predicted at this location.

The GPP at the mooring locations is shown in Figure 8. Comparisons of GPP estimates in the spring and summertime at all of the mooring locations show reasonable agreement between modeled and observed GPP, considering the large variability in the GPP estimates derived from the highly variable and oscillatory high frequency chlorophyll measurements.

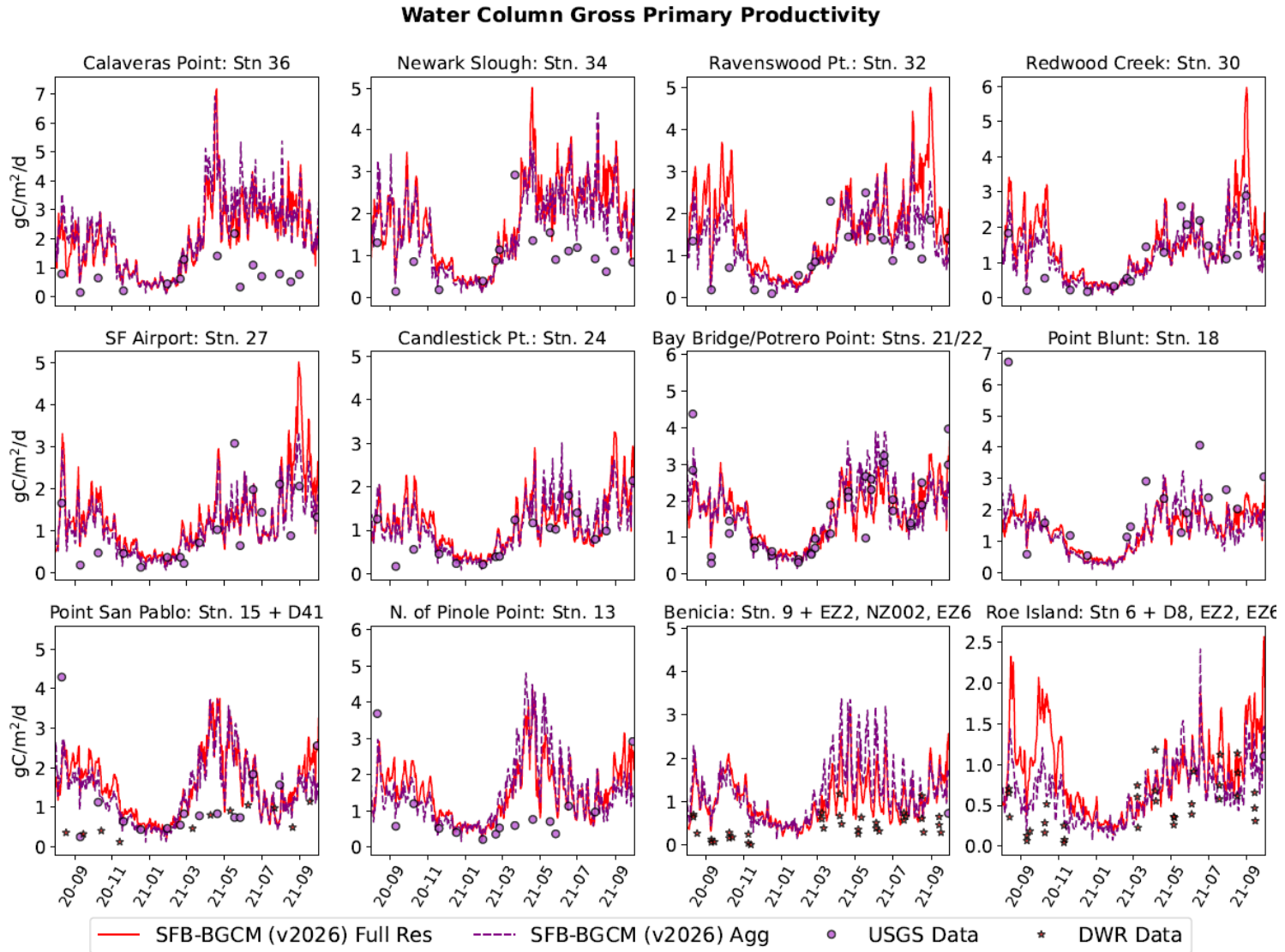


Figure 5: Comparison of daily averaged surface gross primary productivity for the full resolution (red) and aggregated grid (purple dashed) models with the observations from the USGS Peterson cruise data (circles) and DWR Environmental Monitoring Program (stars) at 12 long-term monitoring stations, one shown in each subfigure organized from LSB (top left) to Suisun Bay (bottom right).

Gross Primary Productivity Correlation With Data

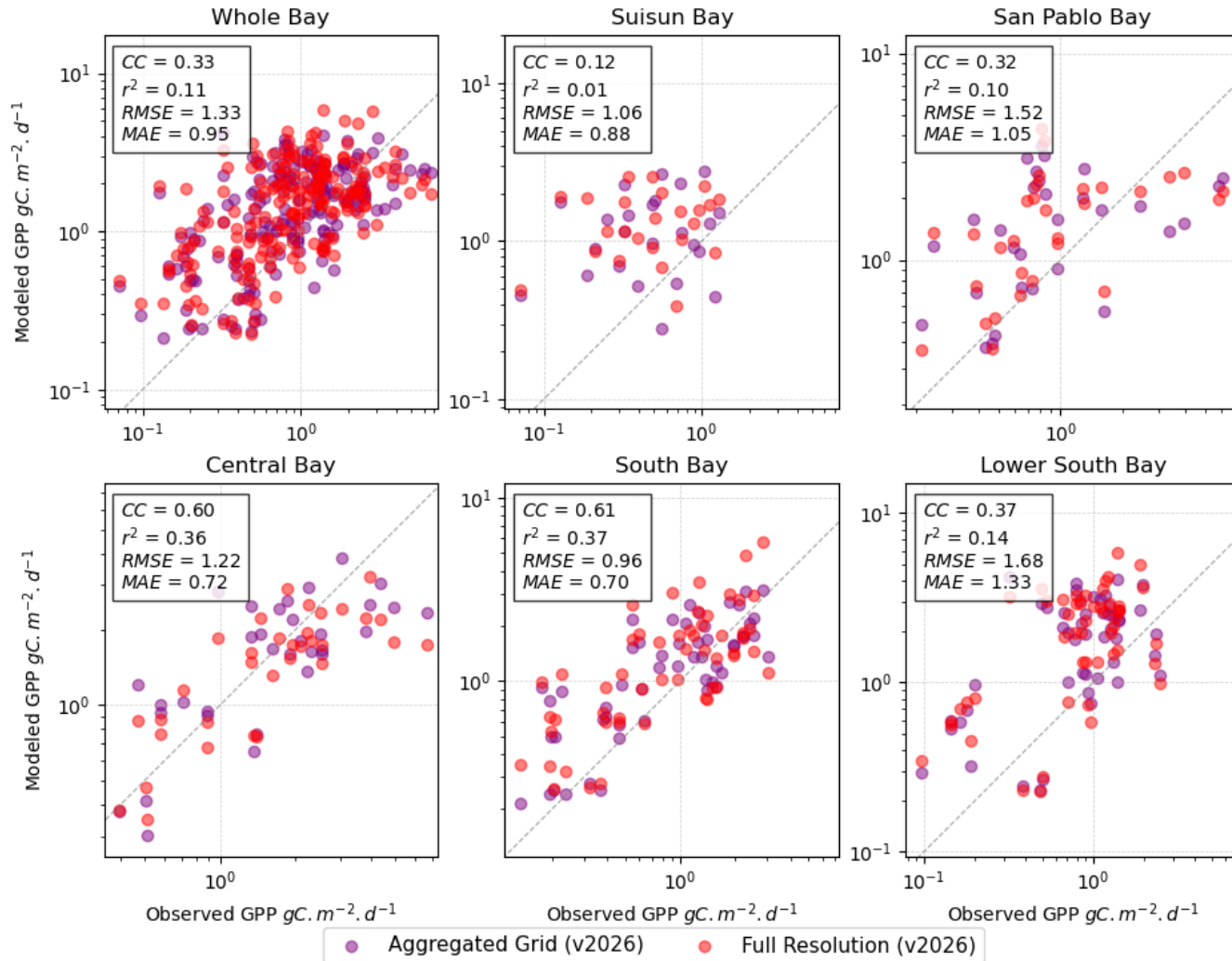


Figure 6: Log-log scatter plots of USGS and DWR observations compared to modeled gross primary productivity (GPP) for the Whole Bay and major subembayments, listed left to right, top to bottom: Suisun Bay, San Pablo Bay, Central Bay, South Bay, and LSB. The modeled GPP for the Full Resolution (red) and Aggregated Grid (purple) are shown along with the 1:1 line (black, dashed). The Pearson Correlation coefficient (CC), Pearson R^2 , root mean square error (RMSE), and mean absolute error (MAE) is shown in the upper left legends in each subfigure for the full resolution model-data comparison only..

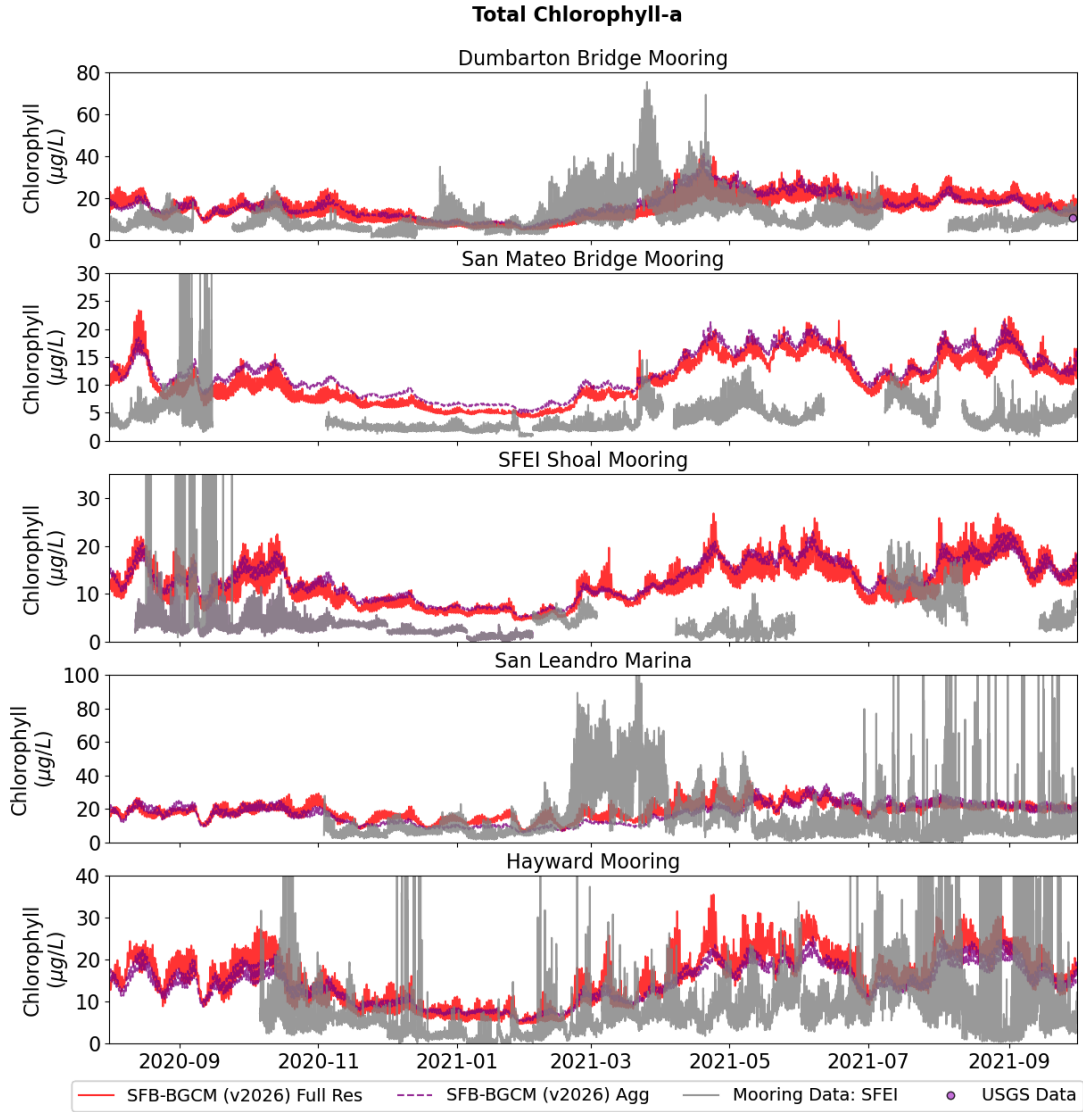


Figure 7: Comparison of surface chlorophyll concentrations at the mooring sites for the full resolution (red) and aggregated grid (purple dashed) models with high frequency observations from SFEI (grey) and USGS (light purple) moorings. A single point measurement collected on the USGS Peterson Cruise in late 2021 at the Dumbarton Bridge is marked with a circle. Model outputs show the full time series.

Water Column Gross Primary Productivity

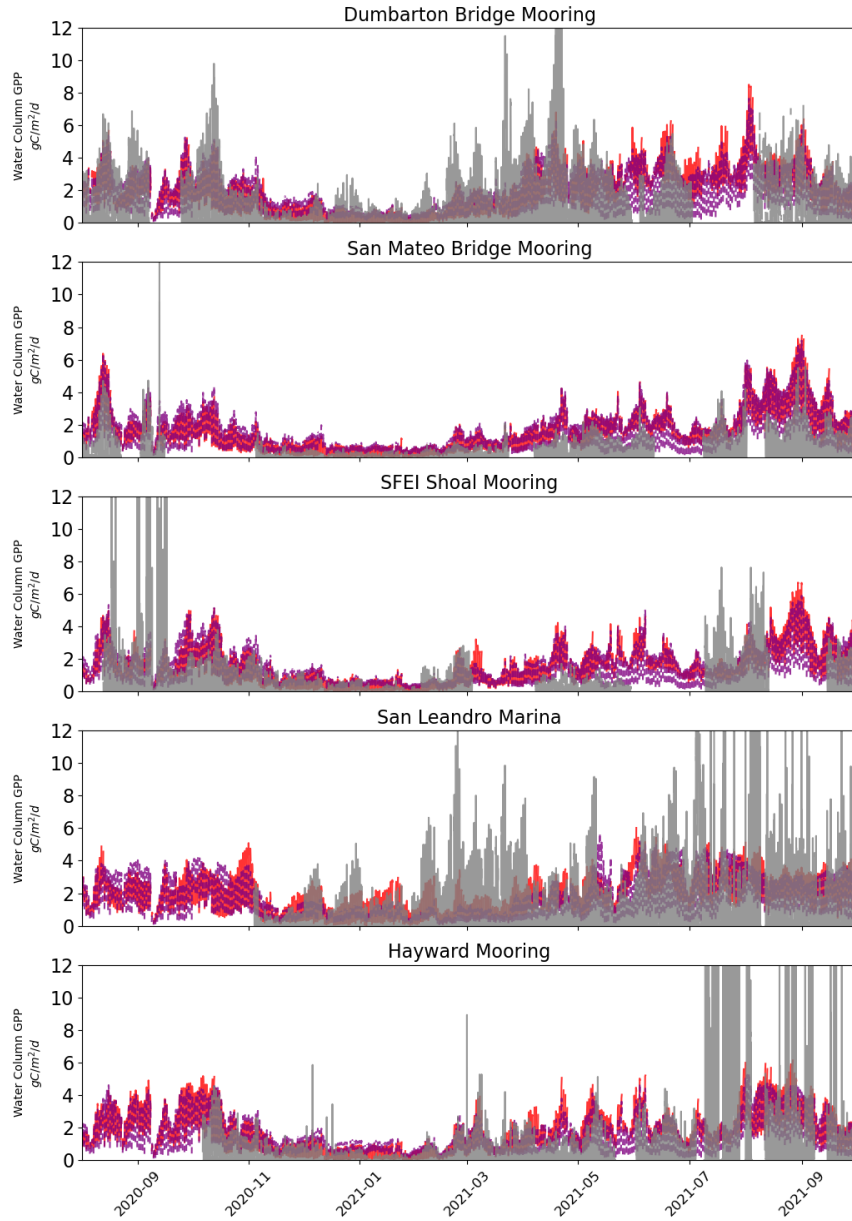


Figure 8: Comparison of surface gross primary productivity (GPP) at the mooring sites for the full resolution (red) and aggregated grid (purple dashed) model outputs with high frequency observations from SFEI (grey) and USGS (light purple) moorings. GPP is shown during daylight hours only.

Comparison to Mapping Cruise Data

The chlorophyll concentration maps are compared to the mapping cruise data in Figure 10. The mapping data comparison shows five days between March and September 2021 covering the extent of the mapping cruises in south bay and central bay. We note that our model output prints out the instantaneous model state at 16:00 local time, whereas the cruises occur between 06:00 and 14:00 hours local time. The timescale for phytoplankton growth and mortality occurs over the course of days to weeks. Therefore, comparing an instantaneous model snapshot relative to a whole-day cruise does not introduce any undue artifacts in comparability between underlying process contributing to modeled and observed biomass production (i.e., growth and mortality) beyond tidal-timescale-level (intra-day) spatial differences in transport. In general, the mapping data shows lower chlorophyll levels relative to the model. However, there are distinct spatial patterns reflected in the mapping data that generally comports with the model simulations. The mapping cruises cannot extend very close to the shore because of the water depth - but towards the eastern extent of the mapping cruises there is consistently higher chlorophyll levels compared to those on the deeper sub-tidal portions. For instance, in March, May and August 2021, the mapping data shows higher chlorophyll concentration near the Oakland Airport, extending down to Hayward. The higher concentrations observed in the mapping cruise along the shoreline are consistent with the large oscillations noted earlier in the high frequency chlorophyll data in Hayward and San Leandro moorings, but not at the SFEI shoal mooring (see Figure 7). The model shows elevated chlorophyll concentration along the eastern shoreline including the Oakland Airport area and generally gets lower westward, mirroring the spatial pattern captured in the mapping cruise, but the model overestimates the chlorophyll concentration away from the eastern shoreline moving towards the channel relative to the mapping cruise data. Finally in September 2021, the modeled chlorophyll is overestimated nearly bay-wide.

Comparisons to Remotely Sensed Data

Comparisons of monthly and seasonal averages between modeled and satellite surface chlorophyll concentrations are shown in Figures A14 and A15. For seasonal averages we use December-January-February, March-April-May, and June-July-August to assess winter, spring and summer biomass, respectively. The months of August through October of 2020 are not presented because they coincide with the model spinup period.

As discussed in Section 4.2.3, the remotely sensed chlorophyll data are primarily used to evaluate the model's ability to reproduce spatial patterns rather than the absolute magnitude indicated in the data due to the inherent uncertainty in the algorithms used to infer remotely sensed chlorophyll concentrations. With this caveat in mind the remotely sensed data shows higher chlorophyll concentrations: in LSB throughout May - Sep; along the east shoal of South Bay in spring (more specifically in March); along east shoal of Central Bay near Emeryville and Berkeley Marina in March through July; from March through September along the north shoal of San Pablo Bay, reaching a peak in June and July; and, in Montezuma Slough in March and April (Figure A14, middle row). The satellite data are generally low towards the channel. The model reproduces some of these spatial patterns, but as with other lines of evidence presented previously, the model overestimates the chlorophyll concentrations in the South Bay east shoals away from the shoreline, near and north of the Berkeley marina in Central Bay, and in the north shoal of San Pablo Bay (in March through June) (Figure A14, bottom row).

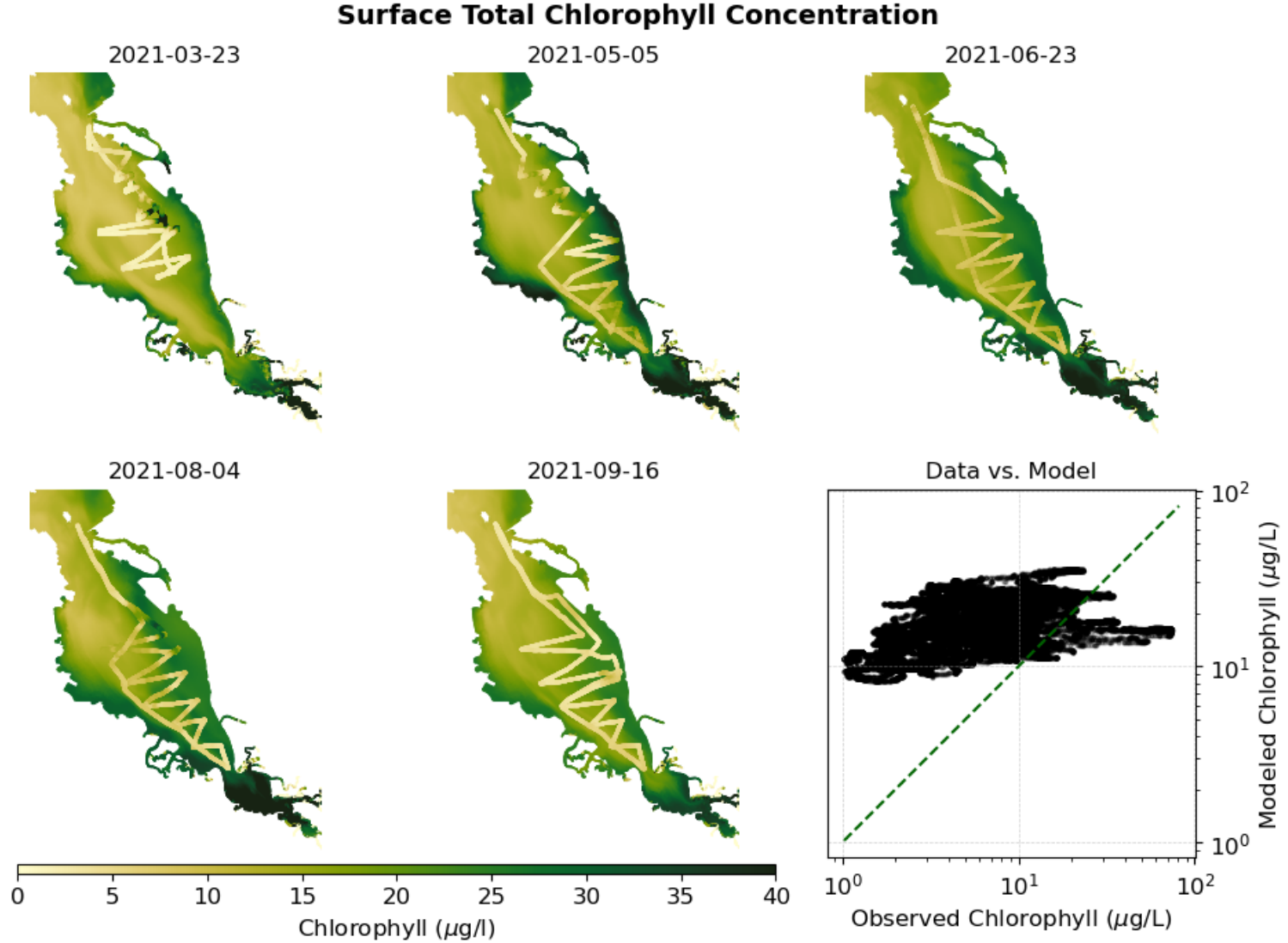


Figure 9: Surface map plots of modeled chlorophyll and the mapping cruise observations (first five subfigures). The instantaneous modeled surface chlorophyll concentration from 16:00 is shown on each date. The surface chlorophyll concentration from the mapping cruise is overlaid on top of the model data using the same colorbar. Scatter plot (bottom right) compares the mapping cruise observations and modeled chlorophyll. The dashed greenline is the 1:1 line.

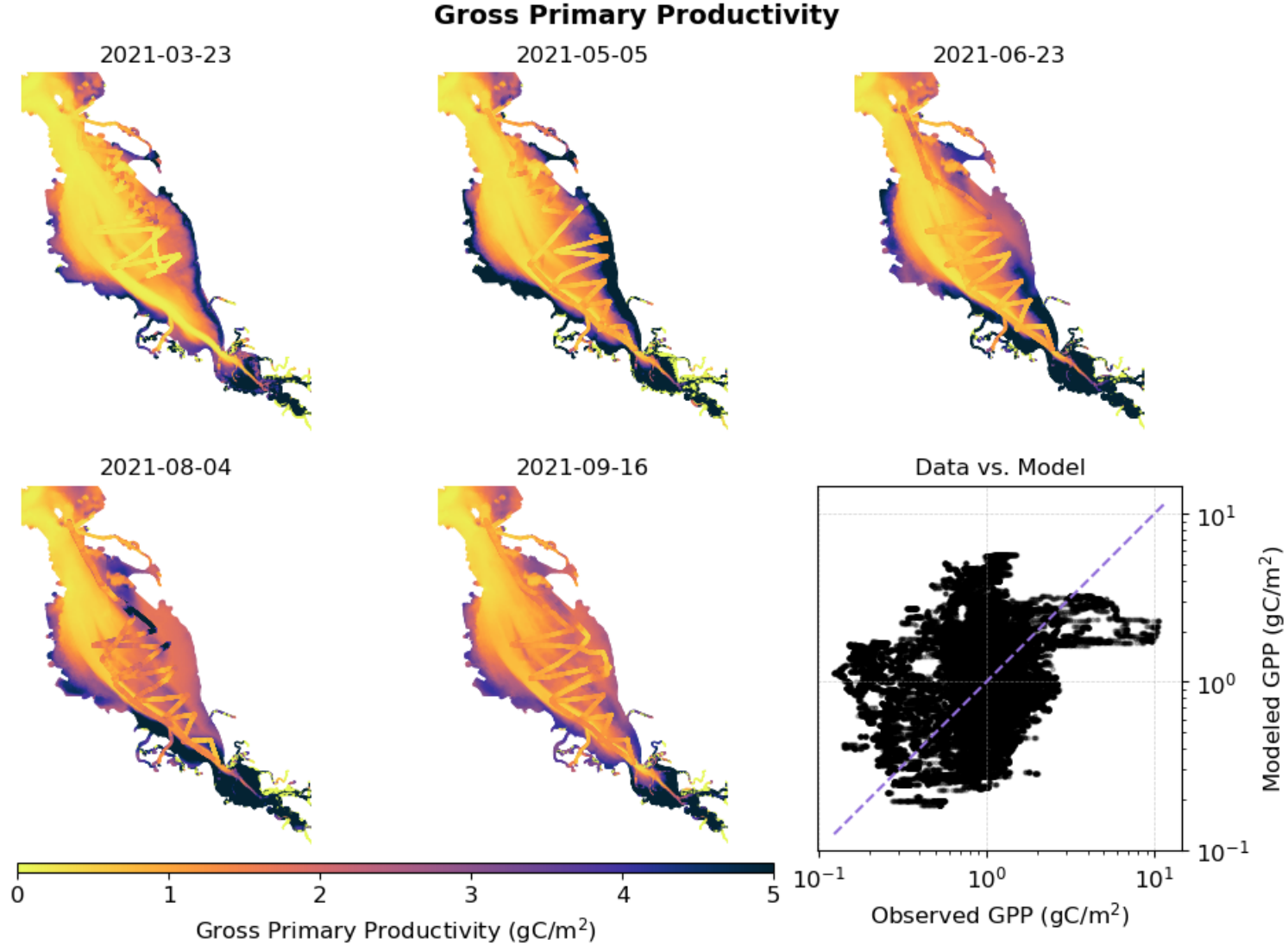


Figure 10: Map plots of modeled gross primary productivity (GPP) and estimates from the mapping cruise observations (first five subfigures). The instantaneous modeled GPP concentration from 16:00 is shown on each date. The GPP from the mapping cruise is calculated following Cloern, 2007 and is overlaid on top of the model data using the same colorbar. Scatter plot (bottom right) compares the mapping cruise observations and modeled chlorophyll. The dashed purple line is the 1:1 line.

The model consistently underestimates chlorophyll levels in LSB in all three seasons (Dec - Aug), but most notably in winter and spring (figure A15, bottom). The Dumbarton Bridge high frequency data similarly shows an increase in chlorophyll in February and March that the model did not capture adequately (Figure 7). This suggests that primary production activity in the LSB marshes and salt ponds are probably not adequately captured in the model because of the coarser spatial resolution. As indicated in our previous reports, this spatial simplification is necessary to keep model run times reasonable (King et al., 2020, King et al., 2023), but a higher resolution LSB model is part of the overall NMS modeling workplan (SFEI 2020 and 2025), and may be resumed in subsequent years depending on other science and management priorities.

Log-normally scaled empirical cumulative distribution function (CDF) and probability density functions (PDF) comparing the modeled and satellite-derived surface chlorophyll concentrations are shown in Figures A16 and A17, respectively. The CDF and PDF show a weakly bimodal distribution in the remote sensing data, with the bulk of the observations centered around 8-9 $\mu\text{g/L}$ and additional sustained high chlorophyll observations between 10-100 $\mu\text{g/L}$ that are not fully captured in the model. The CDF of the model shows that the bulk of the modeled chlorophyll concentration is centered around 10 $\mu\text{g/L}$ with a wider peak. There is a sharp drop off of modeled concentrations less than 100 $\mu\text{g/L}$ (around 30 $\mu\text{g/L}$), whereas the remote sensing product shows a small secondary peak around 100 $\mu\text{g/L}$.

5.1.2 Dissolved Inorganic Nitrogen, Nitrate, and Ammonia

Comparisons to Channel Data

DIN concentration along the channel is compared to the USGS Peterson Cruise data in Figure 11, and scatter plots showing model performance relative to data are shown in Figure 12. Across the USGS Peterson Cruise data, the model aligns well with observations for both the aggregated and full resolution runs. Both capture the increase in DIN over late fall and winter, and both reproduce the subsequent DIN drawdown in spring and summer that are observed in the measurements. The aggregated grid nitrogen concentrations are generally lower when compared to the full resolution. This is likely a consequence of the averaging that occurs when going from a finely resolved high resolution model to a coarse grid. Additionally, the aggregated grid model is tidally averaged, and therefore does not show the same range of daily minimum and maximum values indicated in the full resolution model.

In LSB and South Bay south of station 27, the model shows the return of nitrogen into the late winter months, with good alignment during the spring-time bloom uptake. The full resolution model seems to slightly over-predict the nitrogen in the summertime. In Central Bay (Figure 11, middle row), both models under-predict nitrogen in the winter and wet months (Oct-Feb) by approximately 0.1 mg/L ($\sim 7.1 \text{ mmol/m}^3$). There is a similar pattern after the springtime bloom, likely because of continued uptake in the model relative to the data (see section 5.1.1 discussion on overestimation of summertime chlorophyll). The summertime under-predictions are more exacerbated in San Pablo likely due to the greater primary productivity in the model relative to observations. In Suisun Bay, our modeled nitrogen approximately follows the observed concentration.

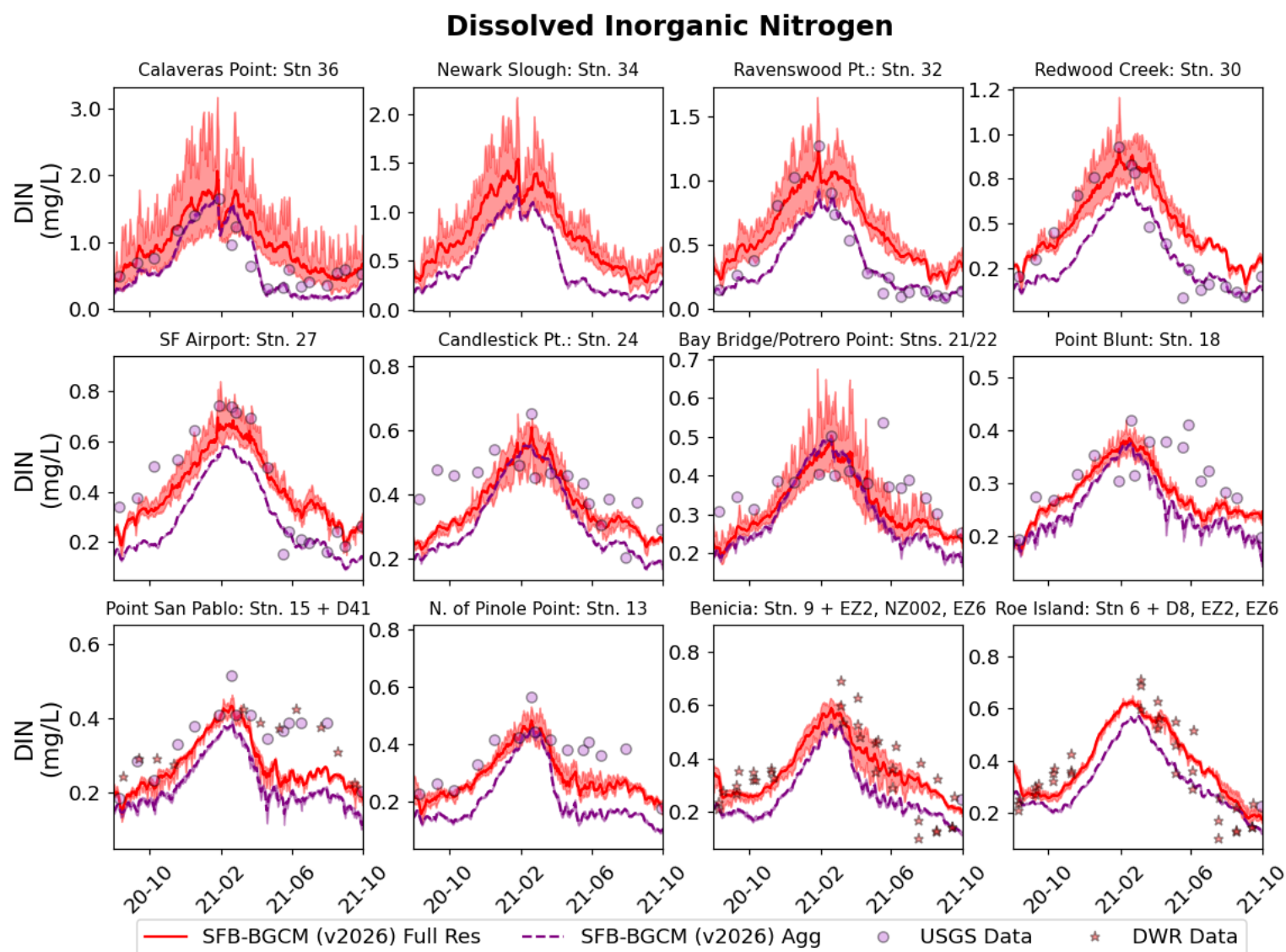


Figure 11: Comparison of daily averaged surface dissolved inorganic nitrogen from the full resolution (red) and aggregated grid (purple dashed) model output with surface observations from the USGS Peterson Cruise Data (circles) and the DWR Environmental Monitoring Program (stars) at 12 long-term monitoring stations, one shown in each subfigure organized from LSB (top left) to Suisun Bay (bottom right). The shaded areas show the daily minimum and maximum values.

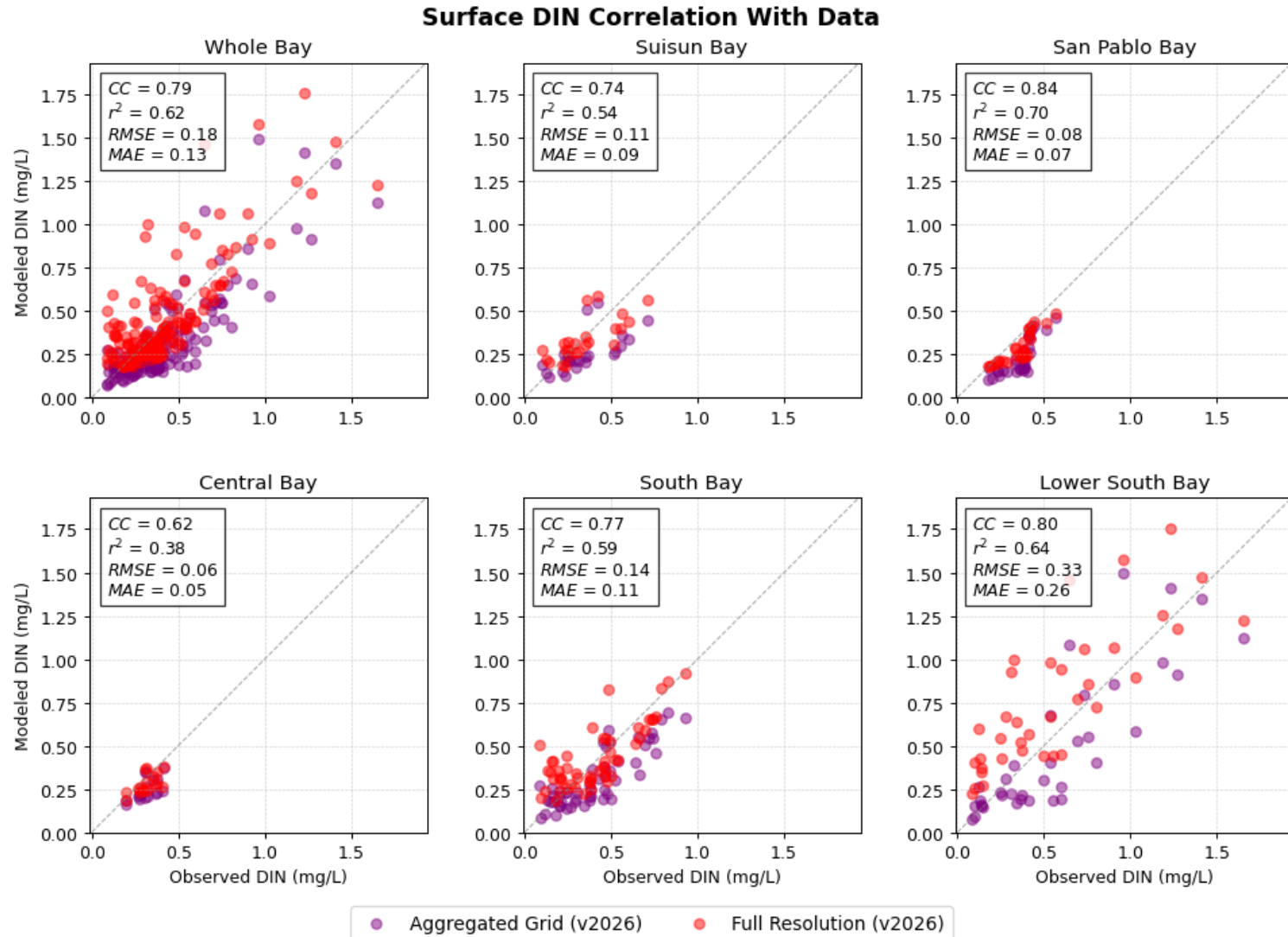


Figure 12: Scatter plots of USGS and DWR observations compared to modeled dissolved inorganic nitrogen (DIN) concentration for the Whole Bay and major subembayments, listed left to right, top to bottom: Suisun Bay, San Pablo Bay, Central Bay, South Bay, and LSB. The modeled DIN for the Full Resolution (red) and Aggregated Grid (purple) are shown along with the 1:1 line (black, dashed). The Pearson Correlation coefficient (CC), Pearson R^2 , root mean square error (RMSE), and mean absolute error (MAE) is shown in the upper left legends in each subfigure for the full resolution model-data comparison only.

Nitrate and ammonium are presented in Figures 13 and 14, respectively. Figure 13 shows that the model is well aligned with nitrate observations, with the exception of the observed data in San Pablo Bay (Stations 15 and 13), where the model underestimates the observed nitrate levels considerably. Otherwise, the bias between modeled and observed data is small. Unlike nitrate, neither the data nor the model exhibit a consistent seasonal pattern for ammonium (Figure 14). Recognizing that ammonium concentrations for both data and model are of the order of 0.01 - 0.1 mg/L (except for some isolated values near 0.2 mg/L in LSB) and nitrate concentrations are at least 5 to 10 times higher throughout the Bay in both the measurements and modeled value, it is evident that nitrate is the dominant component of DIN (as would be expected in a well oxygenated estuarine system where nitrification is expected to be rapid). Overall, the model consistently underestimates ammonium by a factor of 2 - 3, which suggests that modeled nitrification may be somewhat faster than what occurs under ambient conditions or uptake depletes ammonium faster in the model relative to ambient conditions (which corresponds to the higher biomass levels discussed earlier). The modeled ammonium is likely influenced by both factors.

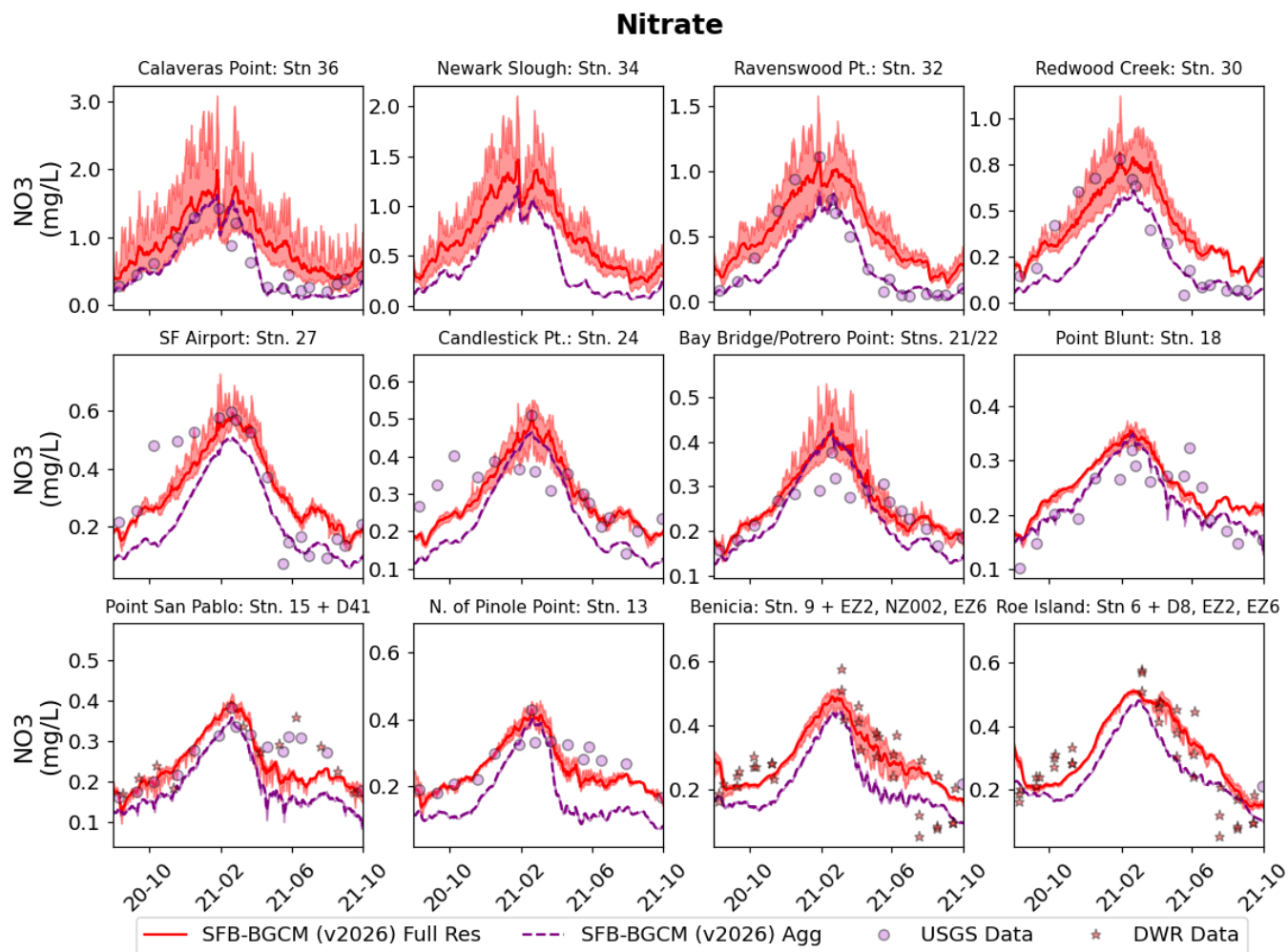


Figure 13: Comparison of nitrate from the full resolution (red) and aggregated grid (purple) model output with the USGS Peterson Cruise Data (circles) and the DWR Environmental Monitoring Program (stars). The shaded areas in show the daily minimum and maximum values showing variability due to tides.

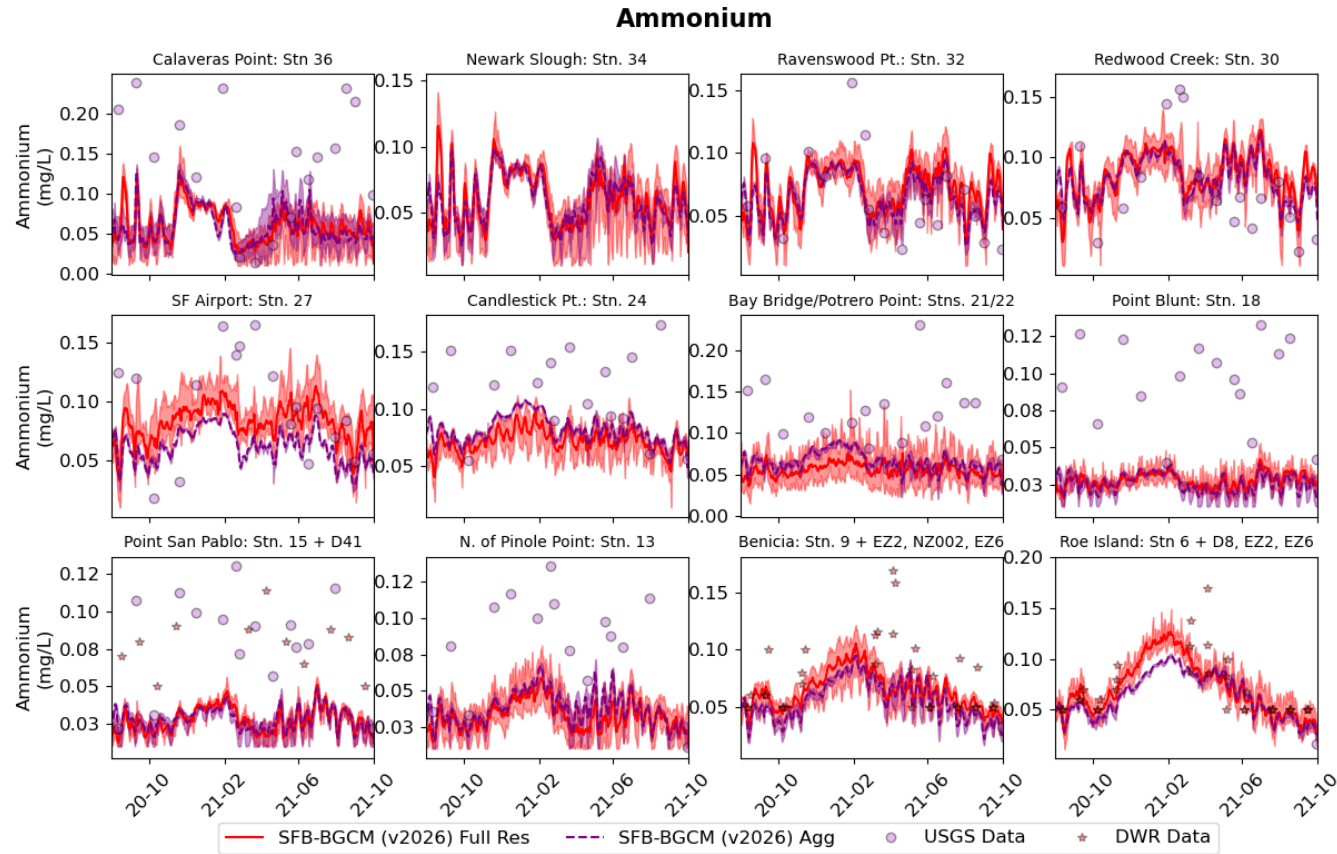


Figure 14: Comparison of ammonia from the full resolution (red) and aggregated grid (purple) model output with the USGS Peterson Cruise Data (circles) and the DWR Environmental Monitoring Program (stars). The shaded areas in show the daily minimum and maximum values showing variability due to tides.

Comparisons to High Frequency Data

A comparison of the modeled nitrate concentrations to observed high frequency mooring is shown in Figure 15. High frequency nitrate data are limited to Dumbarton Bridge and the SFEI Shoal mooring. Currently, data are not available at the San Mateo Bridge, the San Leandro Marina, and the Hayward Mooring due to sensor fouling. The data collected at Dumbarton Bridge and SFEI shoal mooring locations overlap over summer 2021, when both exhibit relatively lower levels of nitrate compared to other times of the year. Notably, the SFEI shoal mooring station has lower nitrate levels and exhibits a lower range in general compared to Dumbarton Bridge.

At both locations, the model matches the seasonal trend of winter high, and a drawdown in spring months leading to sustained lower levels in summer. At the Dumbarton Bridge, the modeled nitrate and observed data are well aligned through most of the simulation period where overlapping data are available. In spring months, the modeled drawdown of nitrate is slower compared to the high frequency measurements and the range in the high frequency measurements are generally higher particularly in the summer months. Both the spring over prediction in the model and smaller range in summer are likely related to the differences in the primary production. As discussed in Section 5.1.1 and as shown in Figure 7 the model does not match the first LSB spring bloom in March, and the fluctuations in modeled chlorophyll at Dumbarton Bridge is smaller than the high frequency observations. Additionally, it is possible that some of the limitations in the spatial resolution and representation of marshes and salt ponds in LSB, which can also influence DIN levels, may be contributing to these larger fluctuations in nitrate observations at Dumbarton Bridge. At the SFEI Shoal Mooring location, the model under-predicts the nitrate concentrations between August 2020 and March 2021 by approximately 0.2 mg/L. When viewed in light of the chlorophyll data in Figure 7 this is somewhat confounding considering that the high frequency chlorophyll data does not show any evidence of significant uptake locally over this period. When viewed in light of other lines evidence (satellite and mapping data), it is likely that the lower nitrate levels observed at the SFEI mooring reflect broader shoal-wide processes (of increased drawdown along the shoreline) rather than localized removal in the vicinity of the shoal mooring. During March 2021, the modeled nitrate is within the bounds of the data. In mid July to mid-August 2021, we slightly over-predict the nitrate by approximately 0.1 mg/L, and fall within the upper and lower limits of the observations after mid-August.

Comparisons to Mapping Data

Surface maps of modeled nitrate is compared to the mapping cruise data in Figure 16. The mapping cruise data is shown overlaid on snapshots of instantaneous model output on the date of the cruise. The nitrate data from the mapping cruises complement the chlorophyll observations presented in 10, exhibiting both a seasonal and spatial signal in nitrate levels. The highest nitrate levels occur in March, and get progressively lower through spring and summer. Additionally, the low nitrate levels near the east shoals that increase progressively towards the channel also correspond to increased productivity observed near the eastern shoreline. The modeled surface nitrogen levels are generally well aligned with the mapping cruise data both seasonally (higher levels in March and progressively lower levels through summer) and spatially (lower levels in the eastern shoreline/shoals and generally increasing towards the channel). The model over-estimates the nitrogen concentration towards the channel of South Bay, particularly in June and August, and also towards the eastern portions moving into Central Bay in the vicinity of Alameda. The scatter plot shows that the model predictions generally fall on either side of the 1:1 line, but with a notable cluster around 0.6 mg/L, which corresponds to the over predictions in June and August noted above.

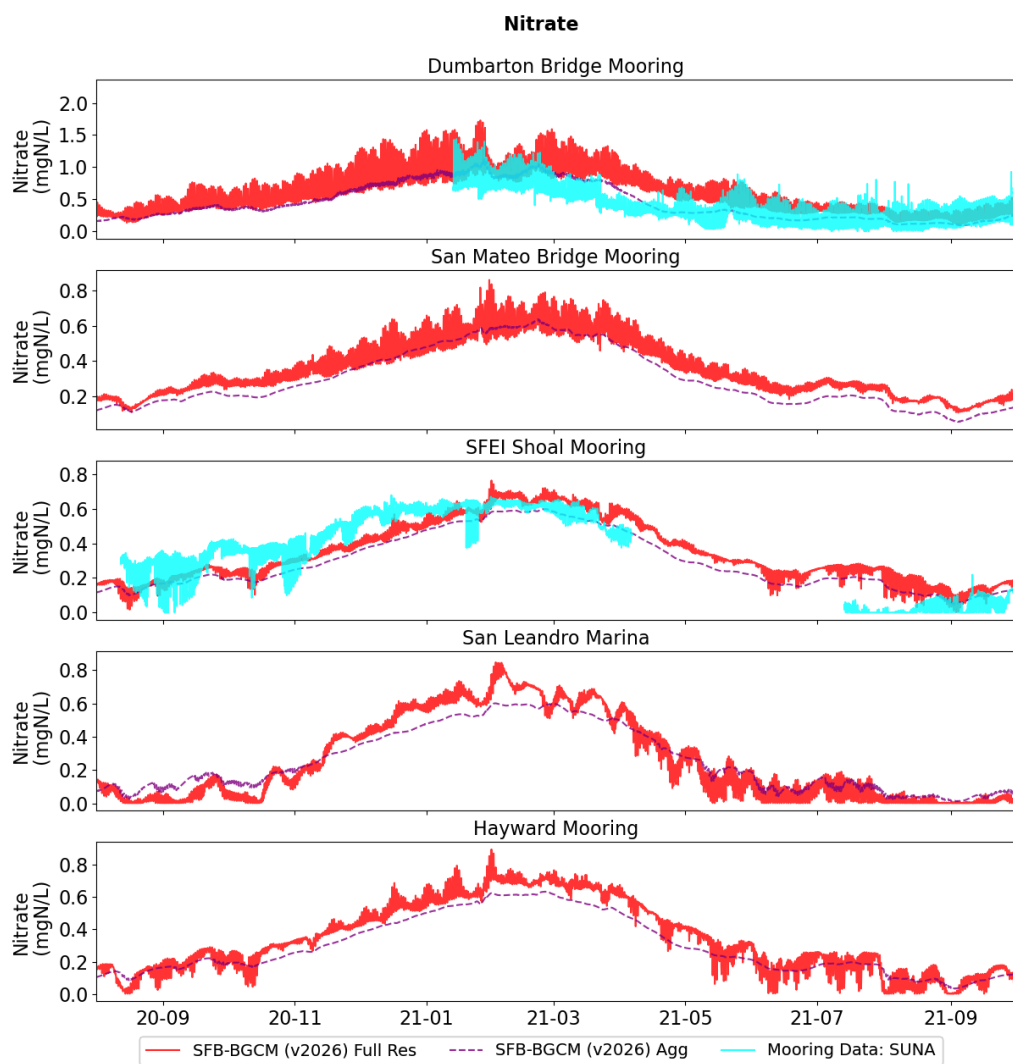


Figure 15: Comparison of surface nitrate at the mooring sites for the full resolution (red) and aggregated grid (purple dashed) model outputs with high frequency observations from SFEI SUNA nitrate sensor (cyan).

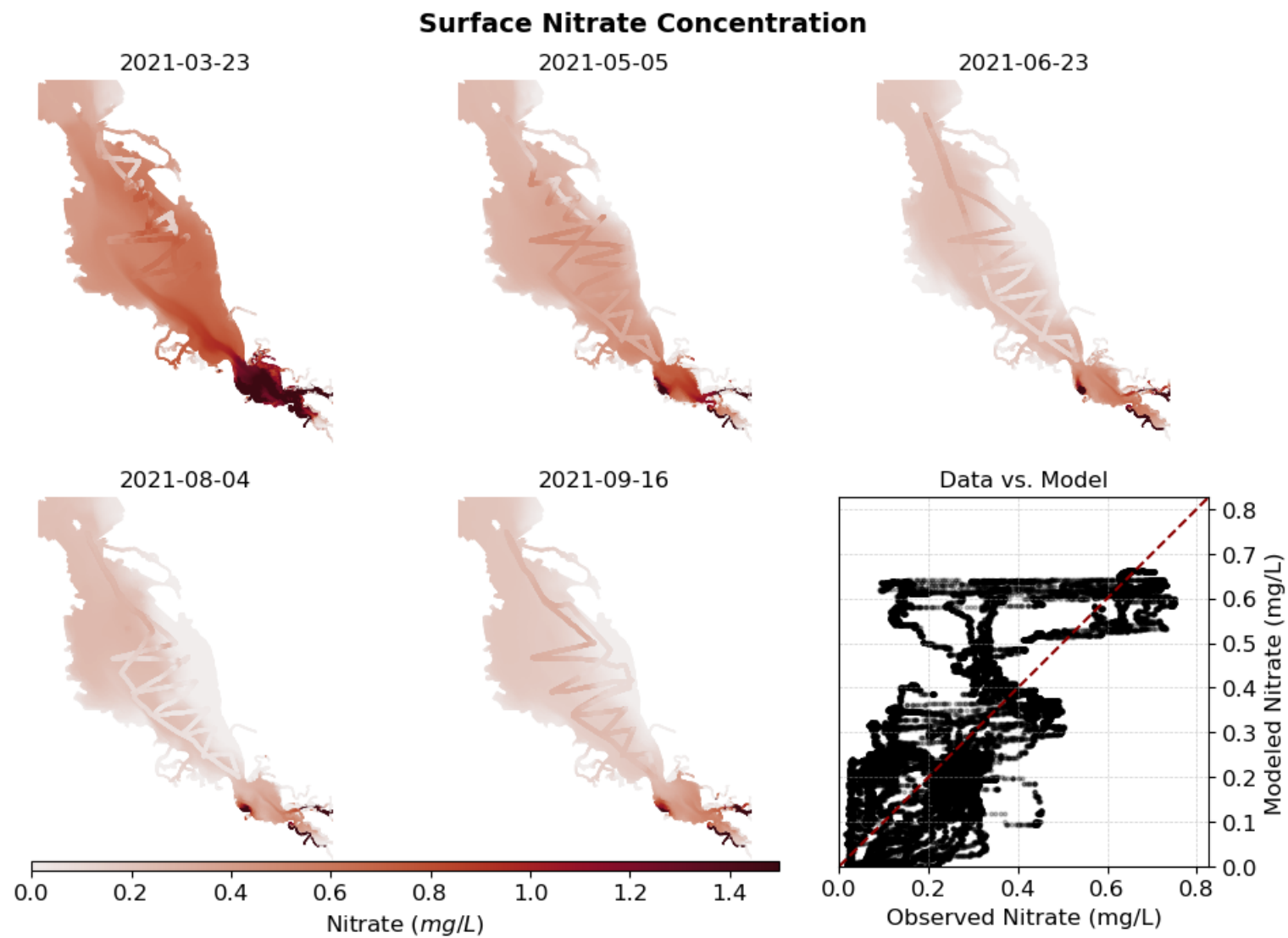


Figure 16: Surface map plots of modeled nitrate data and the mapping cruise data (first five subfigures). The instantaneous modeled surface nitrate concentration from 16:00 is shown. The surface nitrate concentration from the mapping cruise is overlaid on top of the model data. Scatter plot (bottom right) compares the mapping cruise observations and modeled nitrate. The dashed red line is the 1:1 line.

5.1.3 Phosphorus and Silicate

Neither phosphorus nor silica limit primary production within the Bay, and neither of these constituents have high frequency or mapping data unlike chlorophyll, nitrate or DO. Moreover, other key processes such as sorption of inorganic phosphorus to iron in suspended and bed sediments, and the subsequent release of phosphorus due to iron reduction in sediments are not directly represented in the model. We present these model state variables here to provide context on the influence of uptake and subsequent recycling within the water column and sediments (through remineralization), but note that our calibration efforts did not focus on either of these constituents. Furthermore, in the v2023 model, we had determined that phytoplankton uptake alone could not explain the observed ambient changes in dissolved inorganic phosphorus, and demonstrated that sediment-bound legacy phosphorus could reproduce the trend, although we did not definitively establish through other lines of evidence such as field observations that sediment-bound phosphorus is the driver of the ambient water column phosphorus changes. In the v2026 model, since we focus primarily on nitrogen calibration, we no longer use the legacy sediments as a source of inorganic nutrients (which tended to overestimate water column DIN levels even in the v2023 model). With this background we now present the current model performance relative to measurements.

Modeled dissolved inorganic phosphorus concentrations are shown compared to observations from the USGS Peterson Cruise (see Figure 17). Within LSB and South Bay, the model grossly underestimates phosphorus by up to a factor of 2 to 3, even though it shows some seasonal signals in uptake and recovery. These under predictions are not as severe in Central Bay and North Bay. In San Pablo Bay and Suisun Bay, the concentration similarly does not have much of a seasonal signal, and the model does show the same signal, albeit with some bias (low for the full resolution, and high for the aggregated grid). At a typical phosphorus to chlorophyll-a ratio of 1 to 2 (gP:gChla), it would take approximately 75 to 150 $\mu\text{g/L}$ of chlorophyll to produce a reduction of 150 to 300 $\mu\text{g/L}$ of phosphorus that is indicated in the measurements within LSB and South Bay. Such high levels of chlorophyll is not apparent in the observations presented earlier. This suggests that either there is considerable uptake of nutrients in the bay margins or we are missing external (tributary/point sources) and/or internal (legacy sediment) sources of phosphorus. Our review of discharge monitoring report loads for point sources did not indicate any discrepancies in the model input. Tributary inflows to LSB and South Bay are minor compared to point source inflows and tidal exchanges. Thus, the legacy material or uptake in the margins (or a combination of the two) are our strongest hypotheses to explain the discrepancies in the modeled phosphorus. As noted earlier, introducing a temperature-dependent legacy source was successful in reproducing these patterns in the v2023 model. Phosphorus measurements are not available in the margins or in the shoals, and therefore, these hypotheses cannot be verified at this time.

Water column silicate concentrations are compared to observations from the USGS Peterson Cruise (see Figure 18). Similar to DIN and orthophosphate, within LSB and South Bay the silicate data has a clear seasonal cycle which peaks in late summer and early fall and decreases until mid summer. As with DIN and phosphorus, these seasonal signals are muted in Central Bay. Unlike DIN and phosphorus, the magnitude of the silicate influx at the Delta boundary is significantly higher in the spring months suggesting that spring high flows bring in a substantial amount of dissolved silica into Suisun Bay. The model shows a similar seasonal cycle in LSB through Central Bay, with peaks that approximately align with the peaks in the data, however the amplitude of this cycle is not as high as in the data. The model does not capture the large oscillations in the silicate observations at the Delta boundary suggesting that the Delta boundary condition used for the silicate needs further refinement.

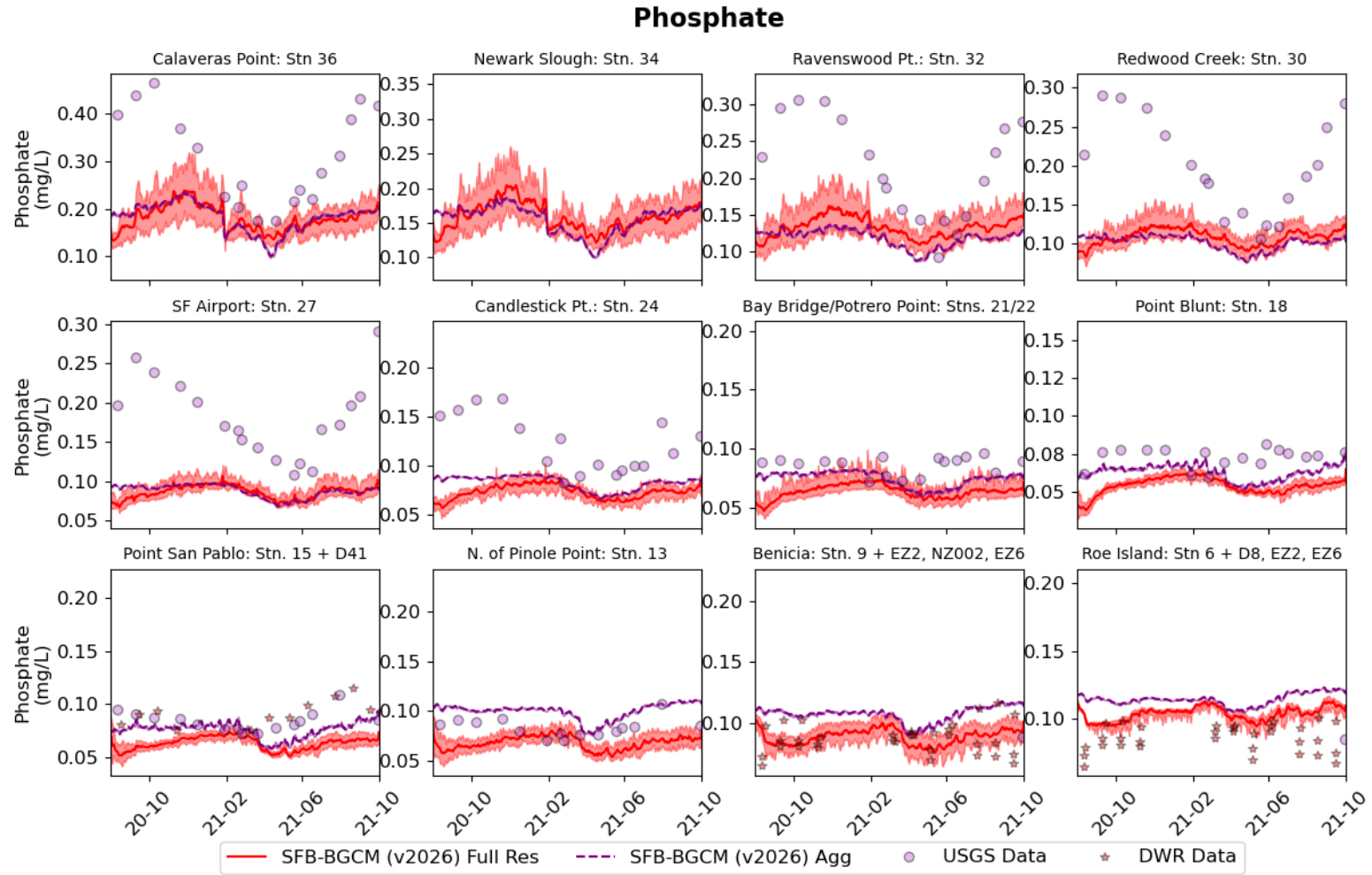


Figure 17: Comparison of daily averaged surface phosphate from the full resolution (red) and aggregated grid (purple dashed) model output with surface observations from the USGS Peterson Cruise Data (circles) and the DWR Environmental Monitoring Program (stars) at 12 long-term monitoring stations, one shown in each subfigure organized from LSB (top left) to Suisun Bay (bottom right). The shaded areas show the daily minimum and maximum values.

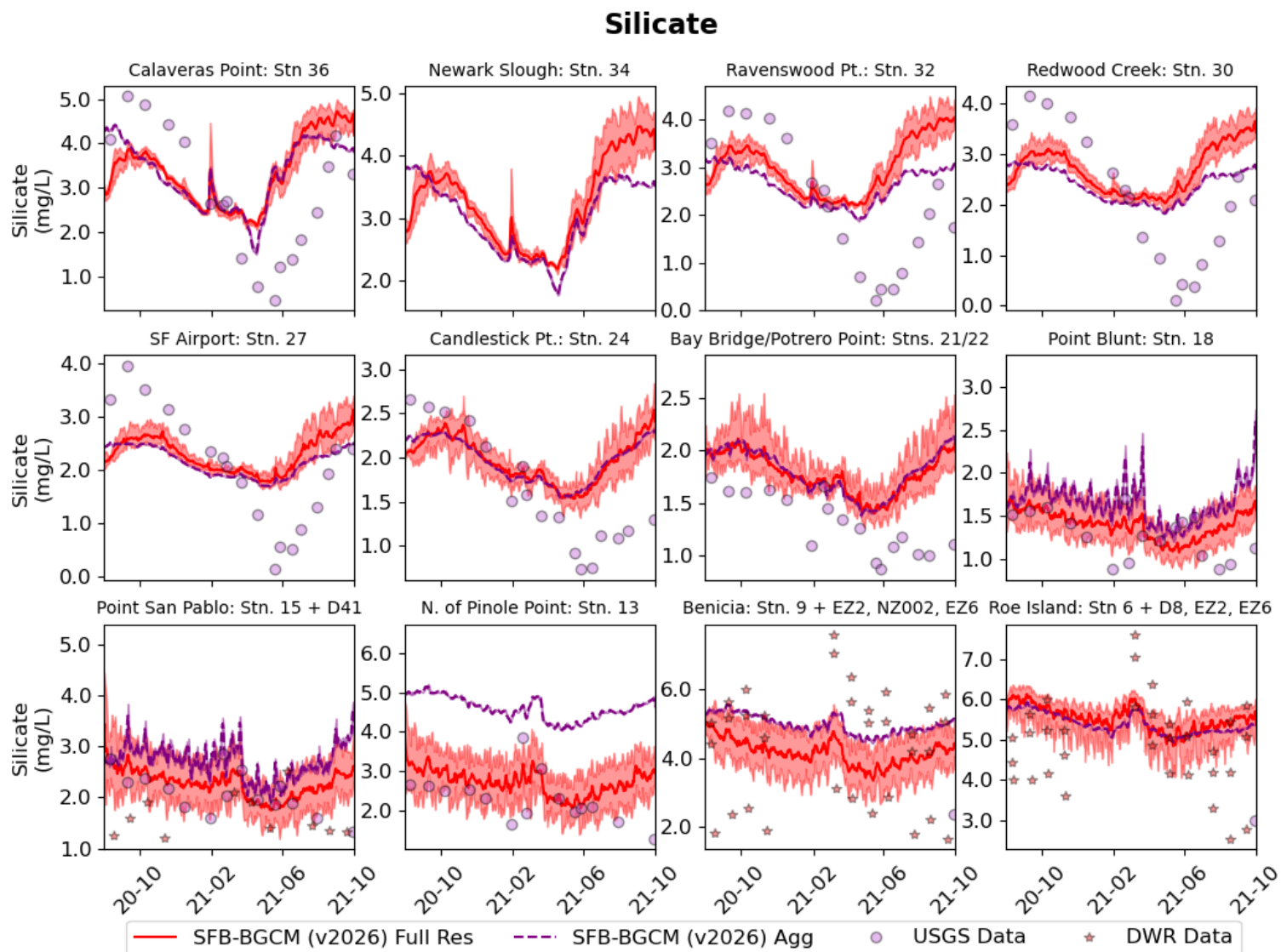


Figure 18: Comparison of daily averaged surface silicate from the full resolution (red) and aggregated grid (purple dashed) model output with surface observations from the USGS Peterson Cruise (circles) and the DWR Environmental Monitoring Program (stars) at 12 long-term monitoring stations, one shown in each subfigure organized from Lower South Bay (top left) to Suisun Bay (bottom right). The shaded areas show the daily minimum and maximum values.

As with the other nutrients, silica is taken up by both groups, but the spring algal group (“diatoms”) is modeled with a higher carbon to silica ratio (approximately 1:1), and thus takes up more silica than the summer algal group (“greens”), which is modeled with a carbon to silica ratio of 2:1. Additionally, the concentration of “greens” is generally smaller in the summer months. This probably explains some of the differences in the early to mid-summer lows that are not captured by the model. Even though measured silica levels are low in summer, it is still higher than the silicate half saturation constant used in the model (20 $\mu\text{g/L}$, which is within the literature range of 5 to 30 $\mu\text{g/L}$), and therefore does not limit primary production. Nonetheless, there are opportunities to improve silica further through a higher carbon to silica ratio for “greens”.

5.1.4 Dissolved Oxygen

Comparison to Channel Data

Depth-averaged model predictions are shown compared to the USGS Peterson Cruise channel locations for DO concentration (Figure 19, Figure A18), DO percent saturation (Figure 20), and the AOU (Figure 21). In combination, these three figures illustrate the DO conditions in the Bay, including periods of supersaturation and DO deficits reflecting the balance between the processes that supply oxygen (primary-production and reaeration) and the processes that consume oxygen (respiration and detrital decomposition). AOU represents the difference between the theoretical oxygen saturation (DO_{sat}) and the actual DO. A positive AOU shows undersaturation, and negative AOU shows supersaturation.

The observed DO shows a seasonal pattern of high DO in the winter and spring, and lower DO levels in summer and fall, which largely reflect seasonal changes bay wide (Figure 19). Locally, the magnitudes of these changes differ based on primary production and oxygen demand in the water column, and the spatial differences in salinity and temperature. This is evident in the larger range in DO within the LSB and Suisun Bay locations relative to other subembayments. The model reproduces these seasonal patterns and the spatial differences in DO generally well. However, the model generally overestimates the DO measurements throughout the Bay, and does not quite capture the DO lows in late fall and summer.

A closer examination of the DO saturation patterns indicates that the model is generally supersaturated in the spring and summer months within LSB and San Pablo Bay (Figures 20 and 21), which correspond to periods of high primary productivity in the model in these areas (see Figure 6). However, the corresponding DO declines in these areas in mid- to late-summer, particularly in LSB, where the observed DO saturation drops to 80% or lower is not captured by the model, suggesting that either the oxygen demand from the decaying biomass is easily satisfied within the model, or the biomass are flushed out faster before settling and exerting an oxygen demand. These are explored further in the control volume analyses presented later in this section, and the mass budget analyses presented in Section 6.

The Peterson Cruise collects depth profiles of oxygen. Modeled depth profiles are compared to data through a Hovmöller plot at the channel locations (Figure 22). In general, there is not much vertical variation both in the modeled and observed oxygen concentration, likely due to efficient vertical mixing. At Station 18, which is the deepest Station along the channel, and in San Pablo Bay (Stations 15 and 13) there is some vertical variation in the oxygen towards the bottom. Considering the highly energetic nature of this area due to the large tidal exchanges across Golden Gate bridge, it is unlikely that this oxygen deficit at the bottom is related to SOD. This likely originates from the conditions in the coastal ocean. At the ocean boundary, the model boundary condition is set to a vertically constant DO profile. This is most likely why the model does not have any vertical variation in the oxygen concentration at these locations.

Depth-Averaged Dissolved Oxygen

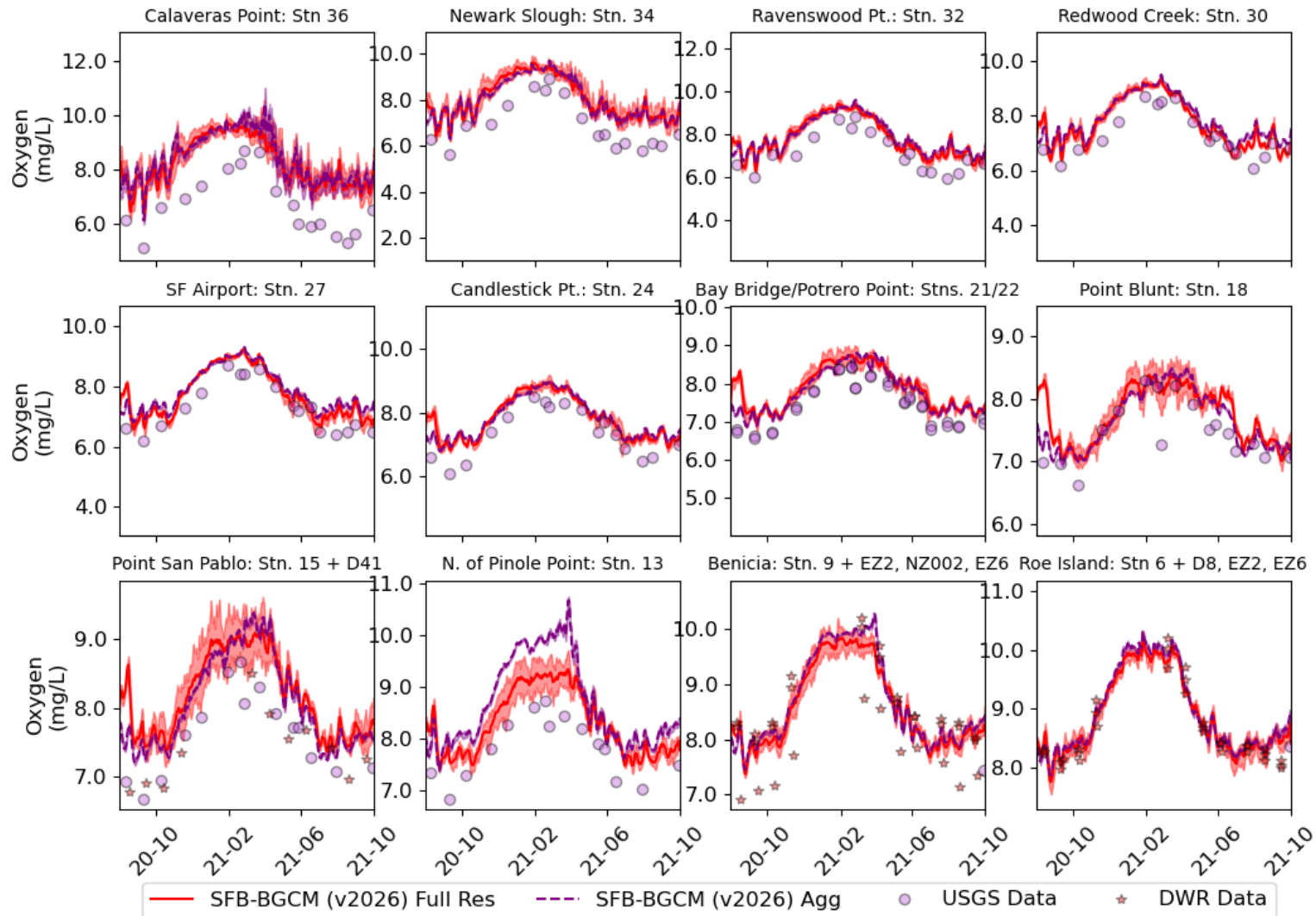


Figure 19: Comparison of daily- and depth-averaged dissolved oxygen from the full resolution (red) and aggregated grid (purple dashed) model output with depth-averaged observations from the USGS Peterson Cruise (circles) and the DWR Environmental Monitoring Program (stars) at 12 long-term monitoring stations, one shown in each subfigure organized from LSB (top left) to Suisun Bay (bottom right). The shaded areas show the daily minimum and maximum values.

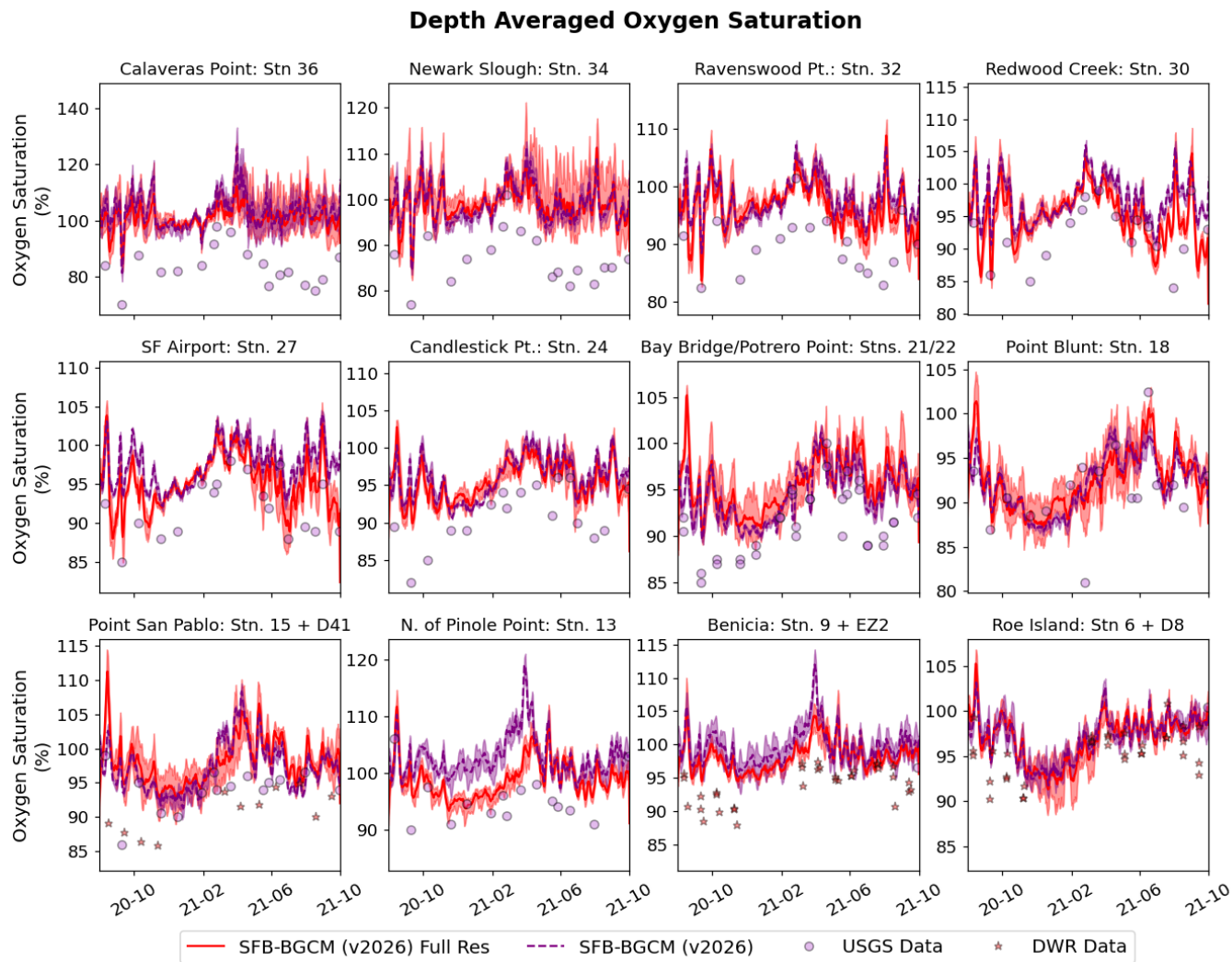


Figure 20: Comparison of daily- and depth-averaged oxygen saturation from the full resolution (red) and aggregated grid (purple dashed) model output with depth-averaged observations from the USGS Peterson Cruise (circles) and the DWR Environmental Monitoring Program (stars) at 12 long-term monitoring stations, one shown in each subfigure organized from LSB (top left) to Suisun Bay (bottom right). The shaded areas show the daily minimum and maximum values.

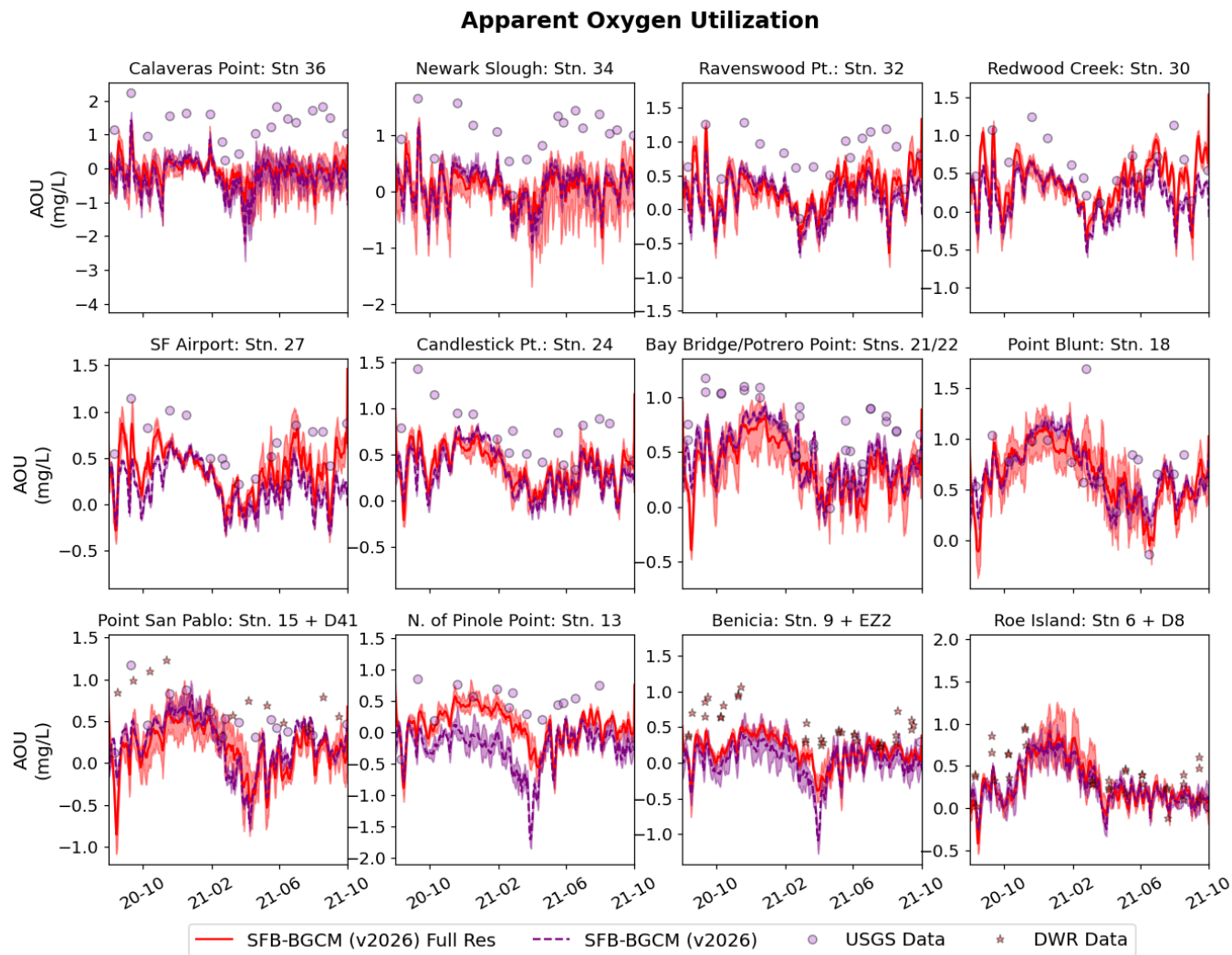


Figure 21: Comparison of daily- and depth-averaged AOU from the full resolution (red) and aggregated grid (purple dashed) model output with depth-averaged observations from the USGS Peterson Cruise (circles) and the DWR Environmental Monitoring Program (stars) at 12 long-term monitoring stations, one shown in each subfigure organized from LSB (top left) to Suisun Bay (bottom right). The shaded areas show the daily minimum and maximum values.

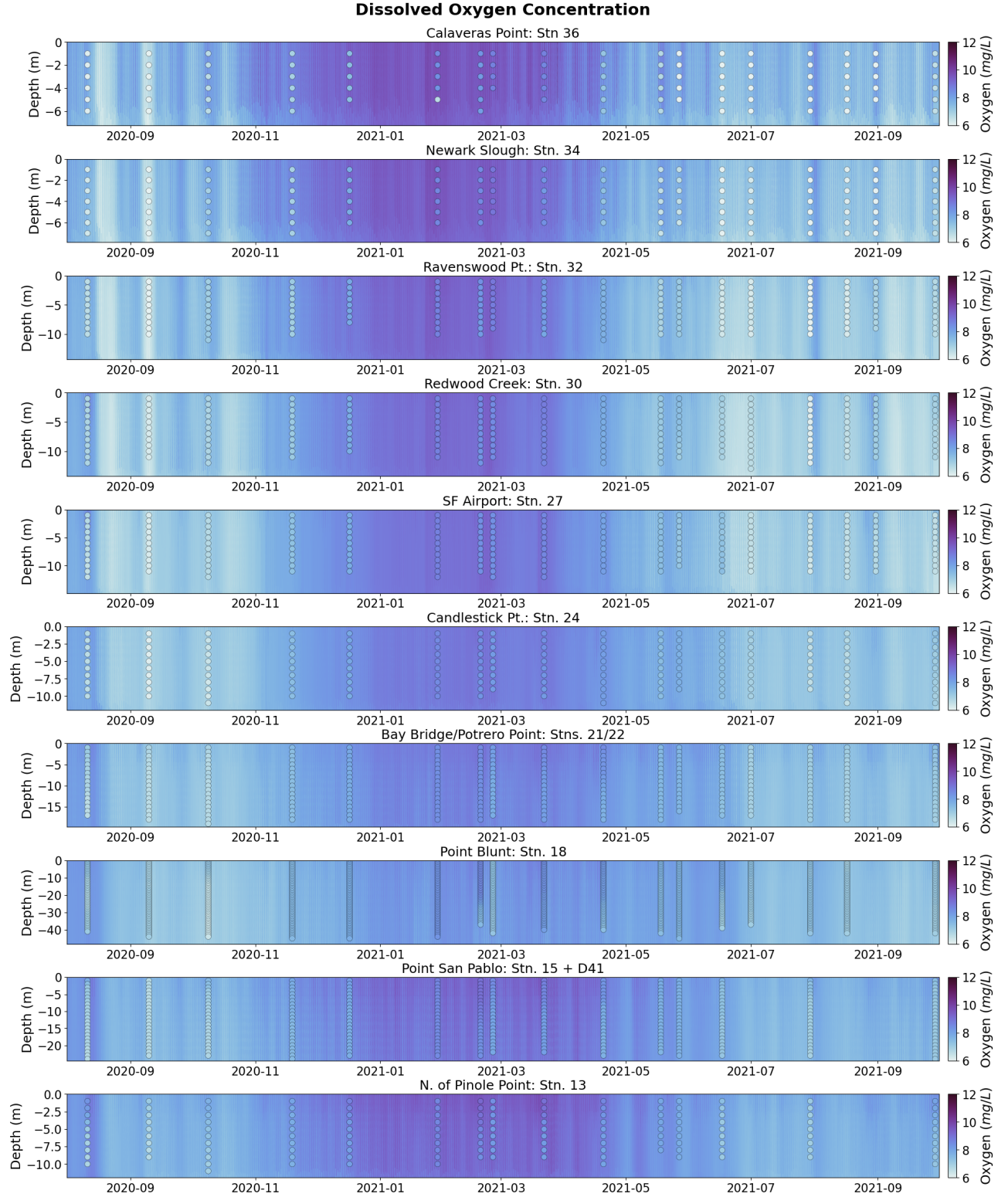


Figure 22: Comparison of depth-averaged dissolved oxygen from the full resolution model output with the USGS Peterson Cruise Data (circles).

Comparisons to High Frequency DO Data

A comparison of the modeled dissolved oxygen and high frequency mooring data is shown in Figure 23. The data and model comparison at the Dumbarton Bridge, San Mateo Bridge, and SFEI shoal mooring have the same general patterns and are discussed together to reduce repetition. At these three sites, the model captures the seasonal trend in the data, but as in the channel locations, the model mostly overpredicts the DO concentrations by approximately 1-1.5 mg/L. However, it is noteworthy that some peaks in the observed DO, that appears to be spring-neap-cycle related, align with the modeled data (for example, in fall of 2020 at Dumbarton Bridge). A closer examination relative to the high frequency chlorophyll data (Figure 7) indicates that the observed DO peaks correlate well with the chlorophyll peaks. The modeled data does not have the same amplitude of oscillations as the data, but does match the timing of some of the peaks. At the San Leandro Marina and the Hayward Mooring, the DO oscillations are substantially larger in magnitude, than at other locations, and also correspond to similarly large-amplitude oscillations in the chlorophyll concentrations, particularly in spring of 2021. The modeled oscillations are muted for both DO and chlorophyll, suggesting that getting the large amplitude oscillations in DO would also require capturing the large fluctuations in primary production and the associated respiration.

Comparisons to Mapping Data

Maps of surface oxygen concentration are compared to the mapping cruise data in Figure 24. Oxygen supersaturation is observed in the eastern shoal, towards the eastern extent of cruise locations near the Oakland Airport and Alameda, and corresponds very well to the primary production and nitrogen depletion in these areas (compare to Figures 10 and 16). The model captures the seasonal differences (higher in March and lower DO in Sep), but does not reproduce some of the spatial differences indicated in the data particularly the very high supersaturation events near Oakland Airport. This is also evident in the wide scatter between 10 to 12 mg/L in the mapping data, that are all grouped closer to near-saturation level concentrations in the model. Unlike the channel data where the model appears to be generally biased high, the model bias is not as apparent relative to most of the mapping data, except for these supersaturation events noted above. As an example, the model does a reasonable job of capturing the shoal to channel gradients in DO observed in the March, June and August mapping cruises.

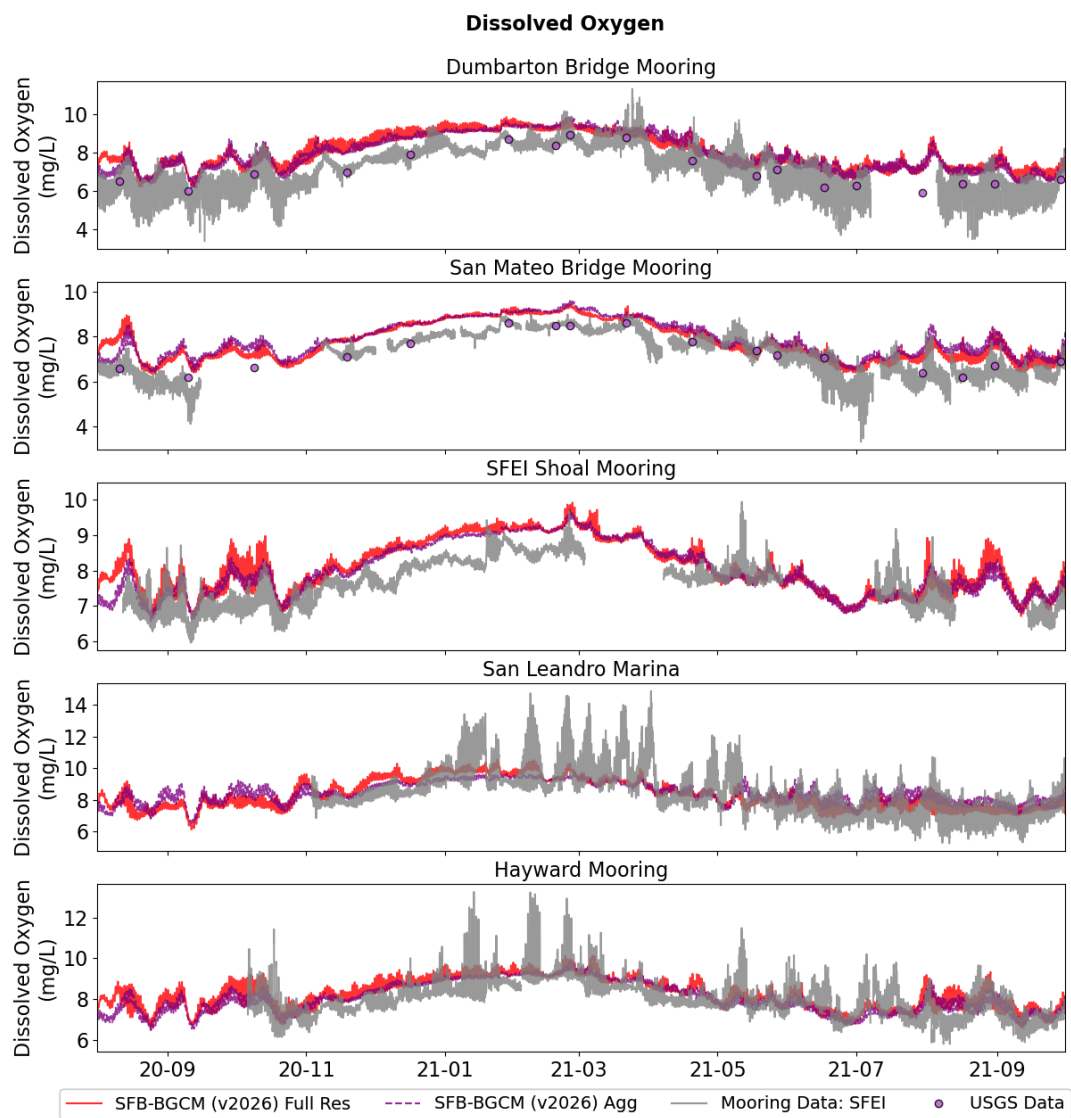


Figure 23: Comparison of surface dissolved oxygen concentration at the mooring sites for the full resolution (red) and aggregated grid (purple dashed) model outputs with high frequency observations from SFEI moorings (grey) and surface observations from the USGS Peterson Cruise (circles).

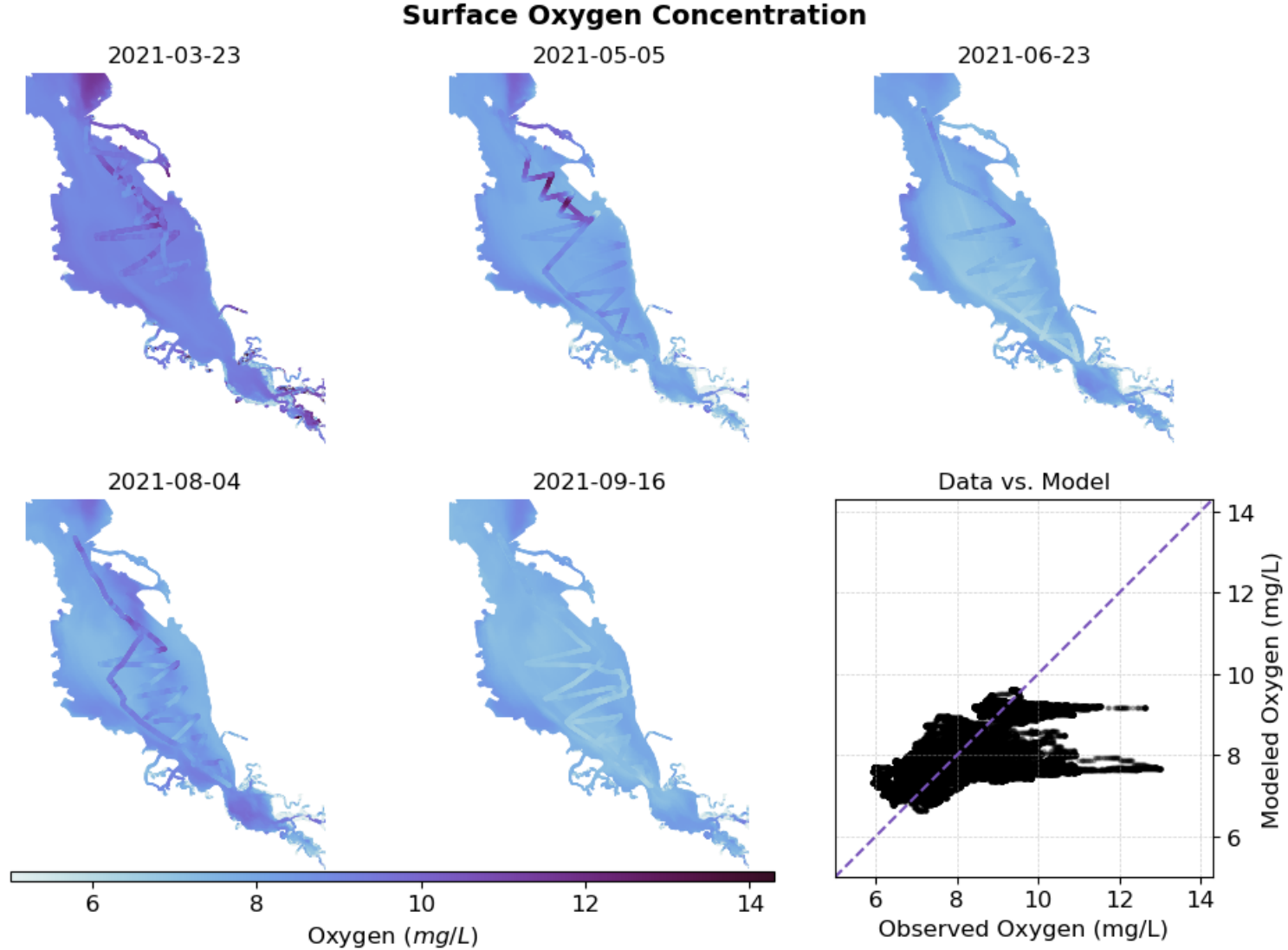


Figure 24: Surface map plots of modeled and mapping cruise oxygen (first five subfigures). The instantaneous modeled surface oxygen concentration from 16:00 is shown on each date. The surface oxygen concentration from the mapping cruise on the same date is overlaid on top of the model outputs using the same colorbar. Scatter plot (bottom right) compares the mapping cruise observations and modeled oxygen. The dashed purple line is the 1:1 line.

5.1.5 Sediment Fluxes

In this section we compare modeled SOD, denitrification and nutrient fluxes to the sediment flux measurements reported in Czapla et al., 2025. Czapla et al., 2025 estimated these measurements from laboratory incubations conducted with duplicate cores collected at each location during the months of February, May, August, and November 2021. Since our modeled water year ends in September 2021, we only consider the data from February, May, and August. Additionally, the full resolution model fluxes are calculated over the control volume polygons within which each sediment core is located. This averages the fluxes over the general area of the sediment core, rather than isolating it to a single model cell to avoid any local heterogeneity in deposition skewing the model results, and provide a more consistent comparison to the aggregated grid (the control volume polygons are essentially the same as the aggregated grid).

Sediment Oxygen Demand

The updates to the sediment-bed module (see Section 4.1.2) provides SOD as an output (Equation 10), from which we developed Figure 25 to provide a comparison to the SOD estimates from Czapla et al., 2025. The height of the bar is defined as the average for duplicate samples, with maximum and minimum values shown as error bars. For the model the 25th and 75th percentile of the spatial average of the corresponding polygons over the month are shown as error bars.

Figure 25 shows that in both the model and the sediment incubations, SOD is lower in winter, and higher and comparable in magnitude in May and August, reflecting the winter slowdown in benthic activity. Moreover, the model predictions as well as the sediment flux measurement exhibit greater SOD in the shoal locations relative to channel. In particular, both the model and sediment flux measurements exhibit greatest SOD in the east shoal of South Bay (S7 near the SFEI shoal mooring, S8 near Oakland Airport, and S9 which is south of San Mateo Bridge), where the mapping cruises and satellite measurements also show higher levels of chlorophyll and primary production. In the summer, at channel locations, the modeled SOD is generally greater than the sediment flux study estimates. This may be due to greater deposition of organic matter in the model. As discussed in Section 3.5.3, we have disabled the effect of shear stress on settling and resuspension, which is probably reasonable for lower energy shoal environments, but the model likely over estimates deposition within the channel, which conveys the majority of the tidal flows in the bay and experiences higher velocities and shear stress compared to the shoals.

Denitrification

Modeled denitrification rates from February and May are shown compared to the estimated rates from Czapla et al., 2025 in Figure 26. Denitrification measurements are not available for summer from Czapla et al., 2025. The modeled denitrification rates exhibit a seasonal variation similar to the measurements, with greater denitrification in May compared to the colder and less productive conditions in February. The modeled denitrification fluxes are within the range of measurements in February, but substantially larger in May. Because the model uses a simplified approach to denitrification, the denitrification directly corresponds to the spatial and temporal patterns of SOD. While this simplification is useful in predicting the overall DIN and GPP patterns in the bay, the spatial patterns in denitrification are probably overestimated in the model.

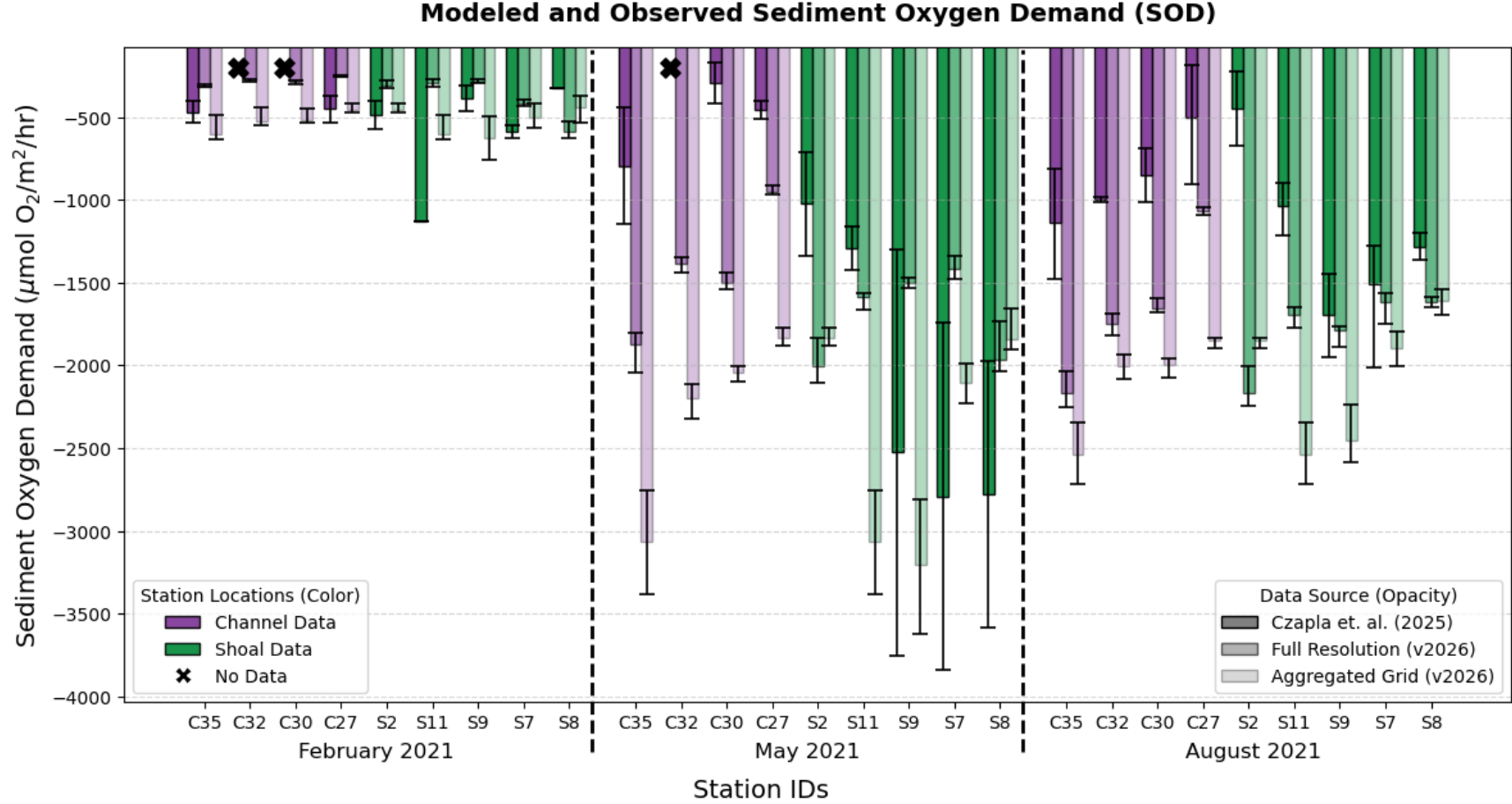


Figure 25: Figure showing modeled and observed estimates of SOD. Observational data is taken from Czapla et al., 2025 (full opacity) and model outputs are from the Full Resolution and Aggregated Grid configurations (v2026) (medium opacity, and low opacity, respectively). The data bar heights are given as an average of all points at each location and time, with the whiskers indicating the maximum and minimum if more than one sample exists. The model output bar heights are derived from the associated control volume from each monitoring location and averaged over the entire month, with whiskers indicating the 25th and 75th percentiles over that month. The channel stations (C35, C32, C30, C27) are shown in purple, and the shoal stations (S2, S11, S9, S7, S8) are shown in green. Note that in Czapla et al., 2025, the channel stations are referred to as C2, C4, C5, C6, respectively. The order of the stations are listed south to north.

The current understanding of patterns of denitrification in the Bay indicate that more denitrification occurs in the sediments in the spring through summer and more mass is removed in South Bay than LSB (Czapla et al., 2025). Czapla et al., 2025 study indicates the largest denitrification occurs in LSB, and in the east shoal locations S8 and S7. They also extrapolated these loss rates for nitrogen at a subembayment scale, and estimated that the greatest losses occurred within South Bay rather than LSB due to the larger area over which denitrification removes nitrogen. Based on the values reported in Table 1 of Czapla et al., 2025, we estimated that the denitrification loss is about 0.9 mega grams per day (Mg/d or equivalently Metric Tons/d) in LSB in both May and August, and within South Bay we estimate it as 0.675 - 2.5 Mg/d (average 2 Mg/d) in May and 0.075 - 4.25 Mg/day (average of 0.75 Mg/d) in August.

To better understand the modeled spatial patterns in denitrification we evaluate the spatial denitrification patterns within South Bay in spring and summer months as shown in Figure 27. A version of this figure for the whole bay is shown in Figure A26. While these figures do not have any direct comparison to data, it allows us to establish a model reference state for estimates and patterns of denitrification in the model. The figure is separated between the timing of the spring bloom (March, April, May), and the sustained summertime productivity (June, July, August). We calculate both the average mass and the average consumption of nitrogen from denitrification from model output directly at each control volume polygon (Figure 27, top and middle rows), from which we estimate the effective denitrification as follows:

$$r_{\text{eff,denit}} = \frac{\text{NO3, dDetDenit (Mg/d)}}{\text{NO3, Mass (Mg)}} \cdot \frac{\text{Volume (m}^3\text{)}_{cv}}{\text{Area (m}^2\text{)}_{cv}}. \quad (24)$$

Note that we take both the area and the volume of each control volume to determine the average depth of the control volume. The base state of the model in MAM shows highly enriched nutrient conditions in the channel, and into central South Bay that diminish at both shoals and over time (top row in Figure 27). The denitrification patterns show high values in both the eastern and western shoals, with the predominant loss of nitrogen occurring along the eastern shoal, where we previously noted higher primary production and associated nitrogen limiting conditions in the mapping cruises. The channel denitrification is much smaller during this time not because there isn't enough nitrogen, but because the delivery of organic matter is far greater to the sediments in the shoal. The range of denitrification rates in the both models in spring months (March, April, May) is approximately 1.5 - 2.5 Mg/d in the eastern shoals near the shoreline, with lower values in the channel and western shoal. This increases slightly in summer in the full resolution model (to about 2.5 Mg/d on the eastern shoal, and to about 0.5 - Mg/d in the channel) and to a larger extent in the aggregated grid model. But these value are still within the range of 0.675 - 3.5 Mg/d for May and 0.075 - 4.25 Mg/d in August reported in Czapla et al., 2025. Similarly, the range of effective denitrification rate (Equation 24) in both the models in the spring months (MAM) is approximately 0.1 - 0.15 m/d in the spring months with lower values (0.05-0.1 m/d) in the channel, and approximately the same values in the western shoal. This increases in the summer in both versions of the model, 0.2-0.5 m/d in the full resolution and 0.3-1.0 m/d in the aggregated grid model. These fall slightly high of the ranges 0.015 - 0.1 m/d for May, and 0.06 - 0.4 m/d for August derived from Czapla et al., 2025.

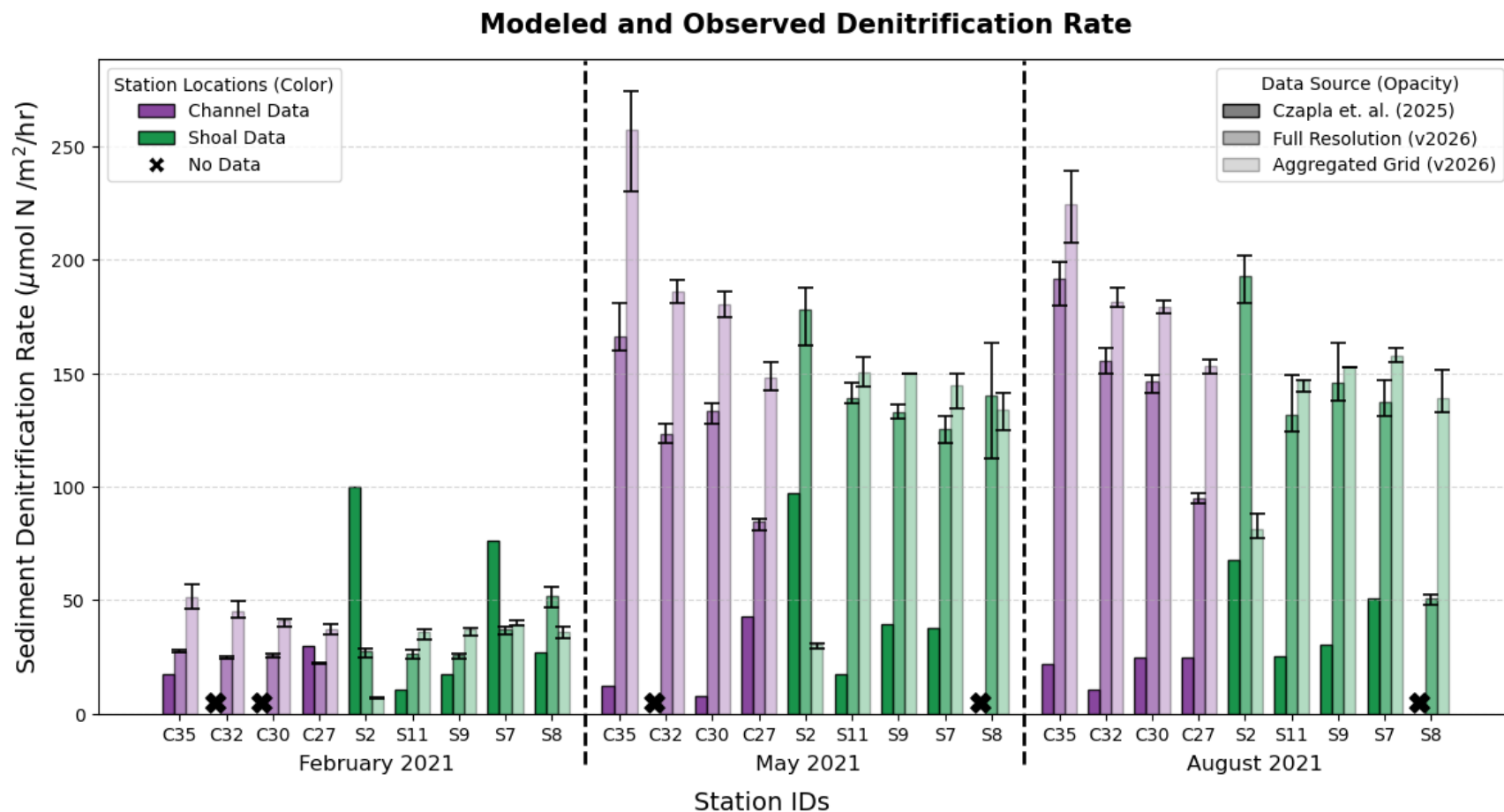


Figure 26: Figure showing modeled and observed estimates of denitrification rate. Observational data is taken from Czapla et al., 2025 (full opacity) and model outputs are from the Full Resolution and Aggregated Grid configurations (v2026) (medium opacity, and low opacity, respectively). The data bar heights are from the D14 denitrification estimates. The model output bar heights are derived from the associated control volume from each monitoring location and averaged over the entire month, with whiskers indicating the 25th and 75th percentiles over that month. The channel stations (C35, C32, C30, C27) are shown in purple, and the shoal stations (S2, S11, S9, S7, S8) are shown in green. Note that in Czapla et al., 2025, the channel stations are referred to as C2, C4, C5, C6, respectively. The order of the stations are listed south to north.

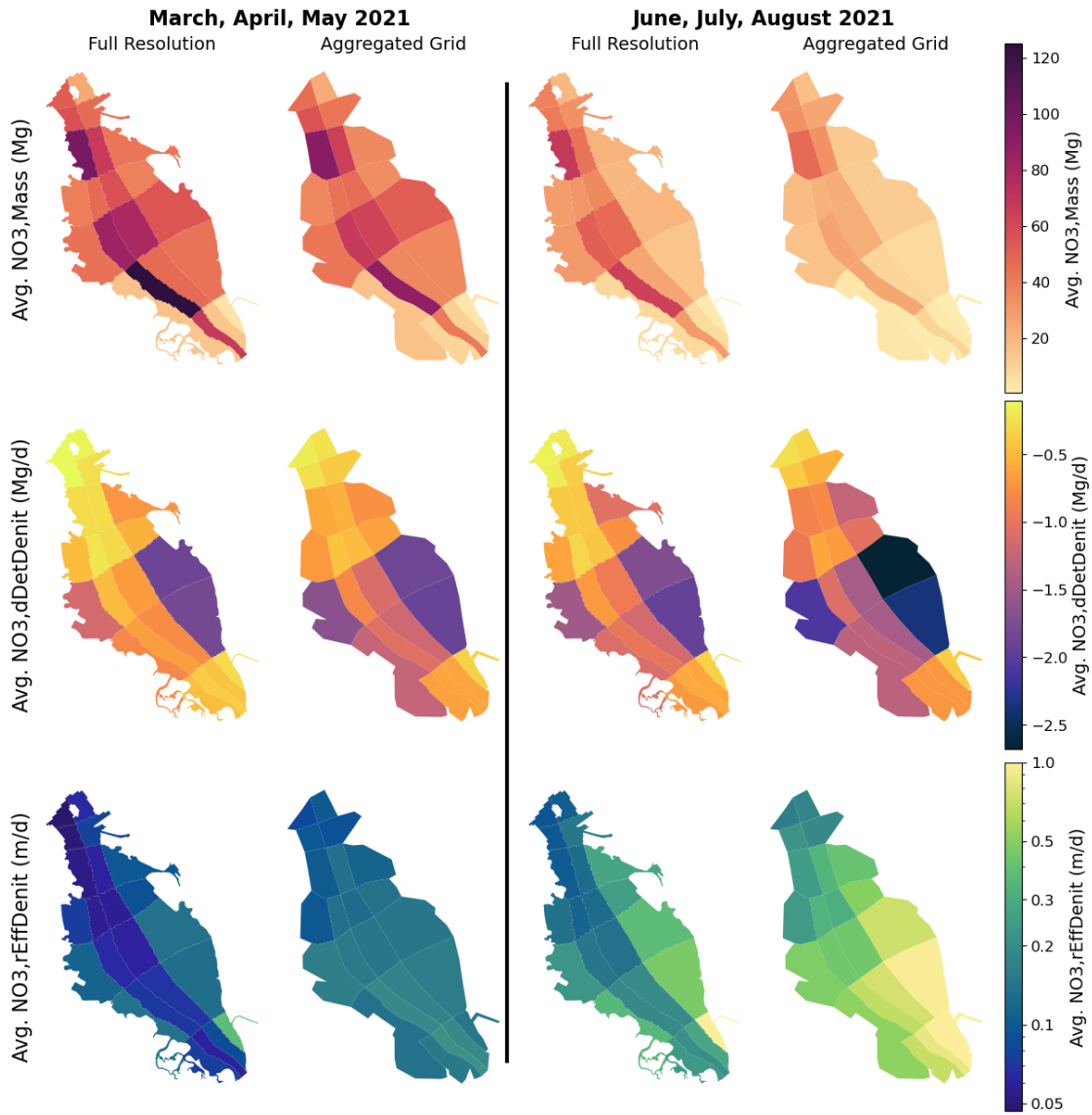


Figure 27: Figure showing control volumes in South Bay for average nitrate mass in each volume (top row), average modeled denitrification flux in each volume (center row), and the effective denitrification rate (bottom row). The left major column shows the model output for the full resolution (left minor column) and aggregated grid (right minor column) for March, April, May 2021. The right major column shows the model output for June, July, August 2021.

5.2 Sensitivity Analysis

The sensitivity analysis presented in this section explores the behavior of the model for each of the biogeochemical model improvements described in Section 3. Table 7 describes the parameters to test the responses in primary productivity, and Table 8 describes parameters used to test responses in nutrient concentrations.

During the model calibration, several parameter combinations were evaluated which yielded insight into some of the influential parameters which are presented below. The results in this section are intended to provide insight into model response to key parameters and processes that are relevant to NMS management and science questions, and provide context for ongoing model calibration efforts. Model sensitivity is evaluated at a sub-embayment level through the following N fluxes that each provide a key component of the focal processes targeted during the improvements that directly affect nitrogen cycling:

1 Phytoplankton uptake of DIN

2 Internal recycling of DIN which is calculated as follows:

$$\begin{aligned} \text{DIN, Recycling (Mg/d)} = & \text{DIN, Net Reaction (Mg/d)} + \text{DIN, Uptake (Mg/d)} \\ & + \text{NO}_3, \text{ Denitrification (Mg/d)}. \end{aligned} \quad (25)$$

Internal recycling represents addition of DIN through phytoplankton respiration, autolysis and remineralization of particulate organic matter in the water column and sediments. The net reaction in the equation above represents all DIN reactions, which includes losses from uptake and denitrification. So, the losses are added back in to calculate the internal recycling fluxes.

3 Denitrification which represents a permanent loss of N from the system.

We present the fluxes above as the yearly averages by sub-embayment for perturbations above and below the calibration value for key parameters in each process. In addition, we also calculated the cumulative export of DIN at Golden Gate bridge to provide context for the changes in the net assimilation of DIN within SFB. We note that, there could be export of organic nitrogen (both dissolved and particulate) that is not accounted in the DIN flux.

The patterns of uptake, recycling, and denitrification are seasonally dependent and so we separate the response of different parameters between spring (March, April, May [MAM]) and summer (June, July, August [JJA]) and present maps of percent difference between the model base state and an experiment by season, with percent difference calculated as follows:

$$\text{Percent Difference} = \frac{(\text{Model Base State}) - \text{Experiment}}{\text{Model Base State}}. \quad (26)$$

Additionally, we show figures showing the yearly averaged rates in each major subembayment including Suisun Bay, San Pablo Bay, Central Bay, South Bay, and LSB. We specifically call out the rates occurring in the South Bay East Shoal where we anticipate major contributions of denitrification and uptake, with the reference average values being approximately of the same order as Central and San Pablo Bay. Time series plots at the USGS channel locations, presented in the appendices, additionally provide sensitivity of the model to the validation data set.

All sensitivity analyses presented herein were conducted with the aggregated grid model due to the large number of simulations necessary. As evident in previous sections, the model predictions between the aggregated and full resolution grids are generally comparable. Thus, the sensitivities indicated herein are applicable to the full resolution model results as well, particularly at the control volume and the sub-embayment scale.

5.2.1 Fraction of Organic Matter Remineralized by Oxygen (frmindod1)

The remineralization rate is determined by a combination of the concentration of organic matter in the sediments and a user defined remineralization rate shown in Equation 9, R_{remin} (rsdet1c). We find that the model is insensitive to values of F_{ox} (frmindod1) less than 0.85, but are more sensitive to values larger than 0.9, at all spatial scales evaluated including the sub-embayment and whole bay level (Figure 28), at the control volume level (Figure 29) as well at the calibration stations (Figure A27 and Figure A28).

This asymmetric response provides useful insights on the relative importance of the permanent loss of DIN through denitrification versus temporary loss through uptake that gets recycled back to the water column. The higher sensitivity at 0.95 indicates that with only 5% of the sediment remineralization mediated by NO_3 , a large proportion of nitrate remains in the water column. This facilitates greater uptake and water column recycling (Figure 28 and Figure 29), higher concentrations of nitrate in the water column (Figure A27), lower nitrate retention and loss in sediments through denitrification and ultimately greater export of nitrate from the bay (bottom panel in Figure 28). In fact, the DIN export at Golden Gate is more than double that is the reference simulation. A majority of these changes occur in summer months (see Figure 29 and Figure A27).

On the other hand, increases in nitrate mediated remineralization does not cause a significant change in either uptake, recycling, denitrification fluxes, water column concentration or DIN export. Denitrification requires both organic matter and nitrate. It appears that at the reference F_{ox} value of 0.9, the shoals in the bay, where majority of the denitrification and primary production occurs, are already nitrogen limited in the summer months limiting both the production and delivery of organic matter to sediments as well the availability of nitrate for denitrification. Further reductions in the F_{ox} are sensitive because the system is already at a steady state relative to the external sources of DIN (POTWs) and the internal assimilation (uptake and denitrification). When F_{ox} is increased, there is more internal capacity for uptake, but not permanent assimilation/loss within the Bay.

Considering that our sediment model is a simplified representation of numerous complex redox and biological processes that occur with the sediments, this is an excellent outcome for our calibration because not only is our bulk parameter F_{ox} sensitive over a relatively narrow range, but the reference (calibrated) value (of 0.9) is at the edge of that sensitive range that helps explain the water column DIN and primary production reasonably well. This effectively bounds the extent of denitrification within our model by reconciling water column observations of both nitrate and primary production. In other words, at a value of 0.9 the model is permanently removing just the right amount of nitrate where the primary production in the water column and external (POTWs) and internal (sediment and water column recycling) sources of DIN remain in balance relative to observations. Reducing F_{ox} further does not seem to affect this balance, while increasing it 0.95 demonstrates higher DIN levels and greater primary production overall that is inconsistent with observations (notice the significantly higher DIN and the large deviations in GPP relative to data in Figures A27 and A28, respectively).

Figure 28 shows maps of the average uptake, recycling, and denitrification in the bay, along with the contribution of internal model processes in each major subembayment. South Bay has the largest overall magnitude in each modeled process. Similar to the channel data, we see that the case in which $F_{\text{ox}} = 0.85$ does not differ much from the reference value in all three modeled processes. This value is 5% less than the reference fraction of organic matter remineralized aerobically. The second case presented is $F_{\text{ox}} = 0.95$, this corresponds to an increase of 5% in aerobic respiration. This experiment is asymmetric to highlight that the predicted model concentrations in the channel is insensitive to values less than 90% of aerobic respiration.

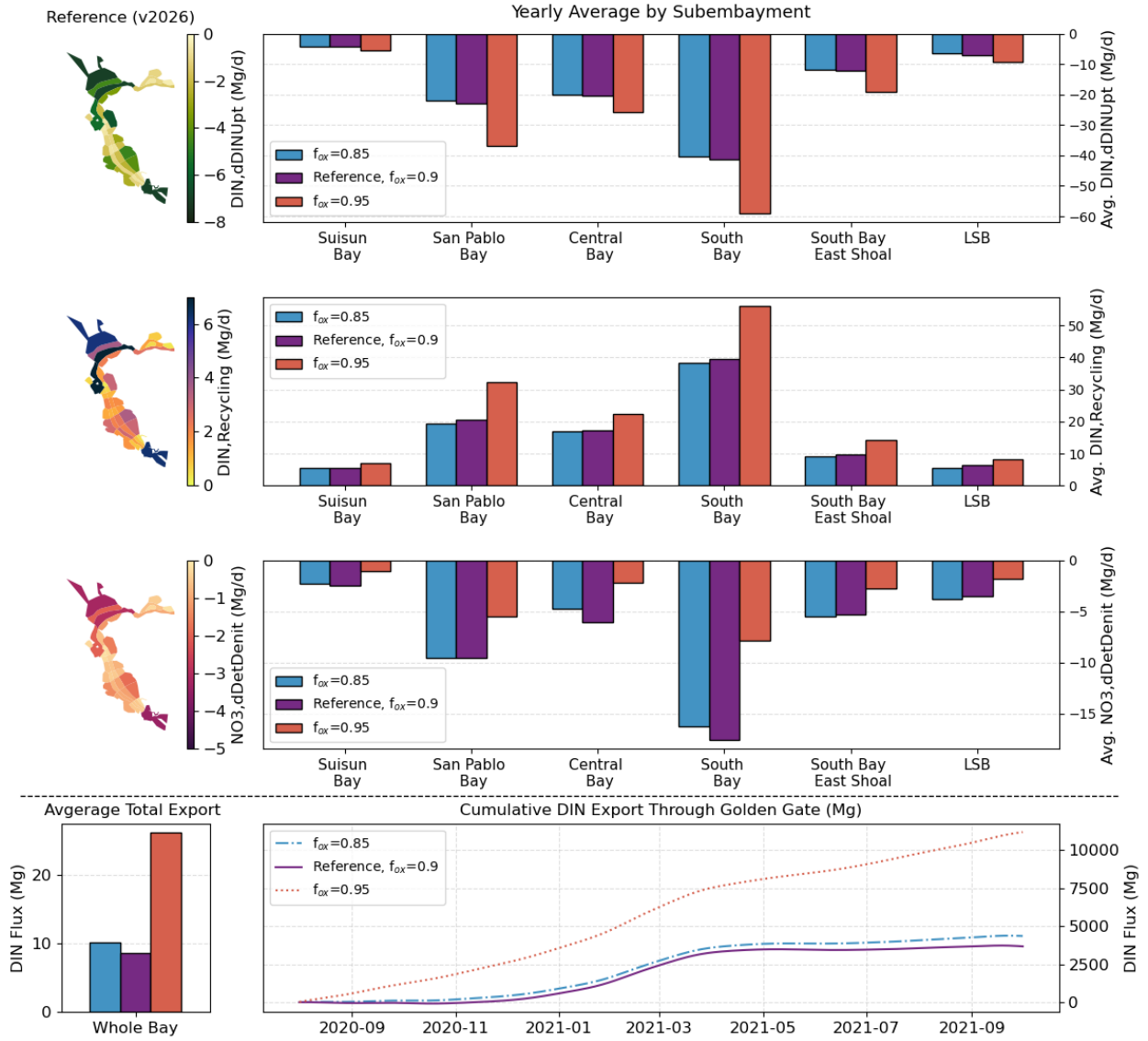


Figure 28: Figure showing control volume analysis for sensitivity to the fraction of aerobic remineralization (f_{ox}) for the whole bay for yearly average dissolved inorganic nitrogen uptake rate in each volume (top row), yearly average DIN recycling in each volume (second row), and yearly average denitrification in each volume (third row). The left column shows a map of the reference model output for the entire water year. The right column shows a bar chart comparing the reference solution to two representative model sensitivity experiments. The bottom row shows the year average (left) and cumulative (right) DIN export through the Golden Gate. All three runs used the aggregated grid.

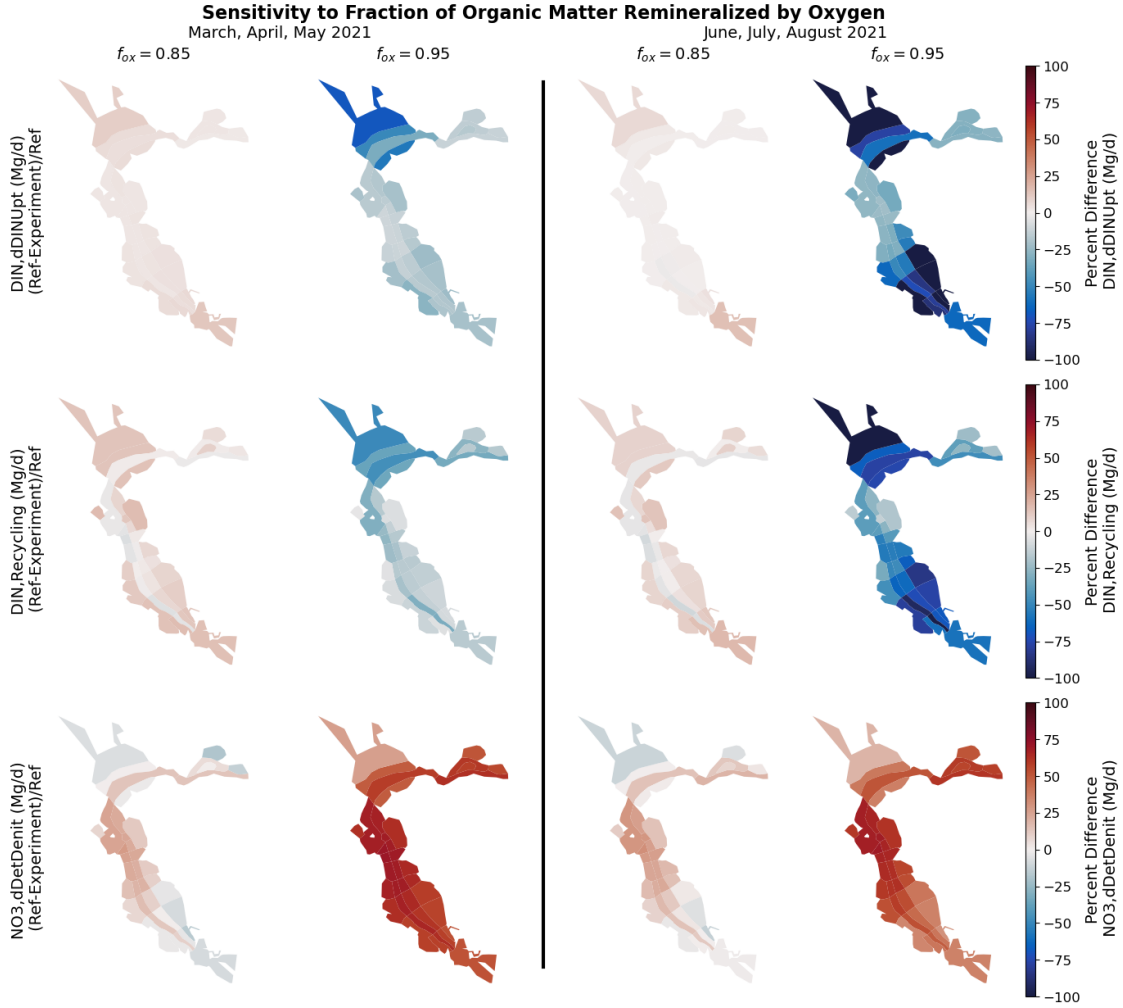


Figure 29: Figure showing control volumes for the whole bay for an experiment focusing on the fraction of aerobic remineralization in the sediments (f_{ox}) using the aggregated grid. The figures show the percent difference between the reference case ($f_{ox} = 0.9$) and experiments with $f_{ox} = 0.85$ (left minor column) and $f_{ox} = 0.95$ (right minor column) of March-April-May (left major column) and June-July-August (right major column) average DIN uptake (top row), average DIN recycling (middle row), and average denitrification in the sediment (bottom row). Positive (negative) values, shown in red (blue), indicate regions of the domain where the reference state has larger (smaller) magnitude than the results in the experiment.

5.2.2 Half Saturation for Radiation in Phytoplankton (RadSat)

Current understanding of controls of phytoplankton blooms and eutrophication resistance relies on suspended sediment concentrations and persistent light limitation (Cloern, 2007). As a consequence, the impact of light limitation on model processes is important to understand. Model perturbations to half saturation for radiation was investigated at 50% and 150% of the reference value. We note here that a decrease in the half saturation coefficient corresponds to a decrease in phytoplankton light limitation. The opposite behavior is true for an increase in the half saturation for radiation.

In the channel, the modeled chlorophyll concentration shows a shift in the timing of the spring-time bloom to earlier in the water year (Figure A30). This shift occurs approximately one month earlier, with a larger sustained fall biomass concentration. On the other hand an increase in the half saturation shifts the timing of the spring bloom two weeks later. There is a slightly asymmetric response around the reference value, with the response of the half saturation coefficient being more sensitive to a reduction in RadSat than increase. As a consequence, the channel DIN shows an increase in overall DIN concentration with more light limitation, and a decrease in DIN with less light limitation for phytoplankton (Figure A29).

The contribution of major model processes controlling DIN concentration for each subembayment is shown in Figure 30. The behavior of uptake in Suisun Bay, Central Bay, South Bay, and LSB follows similar behavior with more average uptake in the subembayment with less light limitation (0.5RadSat), and less average uptake with more light limitation (1.5RadSat). In San Pablo Bay and in the East Shoal of South Bay, both perturbations in half saturation show an increase in the magnitude of uptake, which indicates that there is some non-linearity in the phytoplankton and DIN response to half saturation. In all subembayments, lowering the half saturation coefficient show an increase in DIN recycling. This is a consequence of changes in rates by subembayment and resulting transport (see Figure 31). Over the entire model results indicate that as light limitation decreases the denitrification increases. Because denitrification rate is directly a consequence of phytoplankton production and resulting settling (both phytoplankton and organic matter), the increase in phytoplankton uptake with reduced light limitation drives more organic matter to the sediments and thus increases denitrification. Overall, the net assimilation in the bay increases at decreasing light limitation, which is evident in the reduced exports at Golden Gate; the opposite is evident at higher light limitation with reduced assimilation within the Bay and greater export at Golden Gate.

Control volume maps of DIN uptake in MAM show that a reduction in the half saturation coefficient corresponds to nearly a baywide increase in DIN uptake, with the exception of the San Pablo Shoals (Figure 31). The opposite behavior occurs when the half saturation of radiation is increased. In JJA, the behavior is more complex with Central Bay, Suisun Bay exhibiting less uptake, and San Pablo Bay, South Bay and LSB exhibiting greater uptake. The model generally shows the opposite behavior in uptake with an increase in light limitation. Seasonal denitrification rates show that decreased light limitation correspond to an increase in denitrification rates in MAM, with the exception of the San Pablo North Shoal. Increased light limitation corresponds to decreased denitrification rates in the experimental results. In JJA, less light available for photosynthesis shows a decrease in denitrification rates baywide. The increase in available light shows an increase in production in Central Bay, in the channel, and in Suisun Bay. The shoal regions in San Pablo Bay and South Bay show a decrease in denitrification when compared to the model reference state. This likely is a consequence of the same seasonal reduction in phytoplankton uptake. The response of water column recycling rates on the other hand are offset by the denitrification rates. When denitrification decreases or increases, water column recycling increases or decreases respectively to compensate.

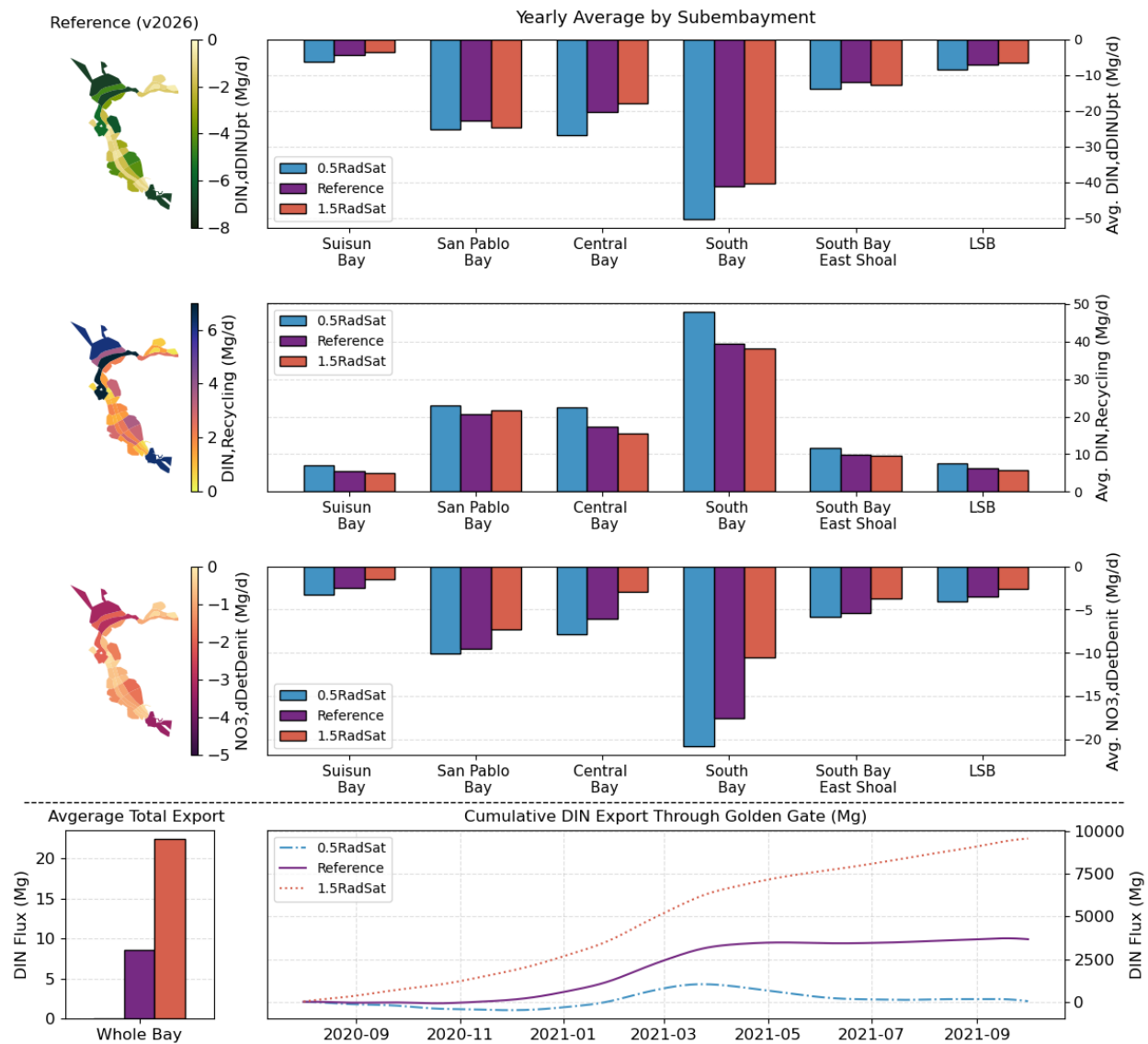


Figure 30: Figure showing control volume analysis for sensitivity to the half saturation coefficient for radiation in phytoplankton uptake (RadSat) for the whole bay for yearly average dissolved inorganic nitrogen uptake rate (top row), yearly average DIN recycling (second row), and yearly average denitrification (third row). The left column shows a map of the reference model output for the entire water year. The right column shows a bar chart comparing the reference solution to two representative model sensitivity experiments. The bottom row shows the year average (left) and cumulative (right) DIN export through the Golden Gate. No bar is shown for the 0.5-RadSat run because the yearly average DIN export was 0. All three runs used the aggregated grid.

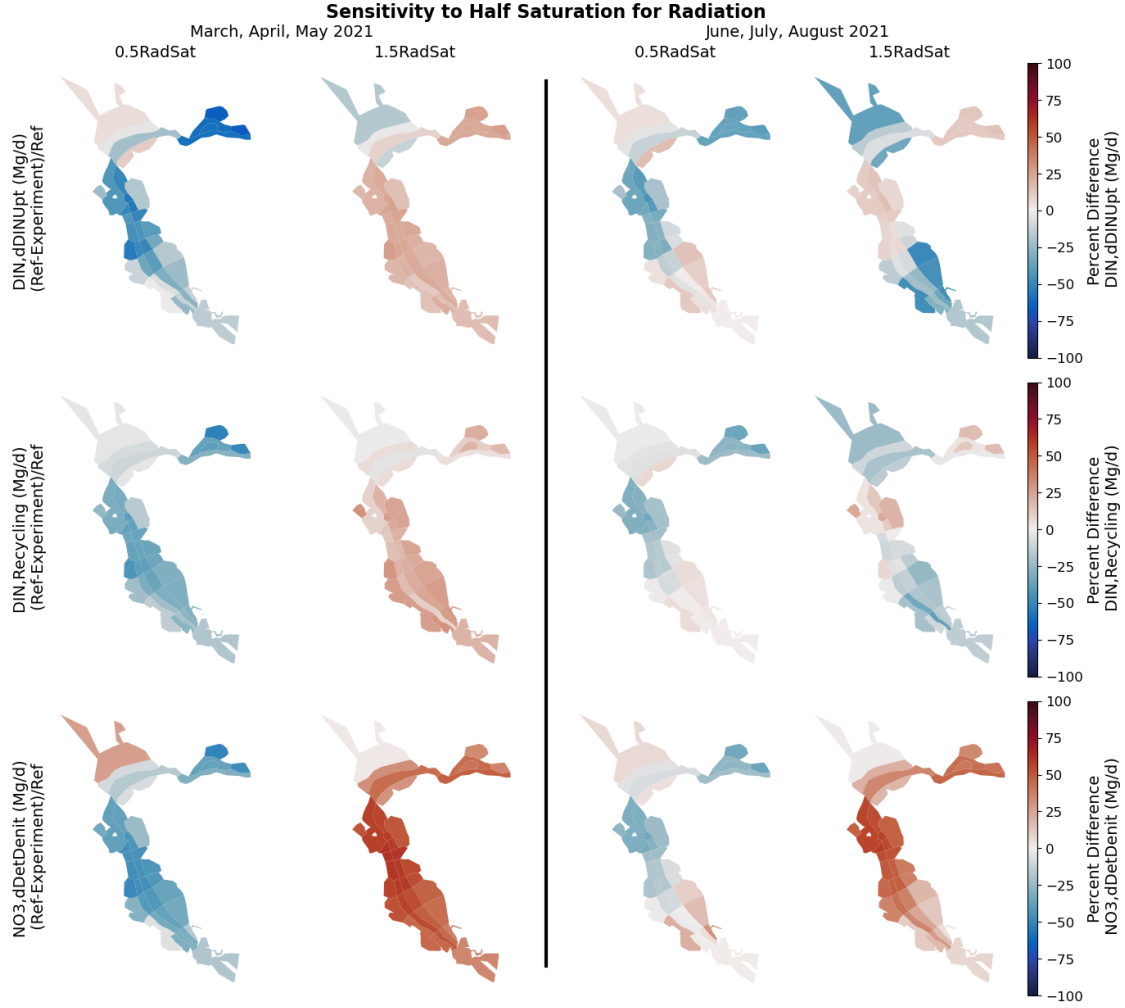


Figure 31: Figure showing control volumes for the whole bay for an experiment focusing on the half saturation for radiation in phytoplankton uptake (RadSat) using the aggregated grid. The figures show the percent difference between the reference case and experiments with 0.5·RadSat (left minor column) and 1.5·RadSat (right minor column) of March-April-May (left major column) and June-July-August (right major column) average DIN uptake (top row), average DIN recycling (middle row), and average denitrification in the sediment (bottom row). Positive (negative) values, shown in red (blue), indicate regions of the domain where the reference state has larger (smaller) magnitude than the results in the experiment.

5.2.3 Maximum Phytoplankton Growth Rate (PPMaxDiat and PPMaxGreen)

The phytoplankton maximum growth rate is an important parameter as it controls the initial process by which nutrient cycling occurs. The maximum growth rate has been shown to be an important control on the total chlorophyll concentration and DIN in King et al., 2020 and King et al., 2023. For this version of the calibrated model we investigated a $\pm 25\%$ change to the reference value concurrently for both phytoplankton groups in the model.

A comparison of the response to the modeled channel DIN and total chlorophyll to the phytoplankton maximum growth rate is shown in Figures A31 and A32. The chlorophyll biomass in the channel shows an increase corresponding to an increase in maximum growth rate. There is a similar reduction in overall DIN concentration however the response is asymmetric around the reference value. With increased phytoplankton growth rate, there is some reduction of DIN concentration in the channel. During the wintertime, this decrease is up to 0.01-0.02 mg/L in Central Bay and up to 0.1 mg/L in South Bay. The Summertime concentrations show little difference after June and July. With a reduction of maximum growth rate, the DIN concentrations increase in the channel up to 0.3 mg/L in Central Bay and 0.5-1 mg/L in South Bay and LSB. These are likely a result of nutrient limiting conditions, or effective top-down control with quadratic mortality or both.

The internal model processes that control the total DIN concentration by subembayment are shown in Figure 32. The uptake rates by subembayment increase as the maximum growth rate increases. Similarly, the total recycling increases. The denitrification in each subembayment, however, does not always increase as phytoplankton production increases. In LSB, Suisun Bay, and in the South Bay East Shoal, the model reports an increase in denitrification as phytoplankton growth rates increase. However, over San Pablo Bay, Central Bay, and through the entire South Bay subembayment the average denitrification decreases. A potential explanation for this could be that phytoplankton mortality does not occur in the same place that uptake occurs and transport processes are important in determining the deposition of particulate organic matter and phytoplankton settling. Another potential explanation is that the denitrification is limited by nitrate availability, especially in the shoals.

Spatial maps of the seasonal average phytoplankton uptake, recycling, and denitrification are shown in Figure 33. In MAM, patterns of DIN uptake follow expected responses given perturbations to the maximum growth rate. For example, a decrease in maximum growth rate corresponds to the model uptake rate decrease in the experimental results, and vice versa. In JJA, the increase in maximum growth rate for phytoplankton corresponds to an increase in the modeled rate. However, a decrease in maximum growth rate does not correspond to a decrease in modeled DIN uptake baywide. In the San Pablo North Shoal and South Bay Eastern Shoal, the model reports an increase in the modeled DIN uptake with decreased maximum growth rate. The model recycling rates in both MAM and JJA correlate to a decrease or increase when the maximum growth rate is decreased or increased respectively. This points to a combination of transport and efficient water column recycling. The modeled denitrification in MAM and JJA shows that when the maximum uptake rate is decreased, the total denitrification rate also decreases at the water column-bed interface. This is a likely consequence of a reduction in organic matter flux to the sediments. When the maximum uptake rate is increased, the denitrification rates in MAM show an increase in the South Bay, but a decrease in Central Bay and San Pablo Bay, a potential explanation for this could be that particles and phytoplankton are exported through the Golden Gate Strait. In JJA, there is a reduction in the total denitrification baywide with the exception of the San Pablo Shoal and regions in Suisun Bay. A potential explanation for this is that there is a functional limit in the amount of nitrogen that can be denitrified at this time in the model. The channel data does not

1192 indicate that the deep channel is nitrogen limited, however, the mapping cruise data indicates that
 1193 nitrogen levels are limiting in South Bay and may result in a reduction in the total denitrification
 1194 rate (see Figure 16).

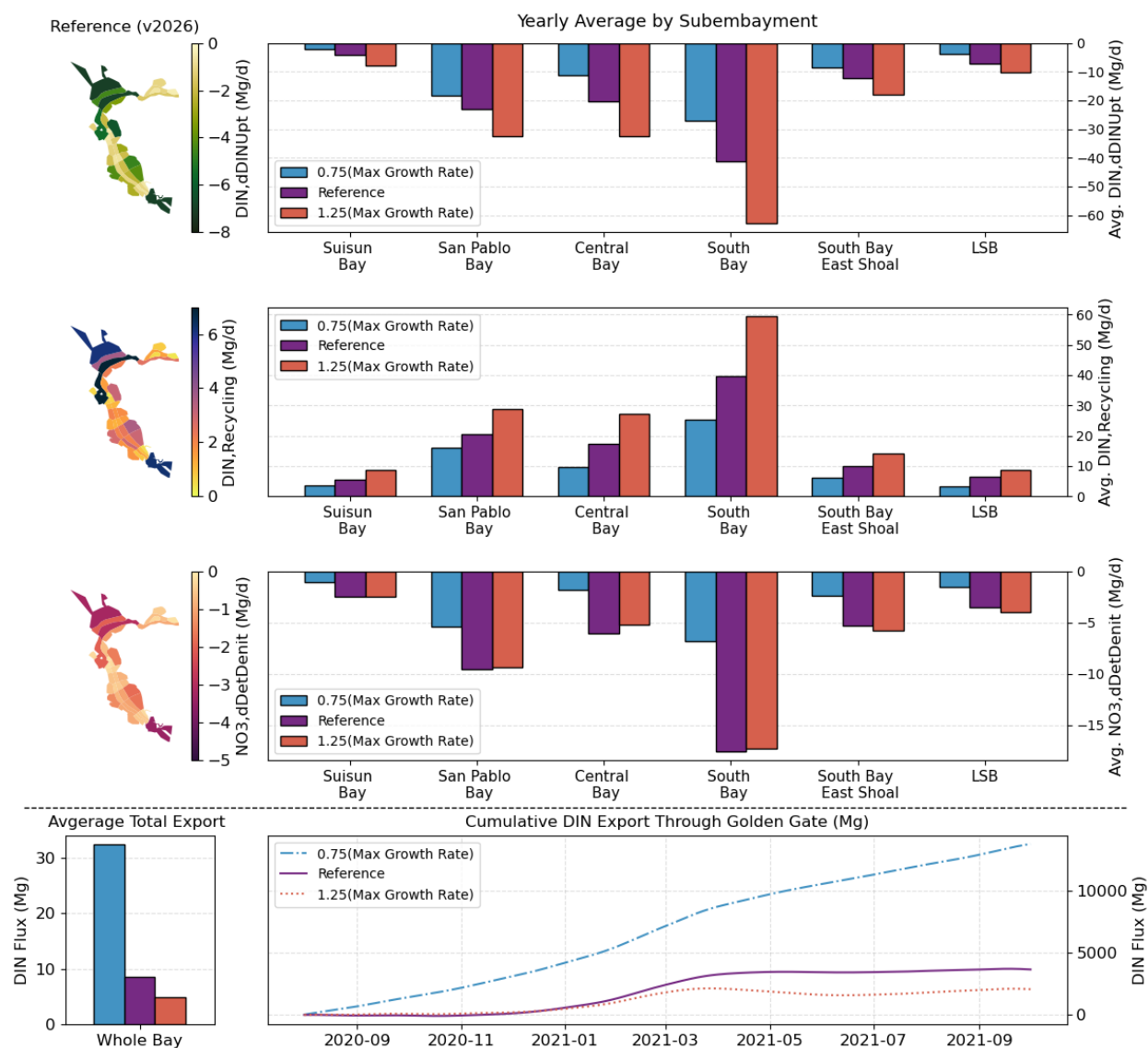


Figure 32: Figure showing control volume analysis for sensitivity to the phytoplankton maximum growth rate (Max Growth Rate) for the whole bay for yearly average dissolved inorganic nitrogen uptake rate (top row), yearly average DIN recycling (second row), and yearly average denitrification (third row). The left column shows a map of the reference model output for the entire water year. The right column shows a bar chart comparing the reference solution to two representative model sensitivity experiments. The bottom row shows the year average (left) and cumulative (right) DIN export through the Golden Gate. All three runs used the aggregated grid.

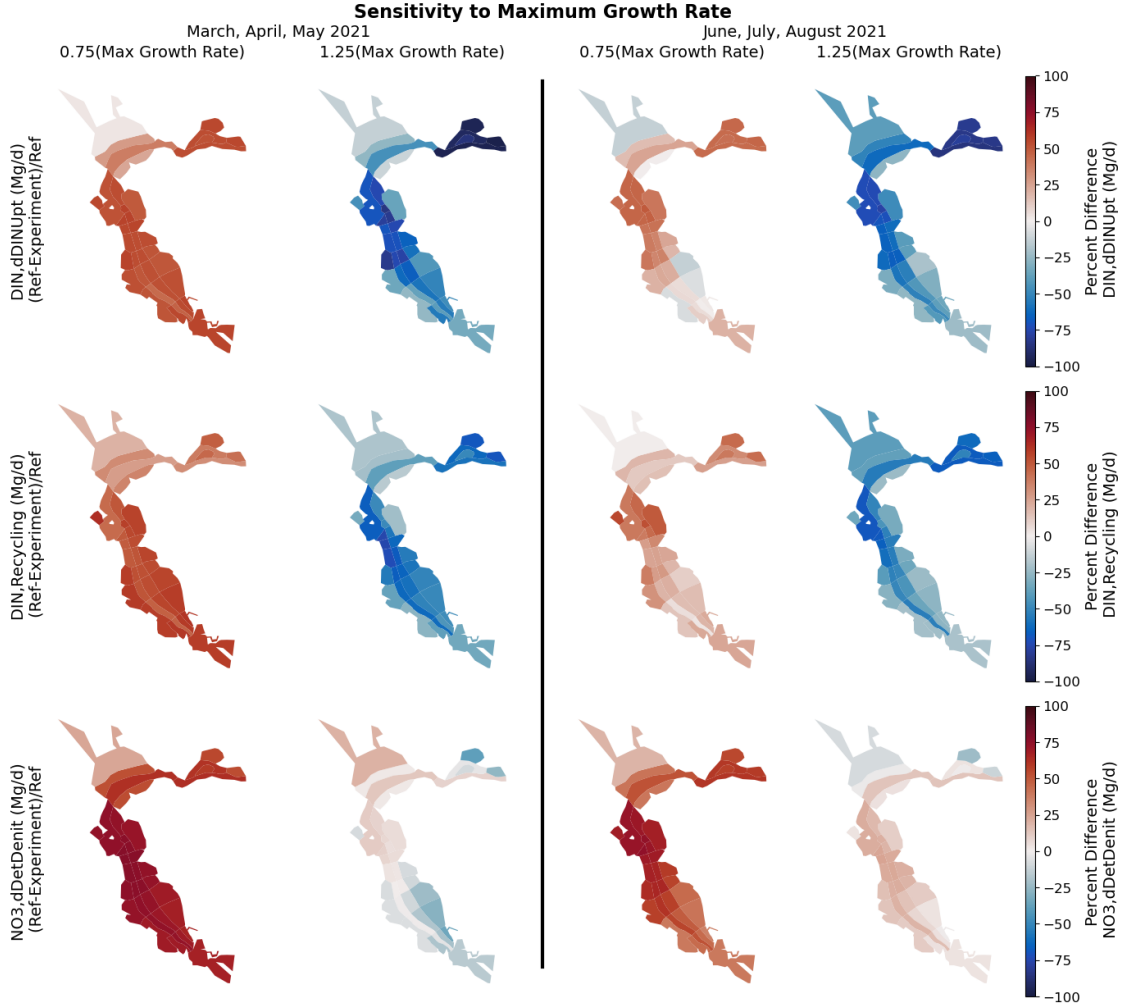


Figure 33: Figure showing control volumes for the whole bay for an experiment focusing on the diatom (PPMaxDiat) and green (PPMaxGreen) maximum growth rates (together: Max Growth Rate) using the aggregated grid. The figures show the percent difference between the reference case and experiments with 0.75·Max Growth Rate (left minor column) and 1.25·Max Growth Rate (right minor column) of March-April-May (left major column) and June-July-August (right major column) average DIN uptake (top row), average DIN recycling (middle row), and average denitrification in the sediment (bottom row). Positive (negative) values, shown in red (blue), indicate regions of the domain where the reference state has larger (smaller) magnitude than the results in the experiment.

5.2.4 Perturbations to Top-Down Control (Mort0 and QMort)

The primary mechanisms for top-down control in this formulation of the model include a combination of linear (natural) and quadratic (pseudograzing) mortality. The quadratic mortality term allows for phytoplankton to bloom at small concentrations, but provide strong top-down control with large concentrations. We investigated a 25% change from reference condition in either direction for all mortality (linear and quadratic, and for both phytoplankton groups). The coefficient for quadratic mortality, QMort, is salinity dependent and the DIN concentration responds non-linearly.

The channel concentrations of DIN and total chlorophyll are shown in Figures A33 and A34 respectively. Decreases in phytoplankton mortality do not show substantial changes in the channel DIN concentrations, with very small increases from the reference model state. Increases in total mortality, on the other hand, show increases in DIN baywide. Chlorophyll concentrations in the channel increase throughout the bay between 2-5 $\mu\text{g/L}$.

Internal model processes that control the DIN concentration response to perturbations in the total phytoplankton mortality by subembayment are shown in Figure 34. The response to perturbation in the total mortality is non-linear in San Pablo Bay and South Bay showing an increase in total DIN uptake with both an increase and decrease in all mortality. Similar behavior occurs in the recycling rates in these regions. There is some linear behavior in LSB, Suisun Bay, and Central Bay. This likely indicates the total phytoplankton biomass in these regions is less than 1 mg C/L (recall that the model does not internally calculate chlorophyll concentrations but instead calculates phytoplankton biomass). This indicates that we likely need to update either the quadratic mortality function with a parameter to indicate when quadratic mortality is dominant. The denitrification rates baywide decrease with both perturbations to the total mortality, this was a surprising result in this parameter exploration. We suspected that increases in total mortality would correlate to an increase in particle production, settling, and sediment denitrification. This however, does not seem like the case on the yearly average. As a result, this module likely requires future development and model calibration.

Seasonal maps of percent difference in average uptake, recycling, and denitrification with perturbations to the quadratic mortality rates are shown in Figure 35. A reduction in mortality shows that the total DIN uptake increases baywide in MAM and JJA. This is expected as a reduction of mortality would make phytoplankton production more efficient. As a consequence the DIN recycling increases when mortality decreases, likely due to less particle formation and export to the sediments. This can be seen in the denitrification field in both MAM and JJA. The model reference state here has higher levels of denitrification when compared to a decrease in both linear and quadratic mortality. When the quadratic mortality increases, the uptake in MAM decreases baywide with the exception of in Suisun Bay. This pattern is additionally seen in the recycling rate. Again, there is a decrease in the total denitrification at the sediment bed. This likely points to inefficient settling of particulate matter. In JJA, an increase in phytoplankton mortality shows an increase in DIN uptake. A potential explanation could be that quadratic mortality is not evenly distributed between diatoms and greens. Greens tend to be more efficient in JJA, and could be the driver of increased production in South Bay and San Pablo Bay. This similarly drives increases in recycling, with a similar explanation to explain the pattern of denitrification as in MAM.

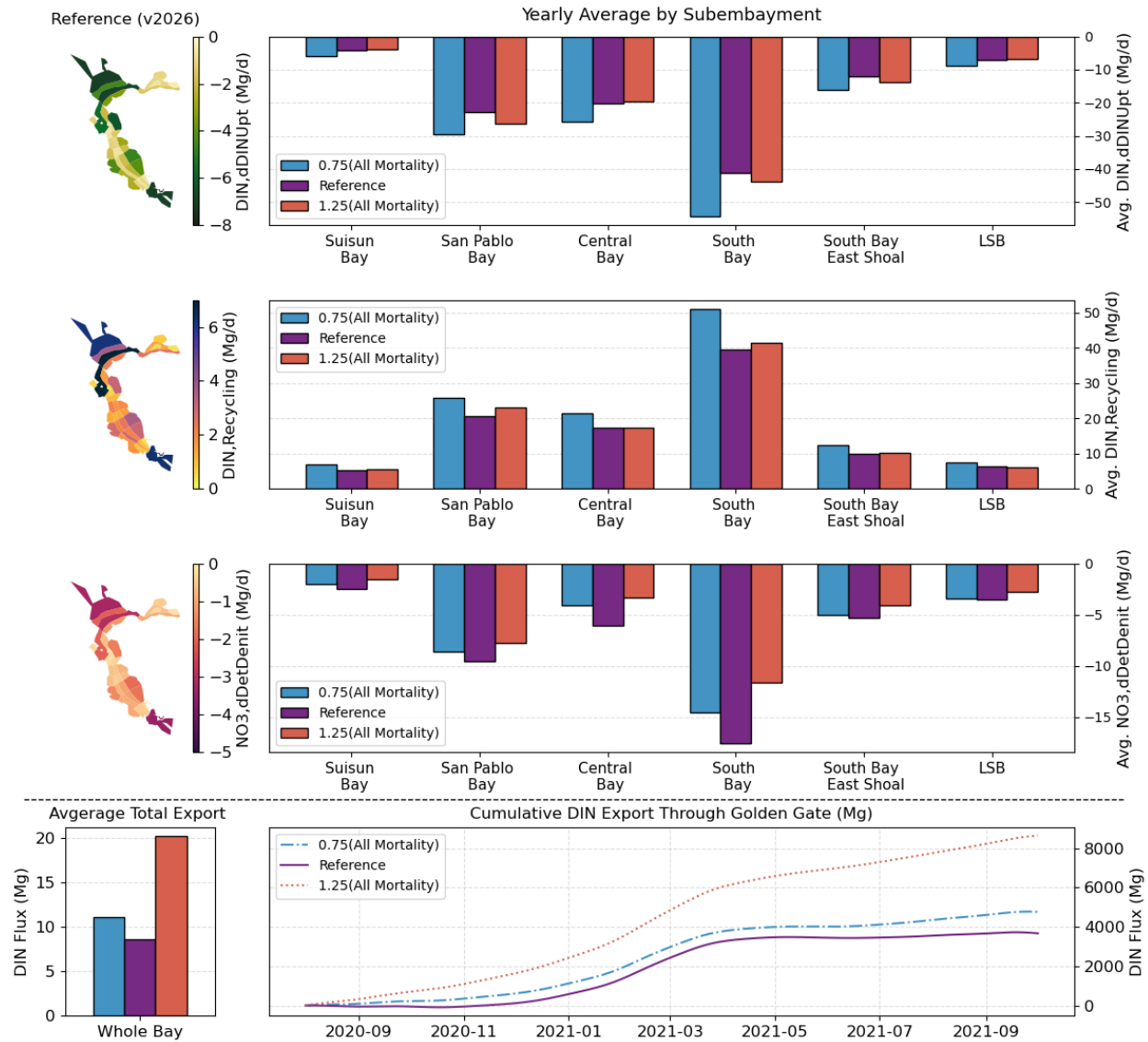


Figure 34: Figure showing control volume analysis for sensitivity to the linear and quadratic mortality rate (All Mortality) for the whole bay for yearly average dissolved inorganic nitrogen uptake rate (top row), yearly average DIN recycling (second row), and yearly average denitrification (third row). The left column shows a map of the reference model output for the entire water year. The right column shows a bar chart comparing the reference solution to two representative model sensitivity experiments. The bottom row shows the year average (left) and cumulative (right) DIN export through the Golden Gate. All three runs used the aggregated grid.

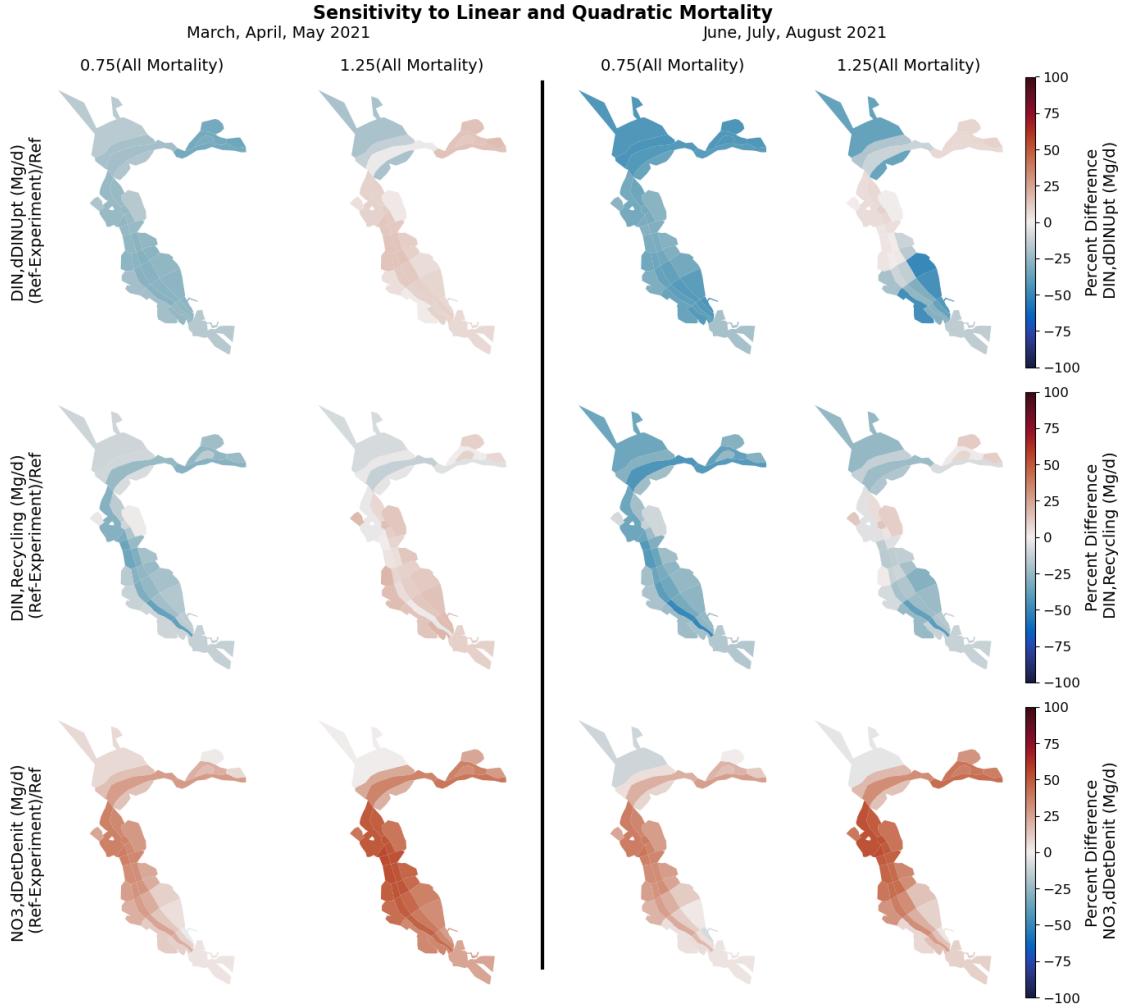


Figure 35: Figure showing control volumes for the whole bay for an experiment focusing on the linear (Mort0) and quadratic (QMort) mortality rates (together: All Mortality) using the aggregated grid. The figures show the percent difference between the reference case and experiments with 0.75·All Mortality (left minor column) and 1.25·All Mortality (right minor column) of March-April-May (left major column) and June-July-August (right major column) average DIN uptake (top row), average DIN recycling (middle row), and average denitrification in the sediment (bottom row). Positive (negative) values, shown in red (blue), indicate regions of the domain where the reference state has larger (smaller) magnitude than the results in the experiment.

5.2.5 Sensitivity to Reaeration Coefficient

As discussed in Section 3.6, for the estuarine reaeration formulation selected, the only adjustable parameter is $k_{l, \text{rear}0}$ which controls the effect of wind on DO transfer at the air-water interface. A sensitivity analysis is performed by varying $k_{l, \text{rear}0}$. The results are presented for three values: as 1) $k_{l, \text{rear}0} = 0.25$, 2) $k_{l, \text{rear}0} = 1.0$ which is the default value in DELWAQ and 3) $k_{l, \text{rear}0} = 4.0$. $k_{l, \text{rear}0}$ is a linear multiplier to a constant coefficient of 0.065 on the square of the wind speed. Thus, sensitivity range evaluates scaling the square of wind speed by a factor 1/4 or 4 relative to the reference simulation. Reducing the scaling factor would be expected to reduce the wind efficiency in mixing at the air-water interface, and vice-versa.

The time series results at channel and South Bay transect locations are shown in Figures 36 and 37, respectively. These figures illustrate that increasing or reducing the scaling factor has different effects on the water column DO depending on whether the water column is already supersaturated, and whether there is greater oxygen demand in the water column. For instance, in spring and summer months, when there is higher primary production, reducing the scaling factor causes high levels of supersaturation, and following the crash of the spring bloom there are parts of the Bay (near SF Airport and Roe Island) where DO levels are lower. These effects are more apparent in the shoals which is evident in the transect time series plots (S2, S6 and S11 are in the channel, while S1 and S4 are on the western and eastern shoals of South Bay, respectively in Figure 37).

Maps of monthly-averaged, depth averaged dissolved oxygen concentrations and the differences from the reference simulation are presented in Figures 38 and 39, respectively. These results are consistent with the time series results, showing that reducing the wind effect leads to higher DO levels, particularly in areas of higher primary production (along the Bay Margins) while reducing the wind effect shows the opposite effect. The effects of the perturbations in either direction are relatively minor within the channel, as expected, because the water current rather than wind speed is more dominant in controlling air-water exchanges in the channel.

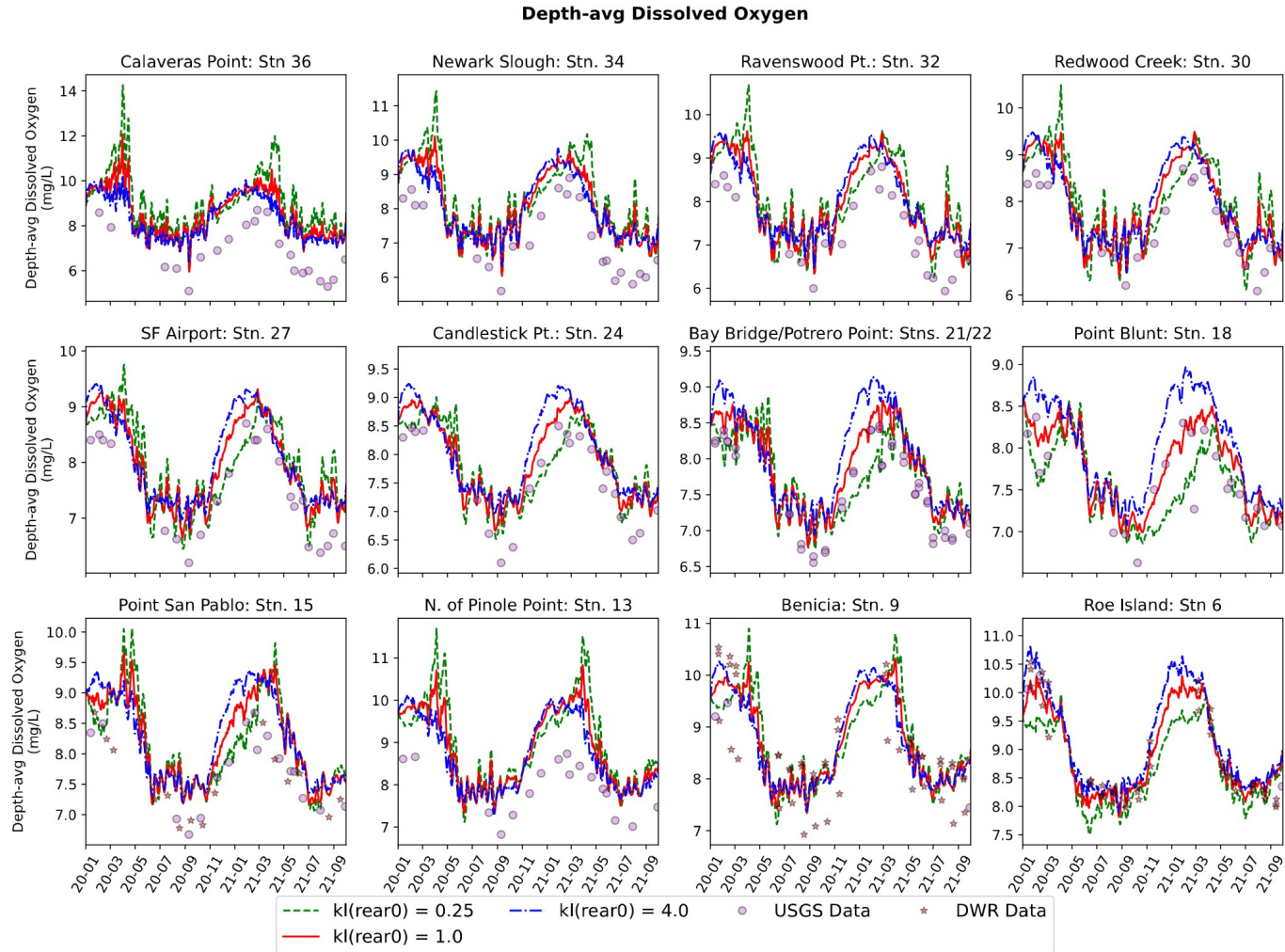


Figure 36: Comparison of daily- and depth-averaged dissolved oxygen from the aggregated grid reference run ($k_{l,\text{rear0}} = 1.0$) and experiments with $0.25 \cdot k_{l,\text{rear0}}$ and $4 \cdot k_{l,\text{rear0}}$ with depth-averaged observations from the USGS Peterson Cruise (circles) and the DWR Environmental Monitoring Program (stars) at 12 long-term monitoring stations, one shown in each subfigure organized from LSB (top left) to Suisun Bay (bottom right).

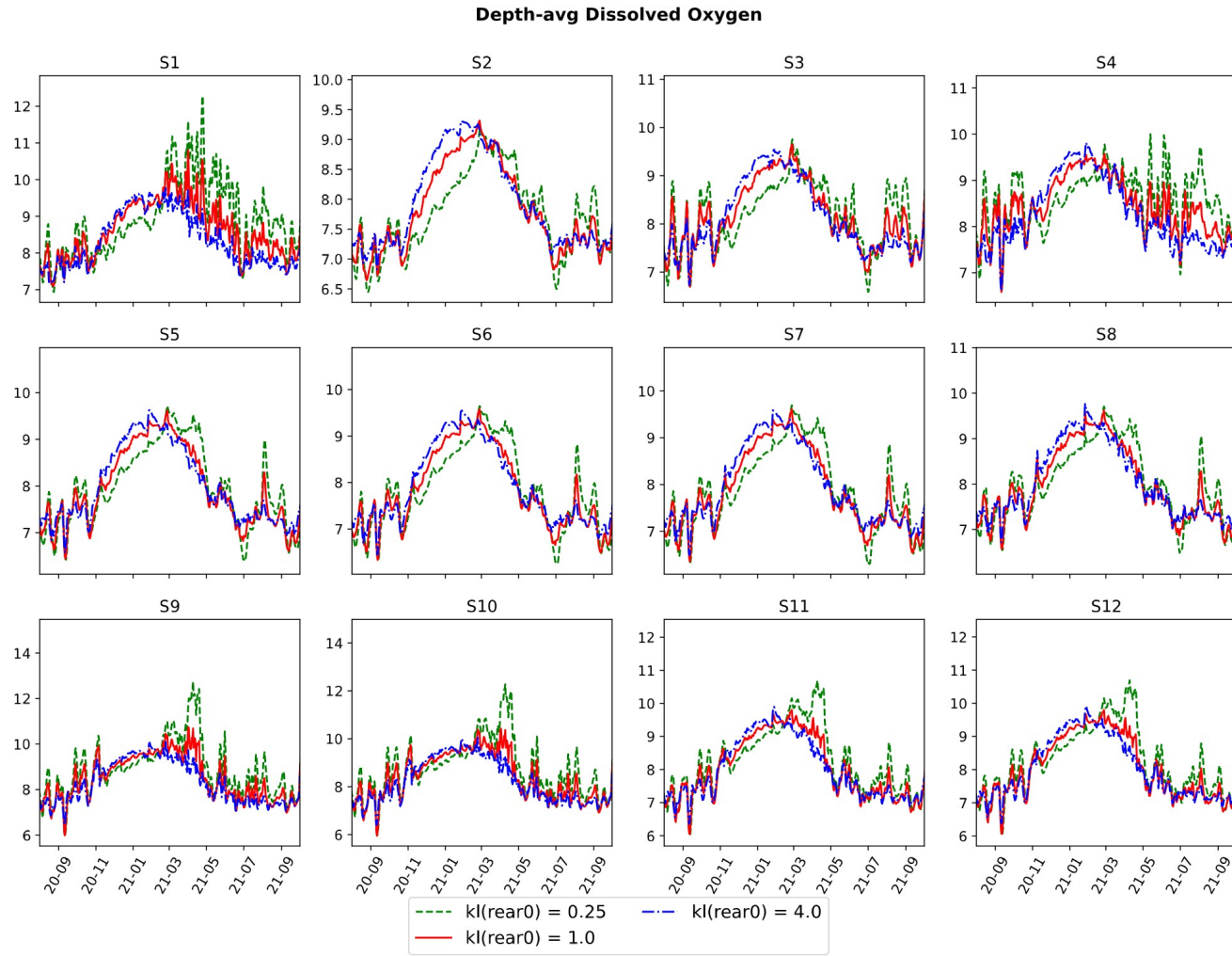
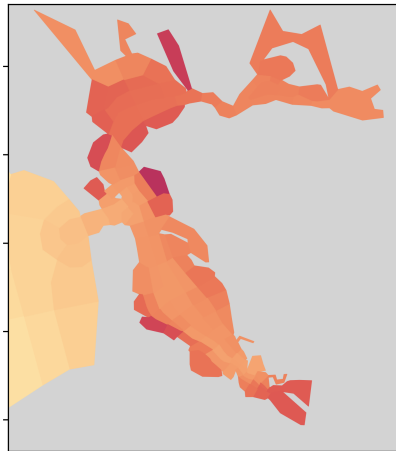


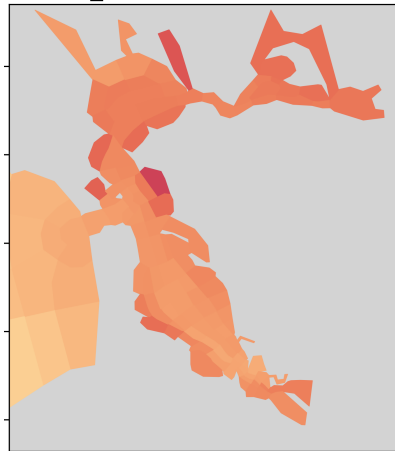
Figure 37: Comparison of daily- and depth-averaged dissolved oxygen from the aggregated grid reference run ($k_{l,\text{rear0}} = 1.0$) and experiments with $0.25 \cdot k_{l,\text{rear0}}$ and $4 \cdot k_{l,\text{rear0}}$ at 12 transect locations across the channel and shoals of South Bay.

DO Absolute Values (monthly-averaged) - May 2021

KL_rear = 0.25 (low reaeration coefficient)



KL_rear = 1 (Base Run)



KL_rear = 4 (high reaeration coefficient)

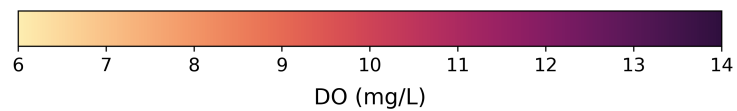
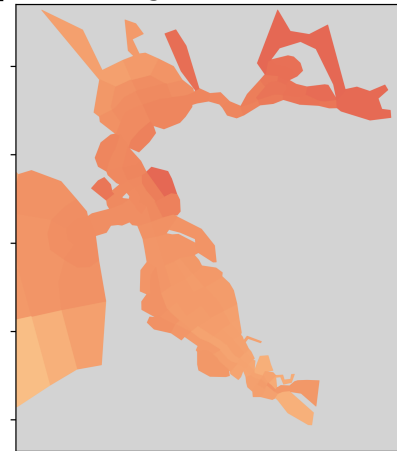


Figure 38: Figure showing depth- and monthly-averaged aggregated grid model output for an experiment focusing on the reaeration coefficient ($k_{l, \text{rear}0}$) from May 2021: $0.25 \cdot k_{l, \text{rear}0}$ (left), the reference run (center), and $4 \cdot k_{l, \text{rear}0}$ (right).

DO Differences from Base Case - May 2021

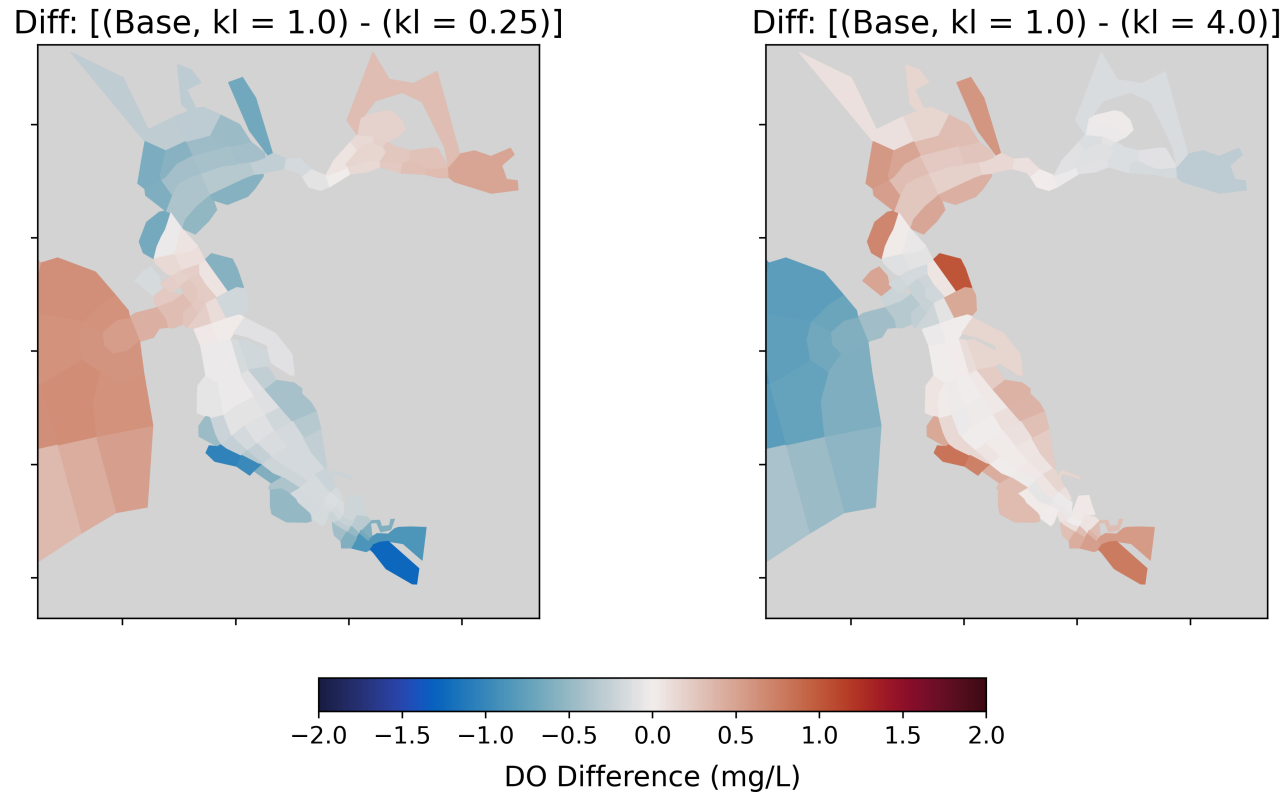


Figure 39: Figure showing depth- and monthly-averaged aggregated grid model output for an experiment focusing on the reaeration coefficient ($k_{l,\text{rear0}}$) from May 2021. Difference between experiments with $0.25 \cdot k_{l,\text{rear0}}$ (left) and $4 \cdot k_{l,\text{rear0}}$ (right) and the reference run. Positive (negative) values, shown in red (blue), indicate regions of the domain where the experiment has larger (smaller) magnitude than the reference case.

5.3 Summary and Discussion

5.3.1 Summary of Model Performance

Overall the model does a reasonably well in capturing the seasonal peaks in chlorophyll, the seasonal changes in DIN uptake and recovery, and the seasonal changes in DO. This assessment is based on reviewing multiple lines of evidence including USGS Peterson Cruise data, high frequency measurements in both the channel and shoals, and mapping cruises and remotely sensed data that were not used in previous modeling efforts to assess performance. The modeled GPP is generally consistent with estimates based on observed discrete as well as high-frequency chlorophyll measurements. The model predicted SOD and denitrification fluxes were generally consistent over space and time relative to sediment flux estimates derived from sediment incubations on cores collected from South Bay and LSB during the model simulation period. Modeled denitrification fluxes were generally higher than the measurements, however, at a subembayment level, modeled losses of nitrogen were comparable to the estimates derived in the sediment flux study.

The model appears to be overestimating chlorophyll concentrations in summer months, particularly in areas farther from the shoals in both South Bay as well as in San Pablo Bay. The model also needs further refinements to improve the predictions of DO in mid-to-late summer, when the model tends to overestimate DO levels. Model performance on phosphorus and silica have not been a major focus of this work, and will be in future phases. While both are important indicators of nutrient cycling processes, neither limit primary production in the bay, and therefore, do not severely limit the utility of the model for developing interpretations of primary production, nitrogen and DO drivers in the Bay.

5.3.2 Discussion

The current model simulation for 2021 provides several useful insights. We are in the process of compiling the control-volume based mass budgets to present a more comprehensive assessment of the primary production, nitrogen cycling and DO conceptual models for the Bay. At a very high level, our modeling results indicate that majority of the primary production occurs in the shoals. This is consistent with the mapping cruises, observations from remotely sensed data, and high frequency data. The observations in the channel primarily reflect the lower trophic level production and nutrient cycling in the shoals and along the bay margins.

Introduction of a simplified denitrification formulation and simulating top-down control through a quadratic mortality formulation reduced the number of degrees of freedom for the model calibration, while still providing the ability to capture key observations. However, these simplifications may also explain some of the differences between the model predictions and observations. For example, clams are known to exert significant top-down control in North Bay. The v2026 model, while successful in eliminating the spurious oscillations observed with the introduction of DEB model to represent clams in previous versions of our models (King et al., 2020, King et al., 2023), it cannot fully represent the spatial variations in top-down control within the SFB system as a model with a full (pelagic and benthic) grazer module could. The inclusion of an active grazer, either as zooplankton and/or as clams, could help to reduce the total chlorophyll concentrations in summer. But this effort would require substantial re-calibration of the model, and would be heavily dependent on having clam and zooplankton biomass levels to properly initialize and/or constrain the model. As discussed in v2023 modeling report (King et al., 2023), while there are monitoring programs that assess clam populations and densities, collecting and processing these data are time consuming and spatial coverage outside of the Delta and North Bay is limited.

The mapping observations also show that areas of elevated chlorophyll-a are more generally

spatially restricted and patchy than the model results would suggest. However, when looking at the biogeochemical data, it is often observed that regions of elevated pH and DO often are more spatially distributed than the area of elevated production itself Killam et al., 2023. This represents a “footprint” of production associated with patches of chlorophyll that are more mobile, due to phytoplankton swimming and aggregation. This clumping behavior of phytoplankton in terms of swimming, reproduction and tidal entrainment may also not be fully captured by the model.

Some of the discrepancies in the modeled to measured chlorophyll may be attributable to the static carbon-to-chlorophyll ratio employed in the model. Previous studies have demonstrated that phytoplankton show seasonally and spatially dynamic concentrations of chlorophyll-a, leading to seasonal variation in carbon to chlorophyll ratio in some temperate estuaries varying from 20 in the winter to over 60 in the summer, as phytoplankton in the winter express proportionally more chlorophyll due to acclimation to lower light levels Jakobsen et al., 2016. When reconciling these ranges in carbon to chlorophyll the model’s comparison to observed chlorophyll is better (Figure A13).

We also suspect here, along with the lack of grazer, the light limitation, which is derived from remotely sensed turbidity, may potentially explain some of the differences between observations and model output. Currently, our model uses derived light extinction coefficients (see Section 3.1), which depend on the availability of remotely sensing data. Gaps in the data are filled by linear interpolation between days, which may leave several important days missing in the light field. Future updates to the model may include using a full sediment transport model to inform light extinction coefficients, but this model is currently in development. Current efforts are also underway to collect vertical profiles of turbidity and light data to supplement the high-frequency mapping data, which will allow for direct validation of photic zone calculations used in the model. With regards to remotely sensed chlorophyll-a, there are currently efforts underway by SFEI and University of California, Santa Cruz (UCSC) partners to improve the relationship between remotely sensed chlorophyll-a and turbidity and direct measurements on the ground through an approach integrating both parameters into a combined algorithm. These improvements will lead to increased precision in $\mu\text{g/L}$ estimates at lower baseline productivity and turbidity conditions outside of those that the Kudela et al., 2024 algorithm was calibrated for.

Notwithstanding these limitations in chlorophyll performance, the current calibration achieves the primary motivation of better representing the observed drawdown of nitrogen during the spring-time bloom and sustained summertime productivity (see Figures 6 and 8). To this extent this calibration is a useful advancement over prior versions of the SFB-BGCM.

The current model calibration accounts for DO sources and sinks, including photosynthetic production by phytoplankton, consumption of oxygen by organic matter remineralization and air-sea oxygen exchange at the surface. In our DELWAQ simulations, reaeration is driven by both wind speed and stream velocity, with DO saturation driven by both temperature and salinity. Currently, the modeled DO concentrations are consistently near saturation (see Figure 20 and 21 for examples), hinting that any oxygen production or drawdown is quickly compensated. The initial hypotheses was that this behavior is primarily driven by an overly efficient reaeration in the model. However, sensitivity analysis by varying the reaeration coefficient (k_l) demonstrated that adjusting the reaeration coefficient does not improve the model’s ability to capture lower DO levels in the summer months. This suggests that internal oxygen sinks may be underrepresented in the model. We are continuing to investigate this issue through increased sediment organic matter retention that can delay oxygen demand to later in the summer, and by rebalancing the sediment remineralization rates.

We expect that some these improvements will also have an effect on silica and phosphorus. However, for phosphorus, we do believe water column production alone may be insufficient to

1354 explain the extent of drawdown and recovery indicated in the data. Some of these processes may
1355 be occurring in the margins at salt ponds and marshes which are not sufficiently resolved in this
1356 model, and may require further refinements to our grid, or a dedicated model to capture these
1357 processes (such as the Lower South Bay model that is being developed as part of the NMS).

1358 Our sensitivity analyses to DO and nitrate changes at the ocean boundary could improve dis-
1359 crepancies between the model and measurements in Central Bay and San Pablo Bay, and portions
1360 of South Bay. SFEI's monitoring team is actively looking for monitoring opportunities in the vicin-
1361 ity of Golden Gate bridge to provide nutrient fluxes, and depth profile of nitrate and DO that
1362 could inform these refinements. Additionally, we also plan to collaborate with the coastal modeling
1363 groups in the region to derive better boundary conditions.

1364 As a closing note, we recognize that model performance for one year does not indicate a calibra-
1365 tion. Results from simulations for WYs 2013, 2017 and 2018 based on the v2026 model parameters
1366 are forthcoming. The next phase of modeling would refine the current calibration across the years
1367 based on these initial results of the hindcast simulations.

References

- Banks, R. B. et al. (1977). “Effect of wind and rain on surface reaeration”. In: *Journal of the Environmental Engineering Division* 103.EE3, pp. 489–504.
- Cerco, C. F. et al. (2017). *The 2017 Chesapeake Bay Water Quality and Sediment Transport Model*. Tech. rep. A Report to the U.S. Environmental Protection Agency Chesapeake Bay Program. Vicksburg, MS: U.S. Army Engineer Research and Development Center.
- Chelsky, Ariella et al. (2026). *San Francisco Bay High-frequency Water Quality Sensor Data*. DOI: 10.6073/PASTA/69B73427D4F9021E2FBCF7B89BBFBF23.
- Cloern, James E (2007). “Habitat connectivity and ecosystem productivity: Implications from a simple model”. In: *The American Naturalist* 169.1, E21–E33.
- Czapla, Kenneth M et al. (2025). “Microbial nitrogen removal in South San Francisco Bay: does it play a role in eutrophication resistance?” In: *Estuaries and Coasts* 48.2, p. 45.
- D-Water Quality Processes Library Description* (2026). Technical Reference Manual 81151. Deltares systems.
- Franks, Peter JS (2002). “NPZ models of plankton dynamics: their construction, coupling to physics, and application”. In: *Journal of Oceanography* 58.2, pp. 379–387.
- Franks, PJS et al. (1986). “Behavior of a simple plankton model with food-level acclimation by herbivores”. In: *Marine Biology* 91.1, pp. 121–129.
- Jakobsen, Hans H et al. (2016). “Carbon-to-chlorophyll ratio for phytoplankton in temperate coastal waters: Seasonal patterns and relationship to nutrients”. In: *Limnology and Oceanography* 61.5, pp. 1853–1868.
- Jassby, Alan D et al. (2002). “Annual primary production: Patterns and mechanisms of change in a nutrient-rich tidal ecosystem”. In: *Limnology and Oceanography* 47.3, pp. 698–712.
- Killam, Daniel et al. (2023). *Characterizing Lateral Variability in South Bay Water Quality to Inform Monitoring Program Design*. Technical Report SFEI Contribution No. 1119. Richmond, CA: San Francisco Estuary Institute.
- Killam, Daniel et al. (2025). *Long-term Forcings: Generating and searching for trends in a whole-bay remote sensed turbidity record 2002-present*. Technical Update SFEI Contribution No. 1277. Richmond, CA: San Francisco Estuary Institute.
- Kimmerer, Wim J. et al. (2014). “Phytoplankton Growth Balanced by Clam and Zooplankton, Grazing and Net Transport into the Low-Salinity Zone of the San Francisco Estuary”. In: *Estuaries and Coasts* 37, pp. 1202–1218.
- King, Allie et al. (2020). *Hydrodynamic and Biogeochemical Model Development for San Francisco Bay*. Technical Update SFEI Contribution No. 1006. Richmond, CA: San Francisco Estuary Institute.
- King, Allie et al. (2023). *San Francisco Bay Numerical Modeling: 2021-2023 Update*. Technical Update SFEI Contribution No. 1113. Richmond, CA: San Francisco Estuary Institute.

- King, Allie et al. (2026). *The SFEI San Francisco Bay Hydrodynamic Model: Performance for Water Level, Velocity, Salinity, and Temperature Over Eleven Years*. Technical Update SFEI Contribution No. 1283. Richmond, CA: San Francisco Estuary Institute.
- Kudela, Raphael M et al. (2024). “Evaluation and refinement of chlorophyll-a algorithms for high-biomass blooms in San Francisco Bay (USA)”. In: *Remote Sensing* 16.6, p. 1103.
- MacCready, Parker et al. (2010). “Advances in estuarine physics”. In: *Annual review of marine science* 2, pp. 35–58.
- Martinez, M et al. (2021). “Interagency Ecological Program: Discrete water quality monitoring in the Sacramento-San Joaquin Bay-Delta, collected by the Environmental Monitoring Program, 1975–2020.” In: *Environmental Data Initiative*.
- McKay, M.D. et al. (1979). “A comparison of three methods for selecting values of input variables in the analysis of output from a computer code”. In: *Technometrics* 21.2, pp. 239–245.
- Murray, AG et al. (1999). “The analysis of alternative formulations in a simple model of a coastal ecosystem”. In: *Ecological Modelling* 119.2-3, pp. 149–166.
- Nechad, Bouchura et al. (2010). “Calibration and validation of a generic multisensor algorithm for mapping of total suspended matter in turbid waters”. In: *Remote Sensing of Environment* 114.4, pp. 854–866.
- Noland, Katie et al. (2025). *Nutrient Management Strategy Moored Sensor Program Semi-annual Report Apr. 2025 - Oct. 2025*. Semi-Annual Report SFEI Contribution No. 1294. Richmond, CA: San Francisco Estuary Institute.
- O’Connor, D. J. et al. (1956). “The mechanism of reaeration in natural streams”. In: *Journal of the Sanitary Engineering Division* 82.SA6, p. 469.
- Richardson, Emily T et al. (2023). “A simple approach to modeling light attenuation in the Sacramento–San Joaquin Delta using commonly available data”. In: *San Francisco Estuary and Watershed Science* 21.4.
- Sandoval-Belmar, Marco et al. (2026). “Biogeochemical effects of Golden Gate Strait exchange and other land-based inputs to the San Francisco and Monterey Bay coasts”. In: *Journal of Geophysical Research: Oceans* 131.2, e2025JC023221.
- Sarmiento, Jorge L (2013). “Ocean biogeochemical dynamics”. In: *Ocean biogeochemical dynamics*. Princeton university press.
- Sathyendranath, Shubha et al. (2009). “Carbon-to-chlorophyll ratio and growth rate of phytoplankton in the sea”. In: *Marine Ecology Progress Series* 383, pp. 73–84.
- Schraga, Tara S et al. (2017). “Water quality measurements in San Francisco Bay by the US Geological Survey, 1969–2015”. In: *Scientific Data* 4.1, pp. 1–14.
- Sutula, Martha et al. (2023). *Findings and Recommendations of an Expert Panel Evaluating a Physical-Biogeochemical Model Supporting the San Francisco Bay Nutrient Management Strategy: February 2023 Workshop*. Technical Report SCCWRP Technical Report No. 1316 and SFEI Contribution No. 1277. Southern Californai Coastal Water Research Project and San Francisco Estuary Institute.

- 1444 Wynne, Timothy T et al. (2022). “Evaluating the efficacy of five chlorophyll-a algorithms in Chesapeake Bay (USA) for operational monitoring and assessment”. In: *Journal of Marine Science and Engineering* 10.8, p. 1104.
- 1447 Zierdt Smith, E.L. et al. (2023). *Bivalve effects on the food web supporting delta smelt—A long-term study of bivalve recruitment, biomass, and grazing rate patterns with varying freshwater outflow*. Open-file Report 2022-1101. U.S. Geological Survey.

A Supplemental Information for the Hydrodynamic Model

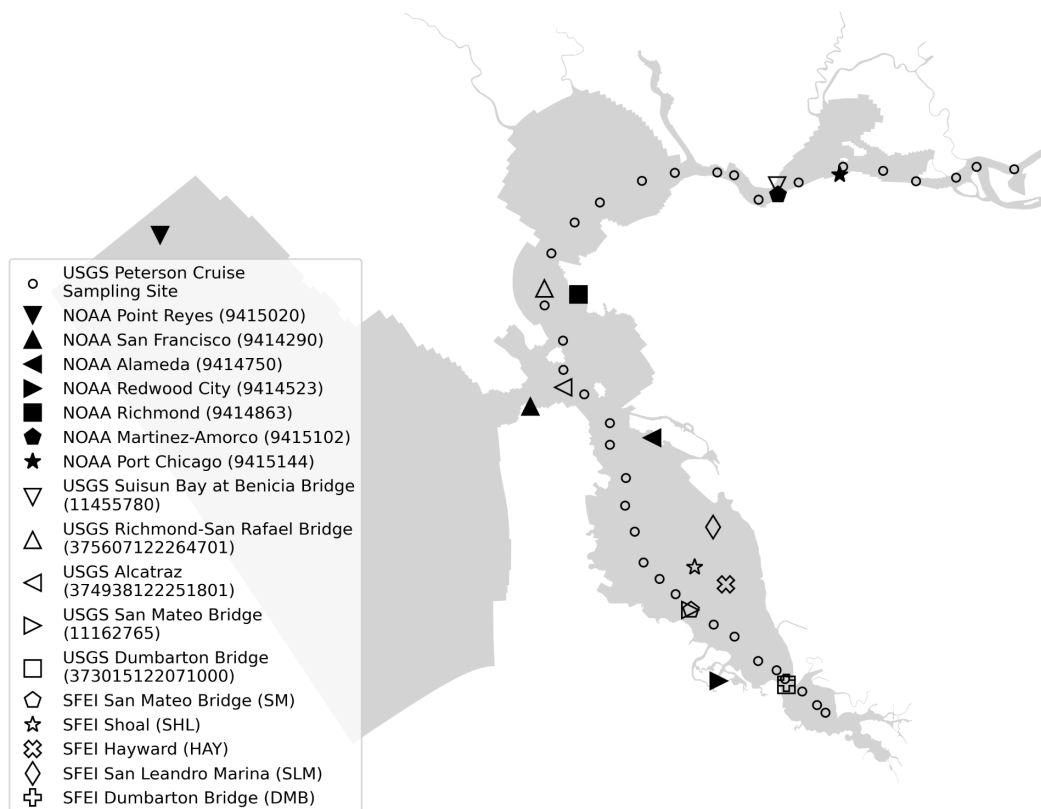
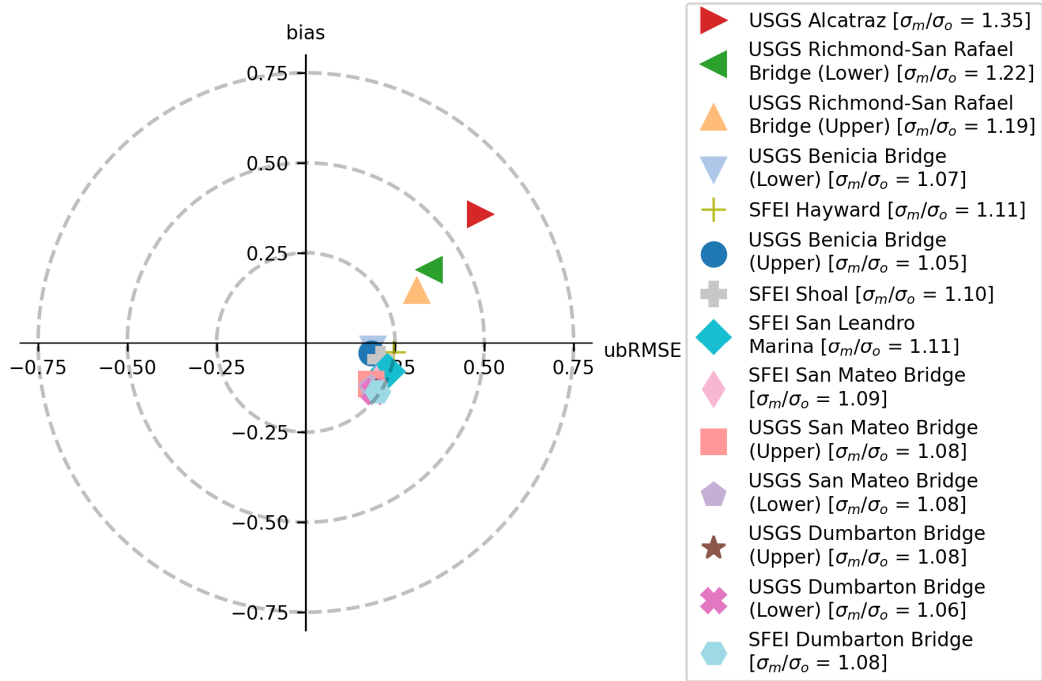


Figure A1: Locations of the moored sensors and other data-collection sites used to evaluate hydrodynamic model performance. From King et al. (2026, Figure 7).



1452

Figure A2: Target diagram of modeled temperature performance at the 14 moored sensors, evaluated on 15-minute data. Bias error and unbiased RMSE are normalized by the standard deviation of the observed temperature, and the legend gives the ratio of modeled to observed standard deviation. Model bias is below 1 °C at all sites. From King et al. (2026, Figure 16).

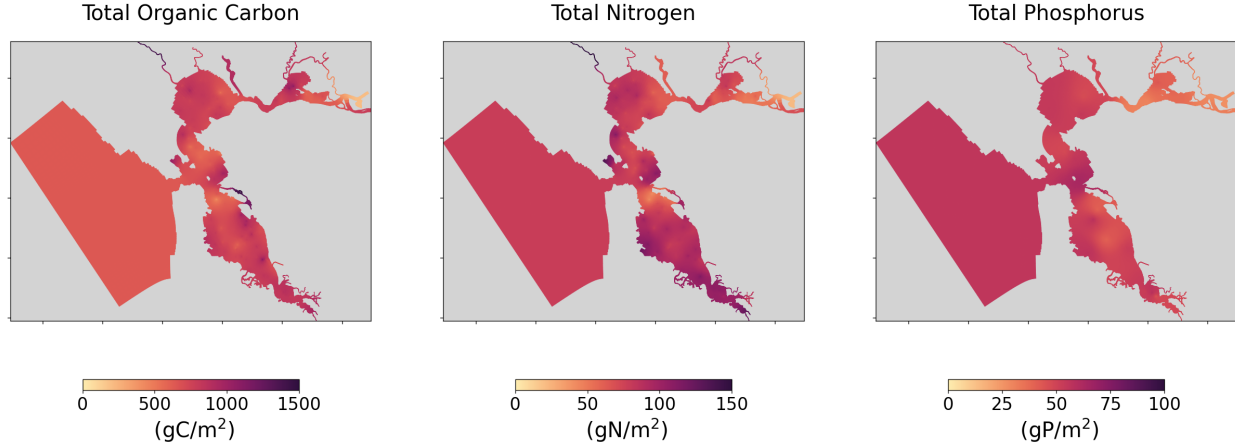


Figure A3: Initial sediment concentrations developed from measured surface sediment data used in previous model runs.

B Supplemental Information for the Biogeochemical Model Base State

In previous runs, the initial concentrations of carbon (C), nitrogen (N), and phosphorus (P) in the sediment bed were estimated using observational data from the Regional Monitoring Program (RMP). The RMP dataset spans 1993 to 2017 and provides measurements of total organic carbon and total nitrogen expressed as a percentage of dry weight (p_{DW}). Because sediment phosphorus measurements are limited to February 2010 and sediment silica data are unavailable, initial conditions for C, N, and P were established by calculating a temporally weighted average centered around October 2012 and spatially interpolating the resulting values across the model grid.

The percentage dry weight measurements were converted into areal mass concentrations (w_{sed} , expressed in $g\ m^{-2}$) using the following equation:

$$w_{sed} = p_{DW} \times (1 - n) \times \rho_{sed} \times h_{sed} \times 1000, \quad (27)$$

where n is the sediment porosity (assumed to be 0.75), ρ_{sed} is the dry density of the sediment ($2500\ kg\ m^{-3}$), h_{sed} is the total thickness of the sediment layer (0.11 m), and 1000 is the unit conversion factor from kilograms to grams. The resulting spatial distributions of the surficial sediment concentrations are illustrated in Figure A3.

For integration into the two-layer DELWAQ (SFB-BGCM) sediment module, the total 11 cm sediment bed was partitioned into an active surficial layer ($S1$, 2 cm thick) and a stable bottom layer ($S2$, 9 cm thick). Following standard assumption for other estuarine sediments (Di Toro 2001), the sediment organic matter was divided into 1% labile detrital matter (DetC/N/P) and 99% refractory detrital matter (OOC/OON/OOP). For silica, spatially uniform initial concentrations of $3\ g\ m^{-2}$ and $300\ g\ m^{-2}$ were assigned to the surficial and bottom layers, respectively, during model calibration.

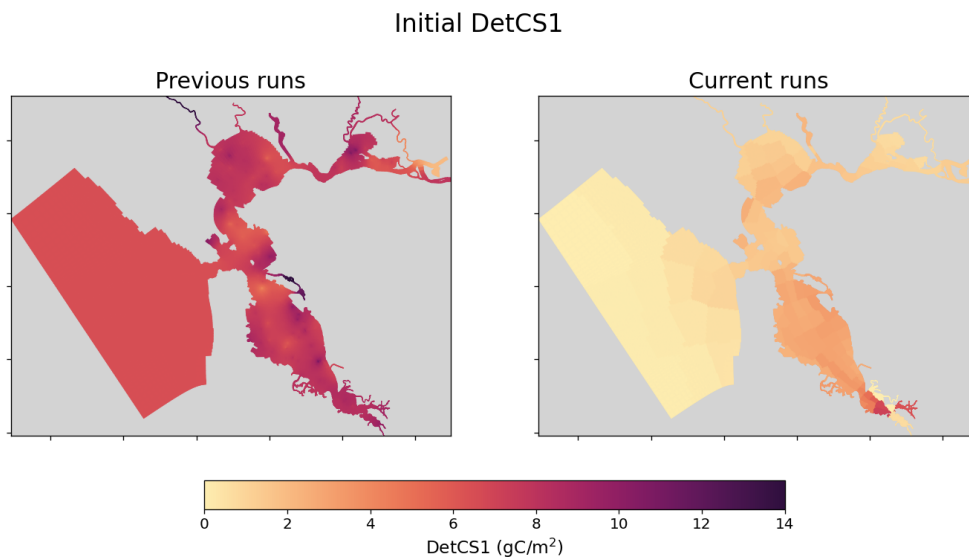


Figure A4: Initial sediment detrital carbon used in old runs and current runs.

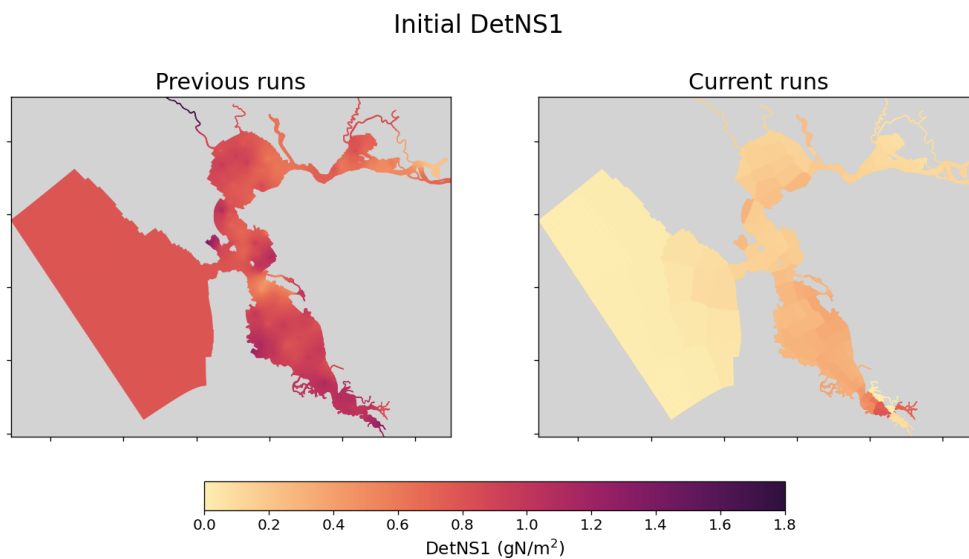


Figure A5: Initial sediment detrital nitrogen used in old runs and current runs.

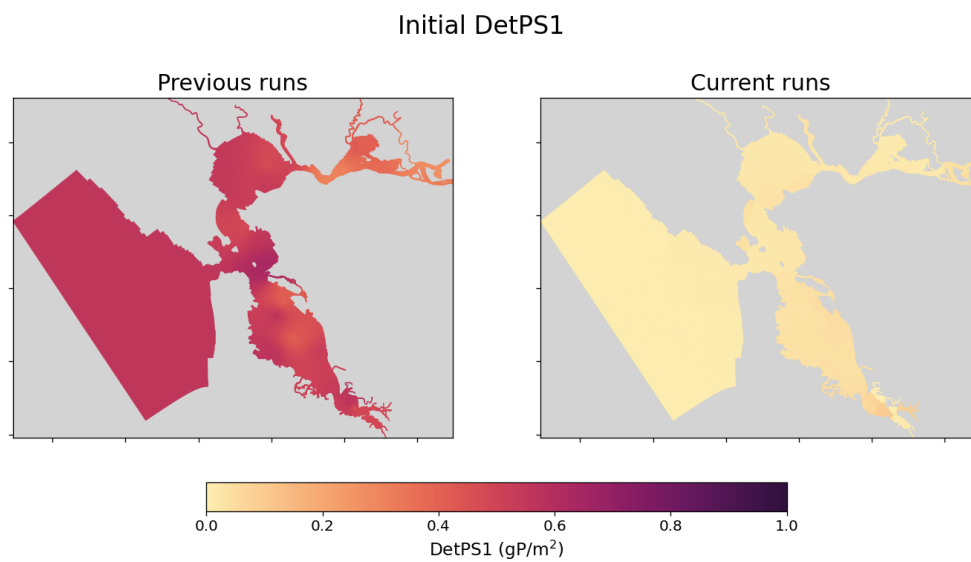


Figure A6: Initial sediment detrital phosphorus used in old runs and current runs.

C Supplemental Results for Biogeochemical Model Validation

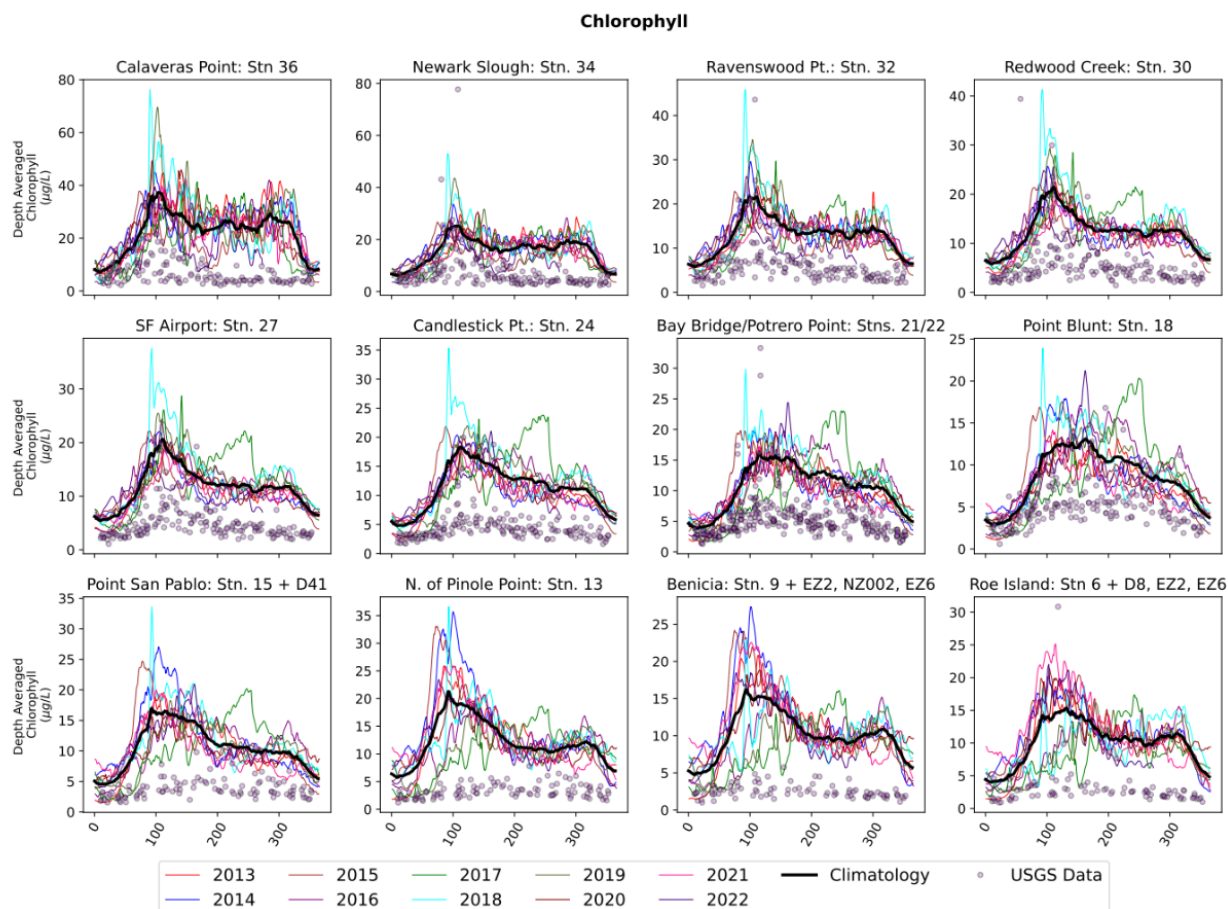


Figure A7: Figure showing model chlorophyll output from a 10 year aggregated grid model run from 1 Aug 2012 to 1 July 2022. The x-axis shows day-of-year with the depth averaged chlorophyll concentration.

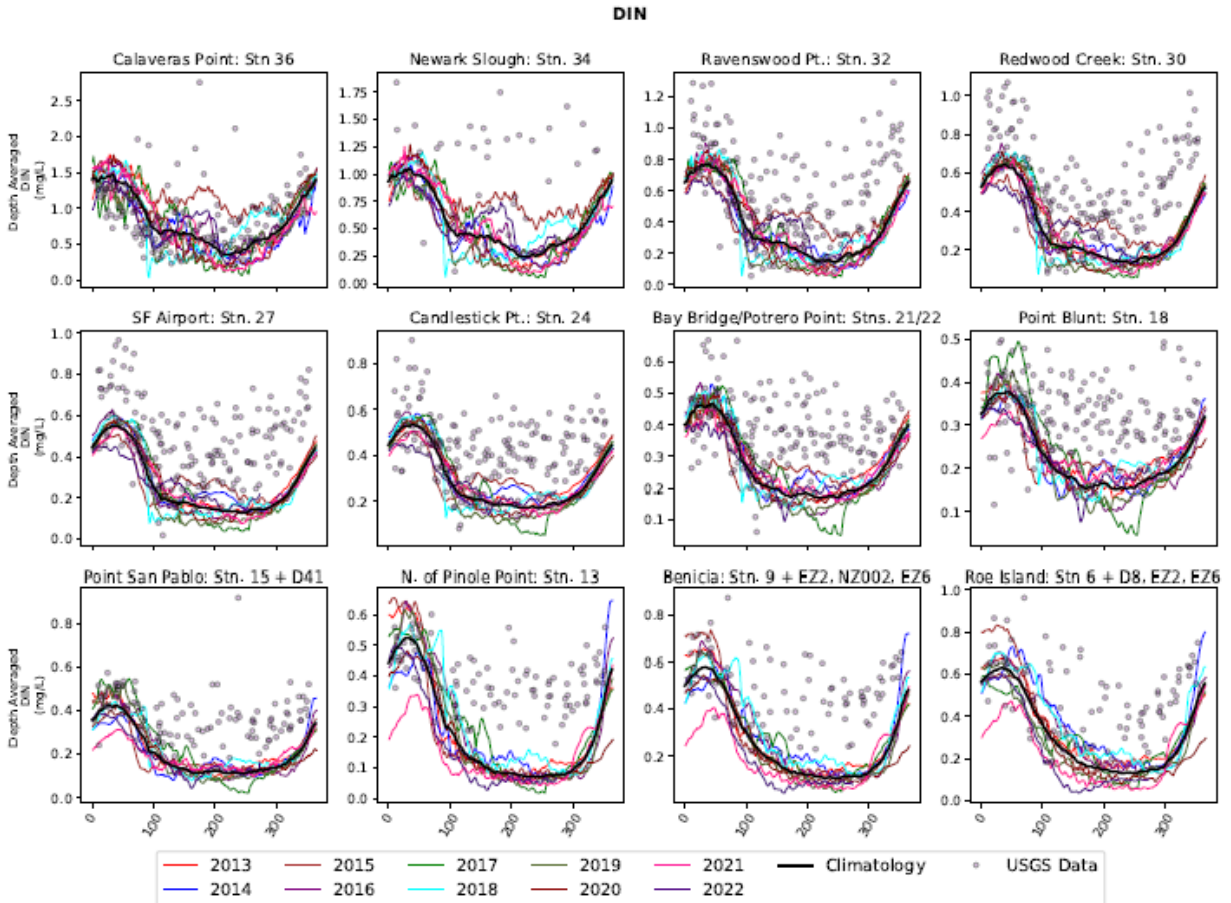


Figure A8: Figure showing model dissolved organic nitrogen (DIN) output from a 10 year aggregated grid model run from 1 Aug 2012 to 1 July 2022. The x-axis shows day-of-year with the depth averaged DIN concentration.

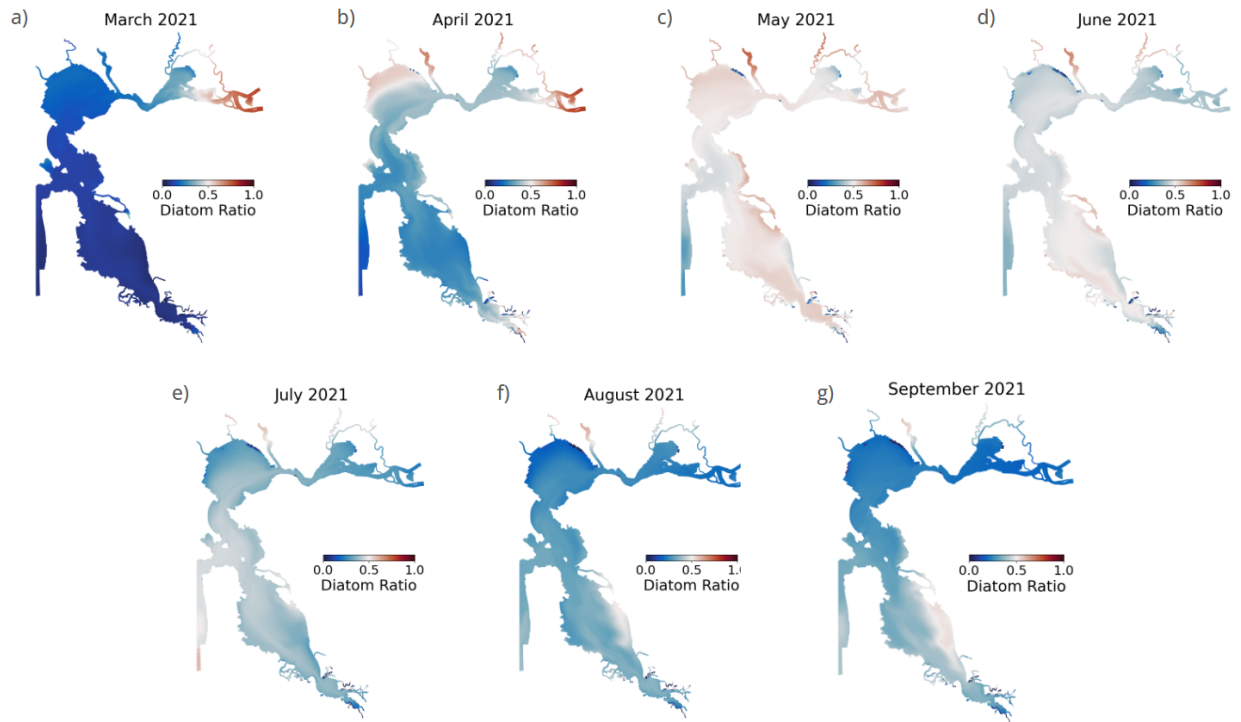
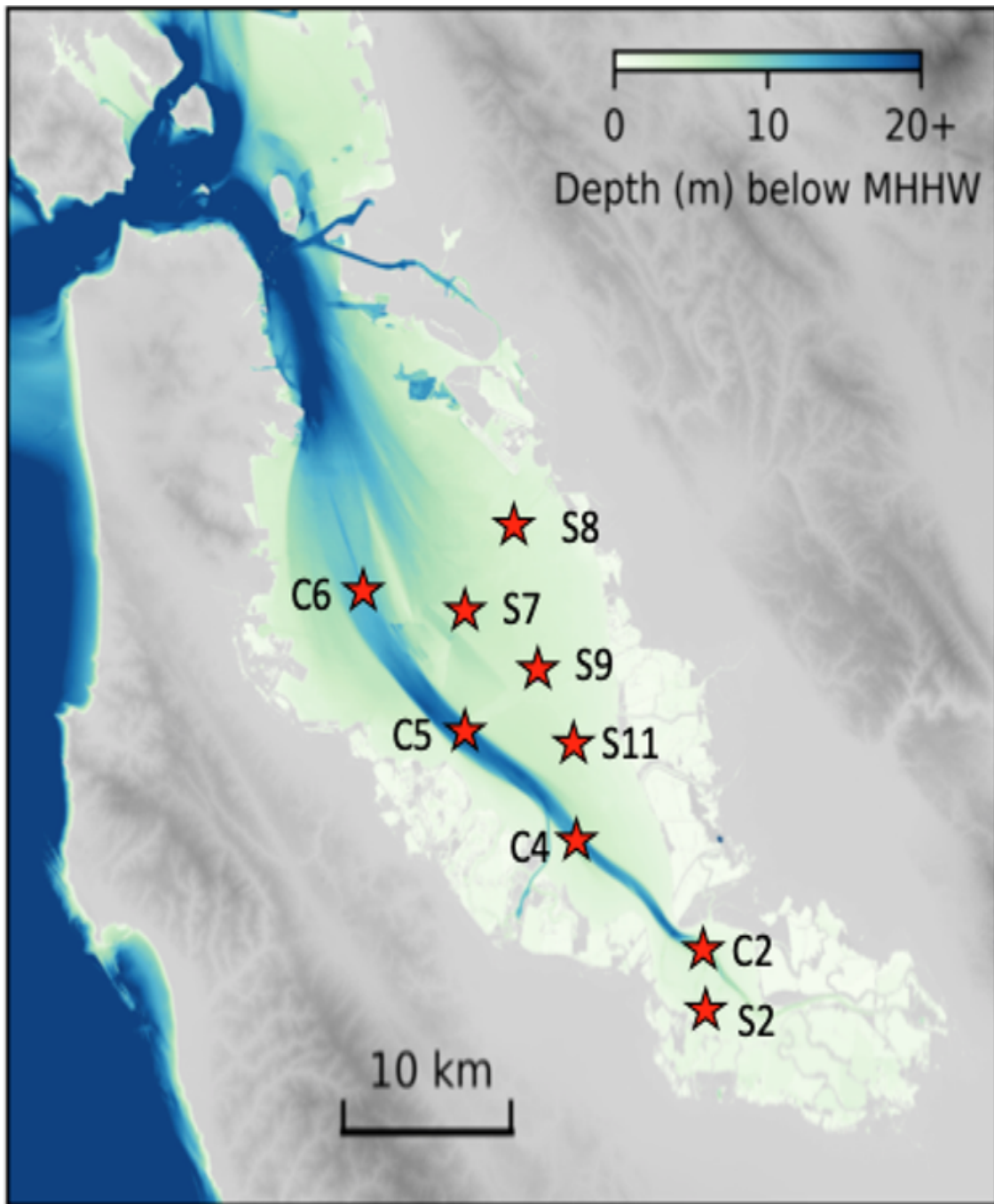


Figure A9: Figure showing the modeled surface diatom ratio for March 2021-September 2021 (a-g). The diatom ratio is the fraction of the biomass of diatoms over the total phytoplankton biomass. Blue colors indicate regions of the model where there are more greens, red colors indicate regions of the model where there are more diatoms. Lighter colors indicate values closer to 50%, darker colors indicate regions where a specific phytoplankton group dominates.



1476

[th!]

Figure A10: Locations where sediment cores were collected in 2020 and 2021 for the sediment flux analyses presented in Czapla et al., 2025. Locations with labels that start with “C” are deep channel cores, and locations that start with “S” are located in the shoals. Note that in this report, stations C2, C4, C5, and C6 correspond to stations 35, 32, 30, and 27 in the USGS Peterson Cruise.

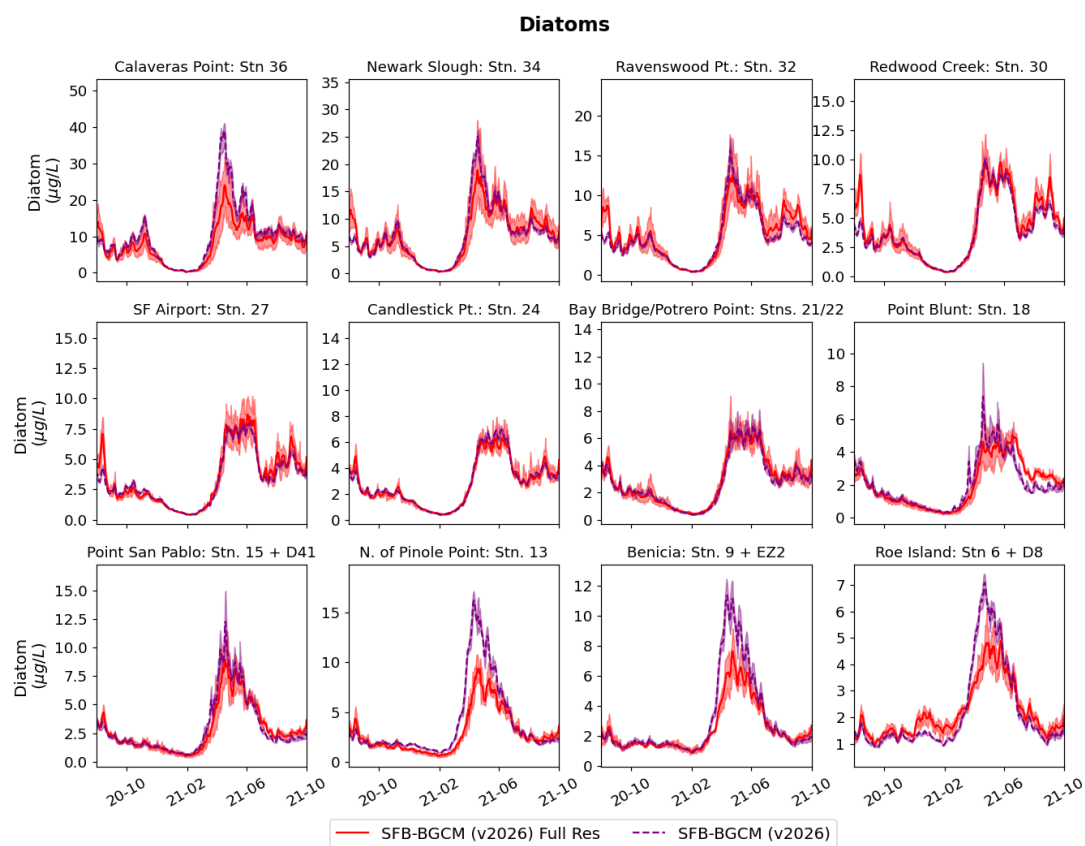


Figure A11: Comparison of chlorophyll from the diatoms in the full resolution (red) and aggregated grid (purple) model output. The shaded areas in show the daily minimum and maximum values showing variability due to tides.

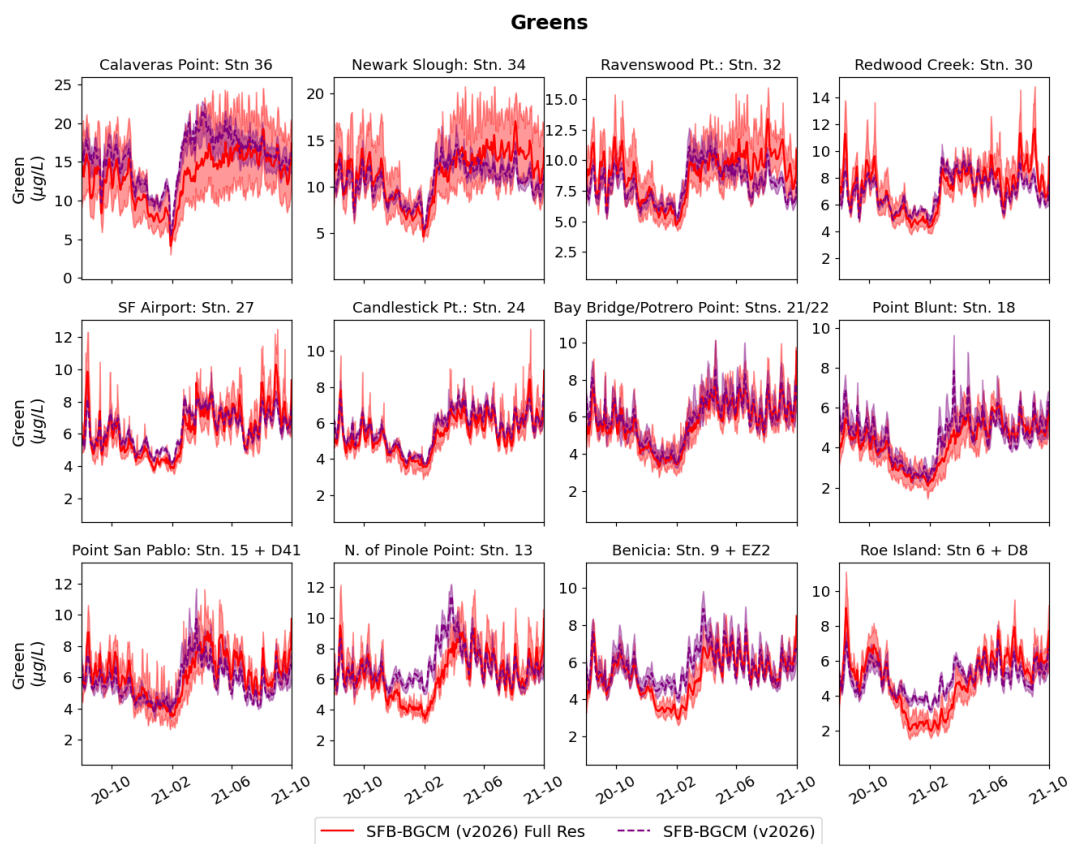


Figure A12: Comparison of chlorophyll from the greens in the full resolution (red) and aggregated grid (purple) model output. The shaded areas in show the daily minimum and maximum values showing variability due to tides.

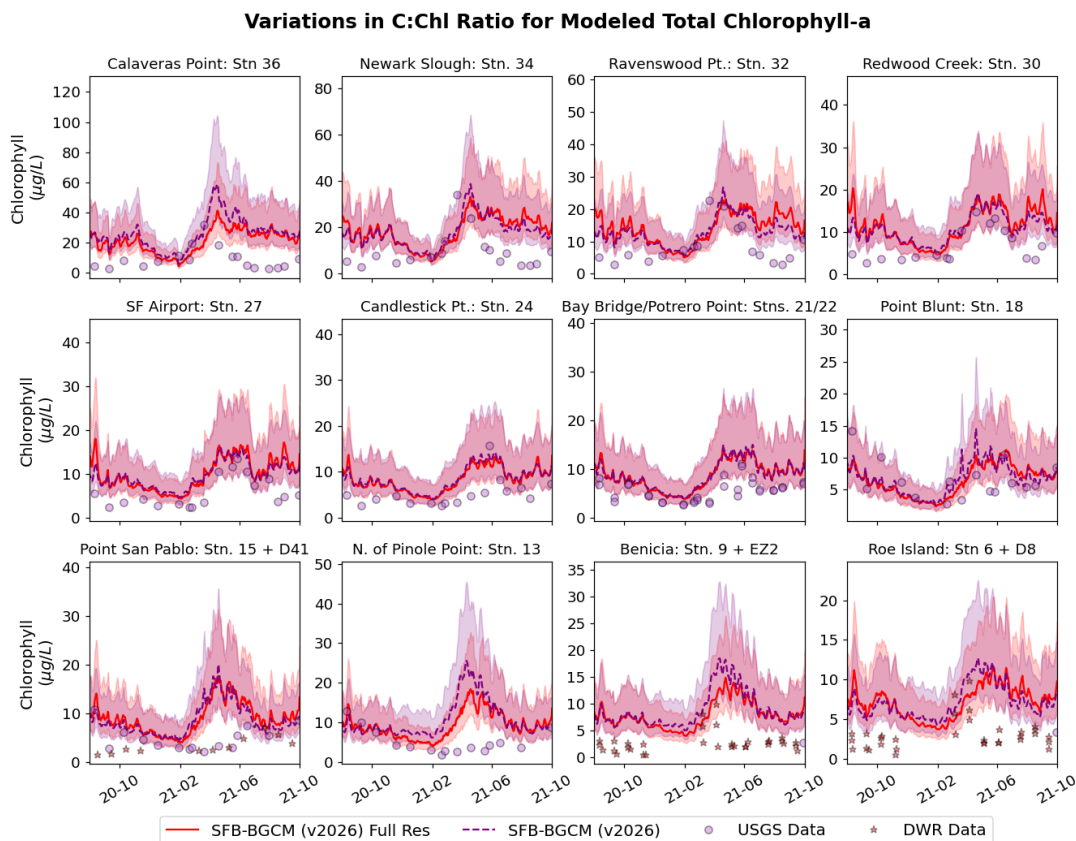


Figure A13: Comparison of chlorophyll concentrations in the full resolution (red) and aggregated grid (purple) model output for the reference ratio, C:Chl=22.5. The shaded areas show variations of carbon to chlorophyll ratio from Sathyendranath et al., 2009 and Jakobsen et al., 2016 ranging from 15 (min) to 40 (max) C:Chl ratio.

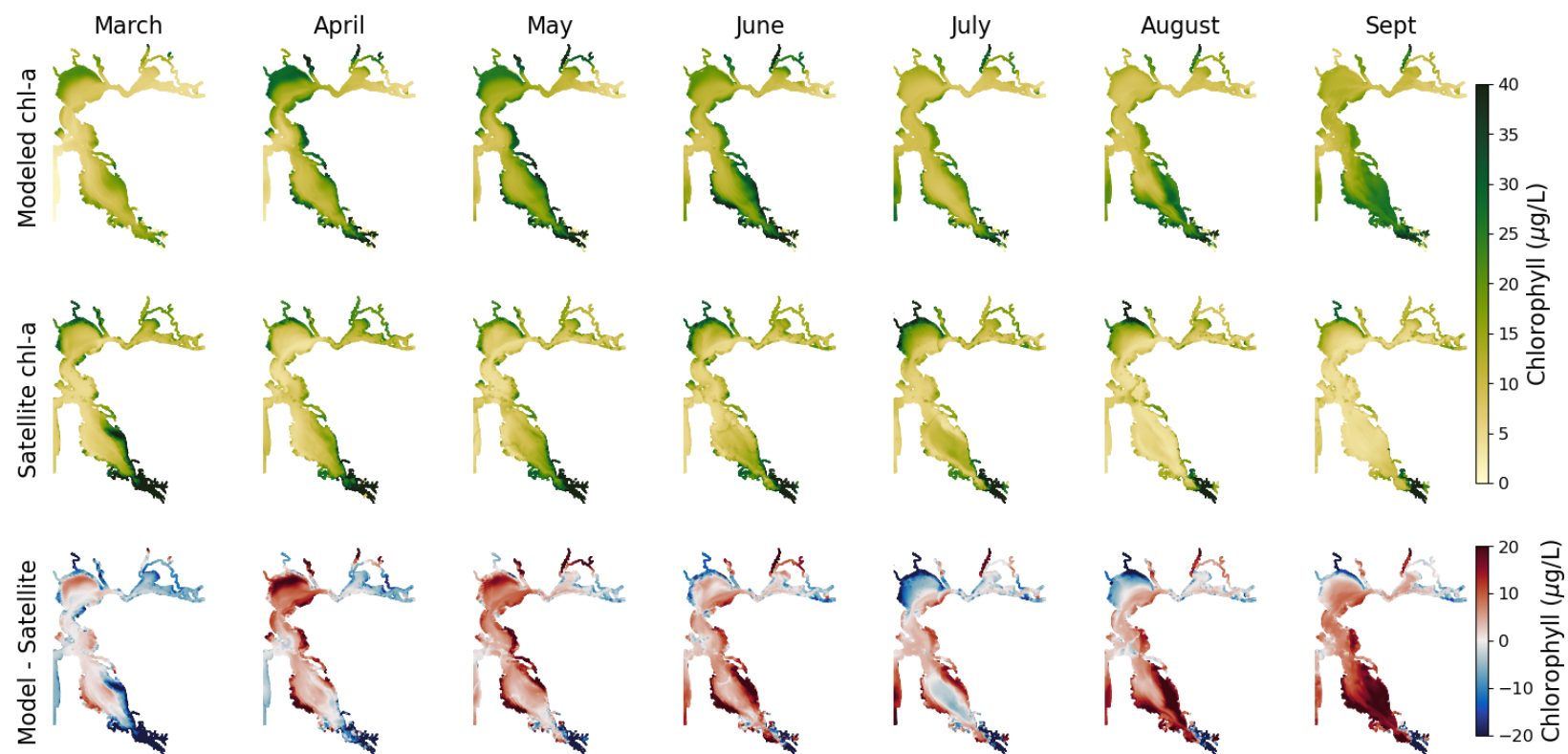


Figure A14: Monthly surface chlorophyll concentrations in comparison to satellite chl-a. The top row shows seasonal averaged modeled surface chlorophyll concentration for March 2021 to September 2021. The middle row shows satellite derived surface chlorophyll concentration for the same time periods. The bottom row shows the difference between seasonal modeled chlorophyll concentration and satellite derived chlorophyll. Positive values (red) indicate that the modeled chlorophyll is greater than the satellite chlorophyll. Negative values (blue) indicate where satellite values are larger

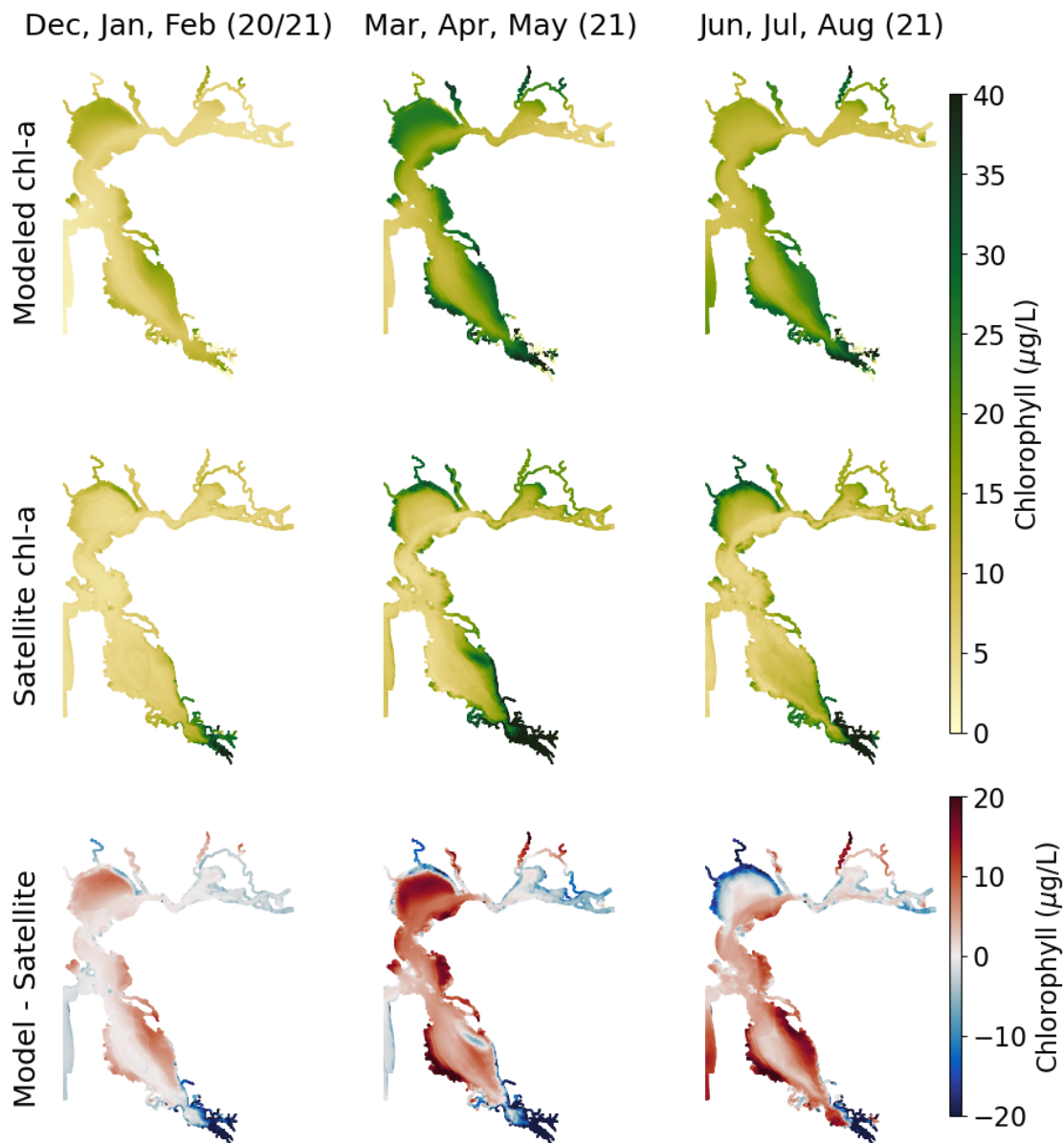


Figure A15: Seasonal surface chlorophyll concentrations in comparison to satellite chl-a. The top row shows seasonal averaged modeled surface chlorophyll concentration for Dec,Jan,Feb 20/21 (left column); Mar,Apr,May 21 (middle column); Jun,Jul,Aug (right column). The middle row shows satellite derived surface chlorophyll concentration for the same time periods. The bottom row shows the difference between seasonal modeled chlorophyll concentration and satellite derived chlorophyll. Positive values (red) indicate that the modeled chlorophyll is greater than the satellite chlorophyll. Negative values (blue) indicate where satellite values are larger.

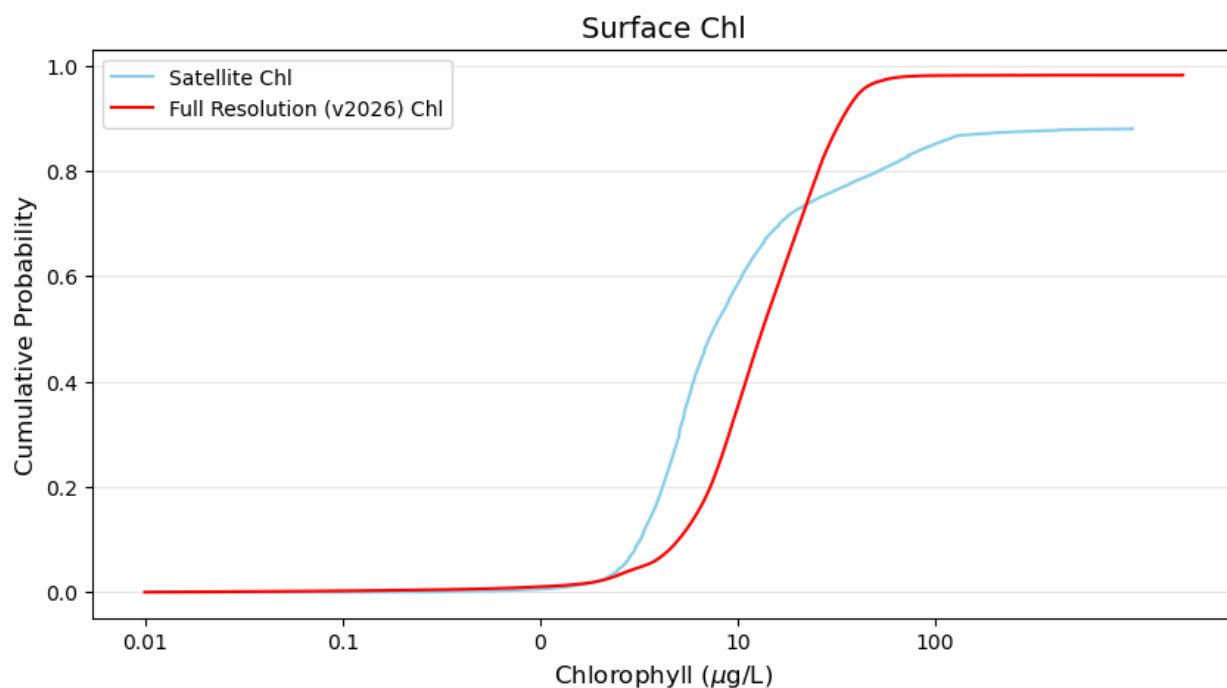


Figure A16: Figure showing the empirical cumulative density function for the modeled total surface chlorophyll concentration (red) and the remote sensing chlorophyll data (blue).

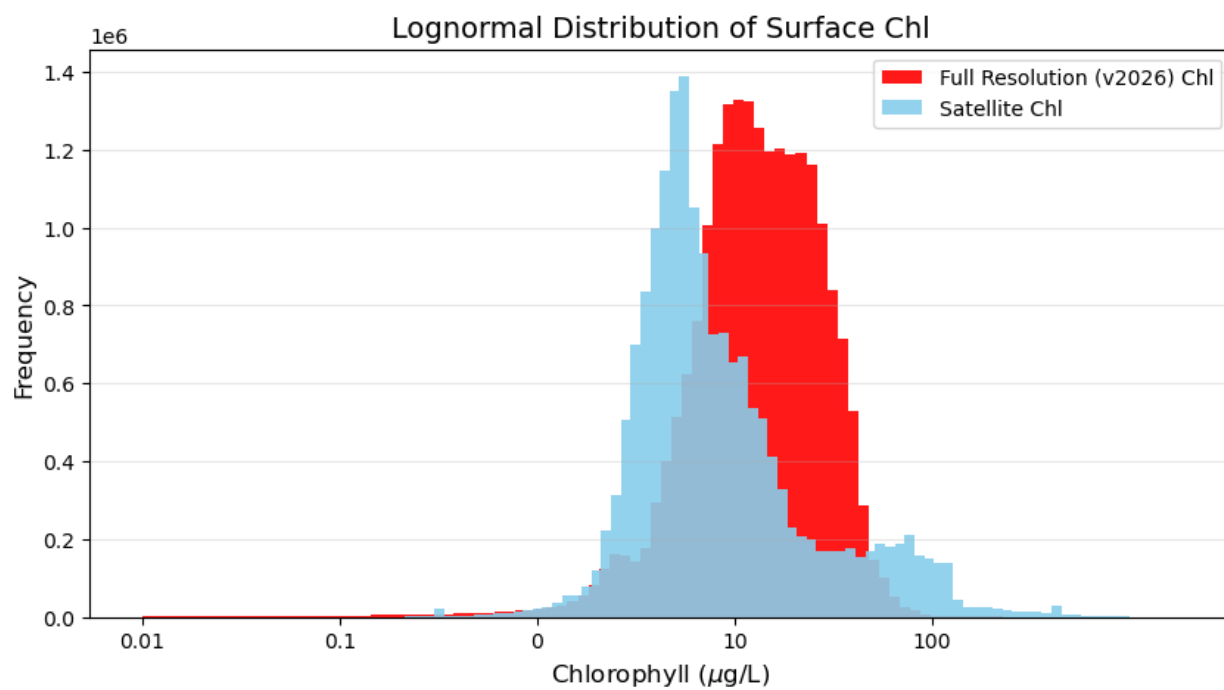


Figure A17: Lognormal distribution of surface chlorophyll concentration from the remote sensing data (blue) and full resolution model (red)

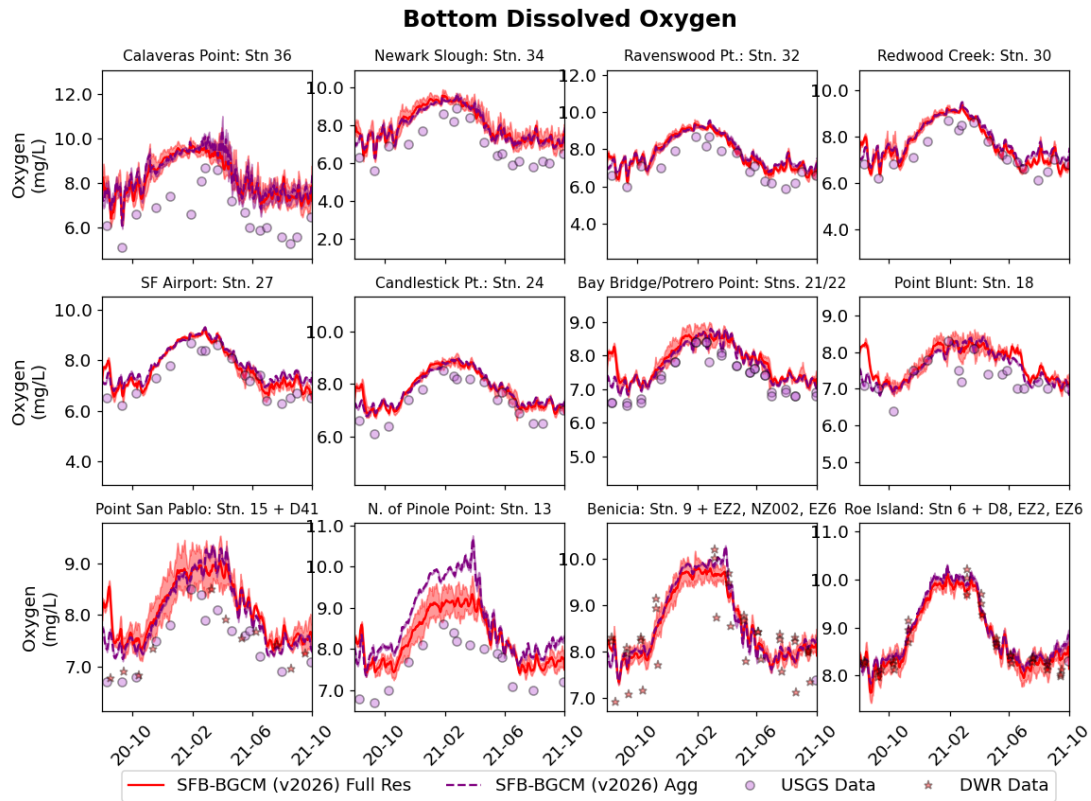


Figure A18: Comparison of bottom dissolved oxygen from the full resolution (red) and aggregated grid (purple) model output with the USGS Peterson Cruise Data (circles) and the DWR Environmental Monitoring Program (stars). The shaded areas in show the daily minimum and maximum values showing variability due to tides.

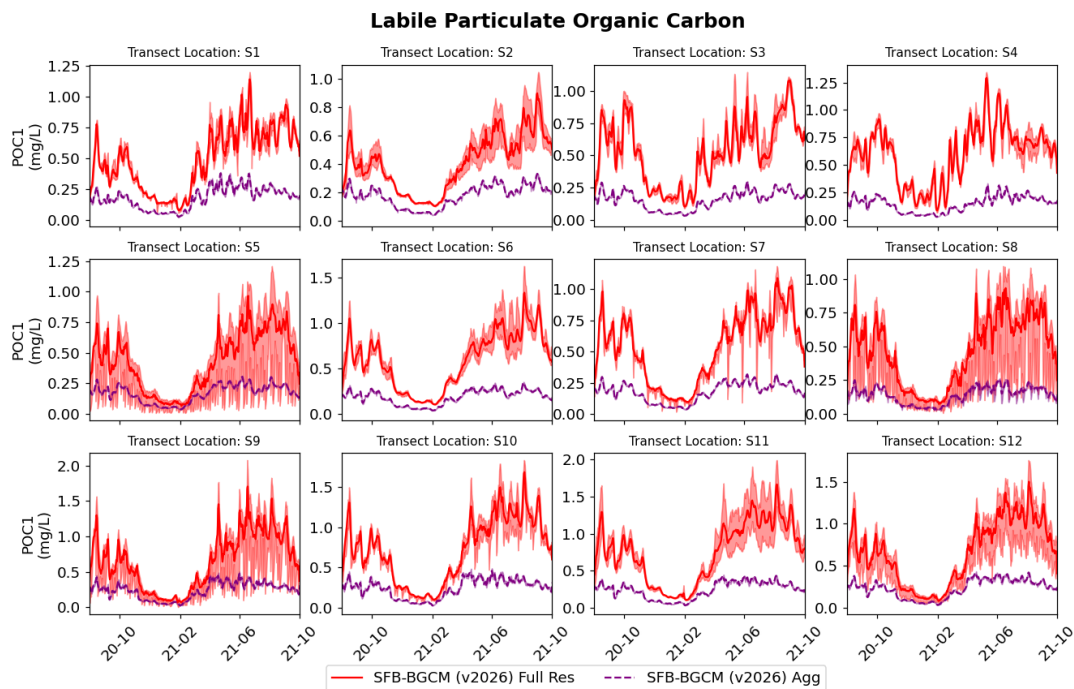


Figure A19: Figures showing the labile water column particulate organic carbon from the full resolution (red) and aggregated grid (purple) model runs. The shaded areas in show the daily minimum and maximum values showing variability due to tides.

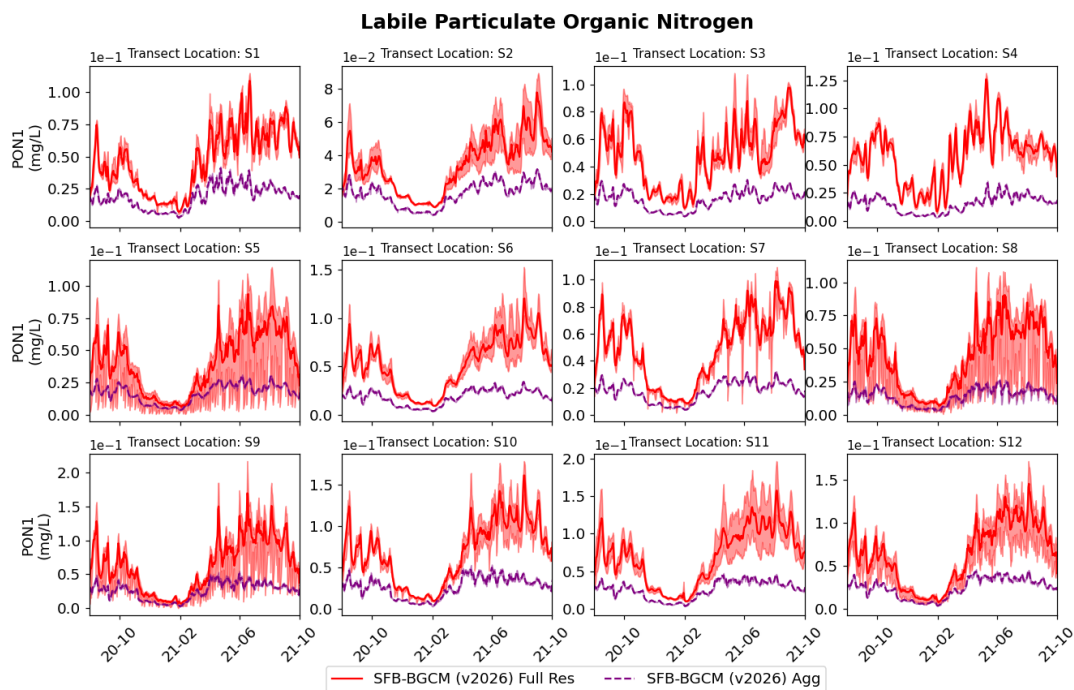


Figure A20: Figures showing the labile water column particulate organic nitrogen from the full resolution (red) and aggregated grid (purple) model runs. The shaded areas in show the daily minimum and maximum values showing variability due to tides.

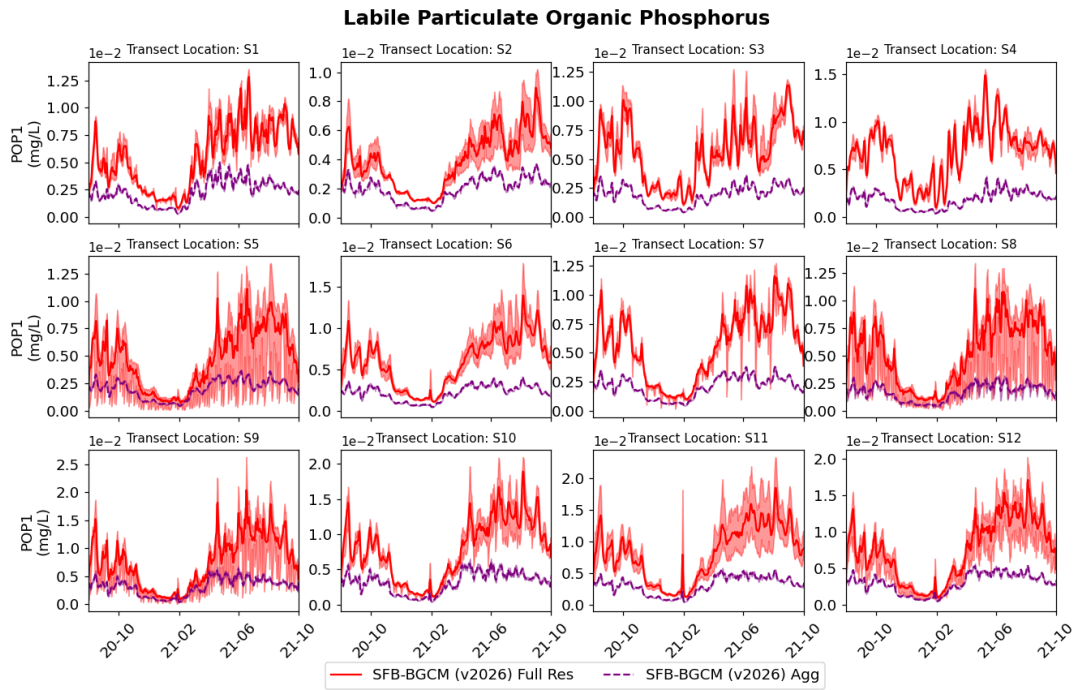


Figure A21: Figures showing the labile water column particulate organic phosphorus from the full resolution (red) and aggregated grid (purple) model runs. The shaded areas in show the daily minimum and maximum values showing variability due to tides.

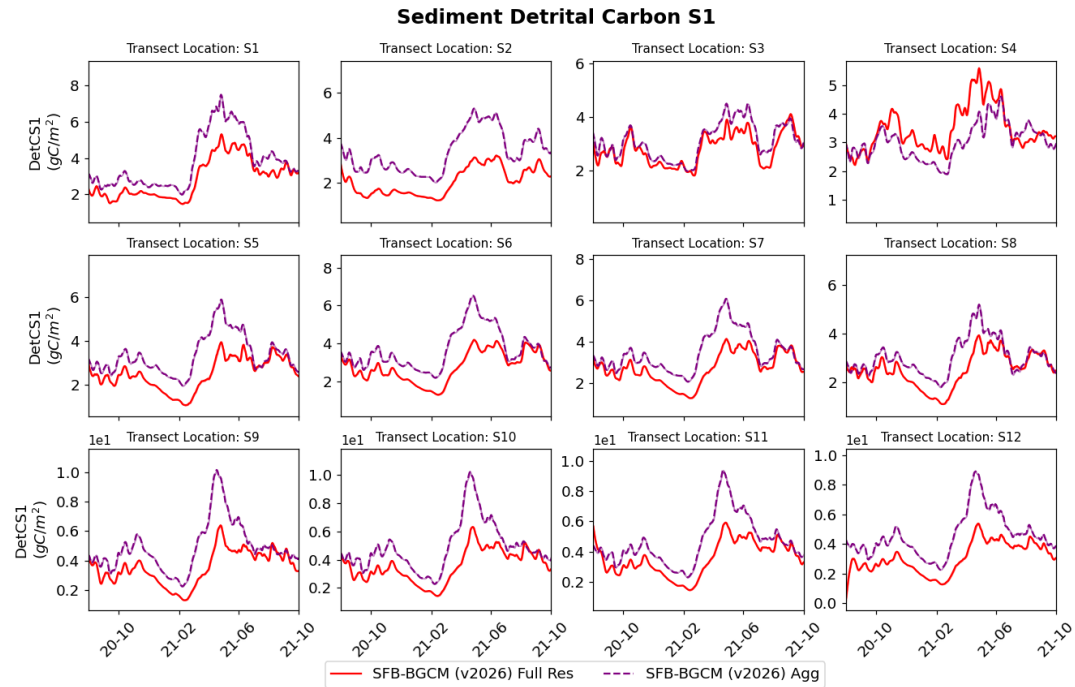


Figure A22: Figures showing the sediment organic carbon in the top (S1) layer from the full resolution (red) and aggregated grid (purple) model runs.

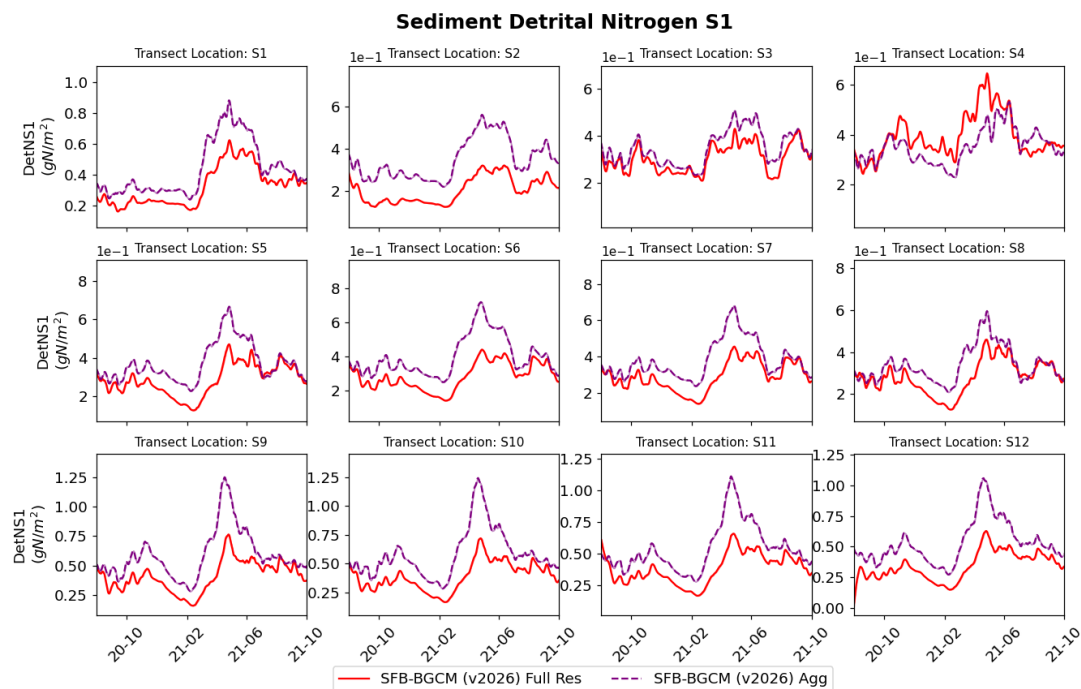


Figure A23: Figures showing the sediment organic nitrogen in the top (S1) layer from the full resolution (red) and aggregated grid (purple) model runs.

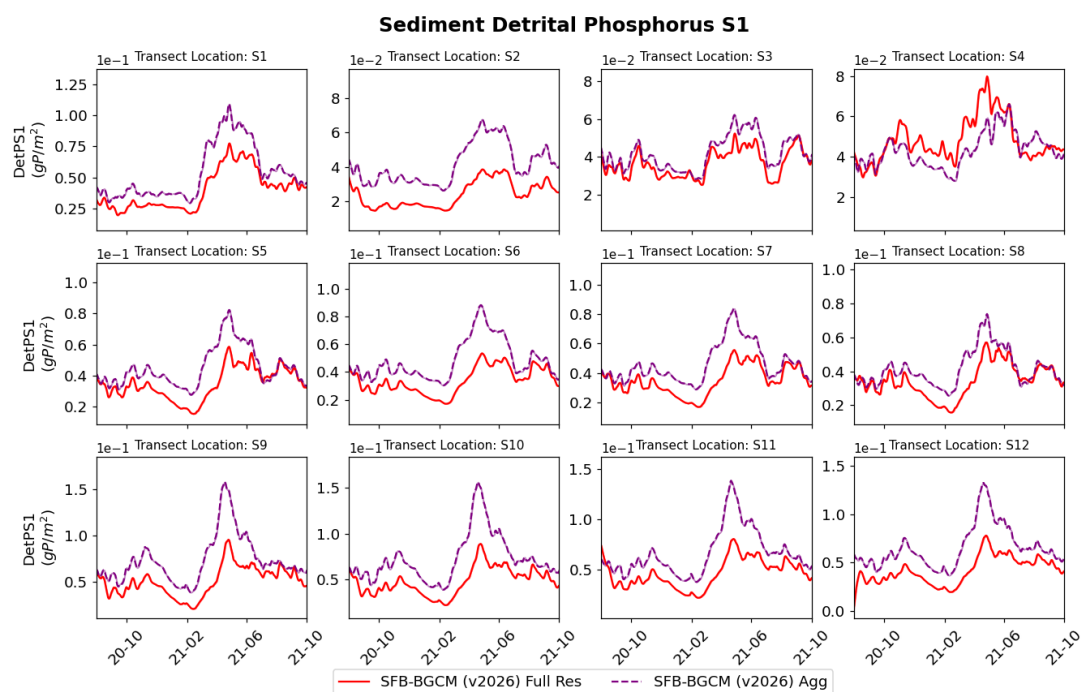


Figure A24: Figures showing the sediment organic phosphorus in the top (S1) layer from the full resolution (red) and aggregated grid (purple) model runs.

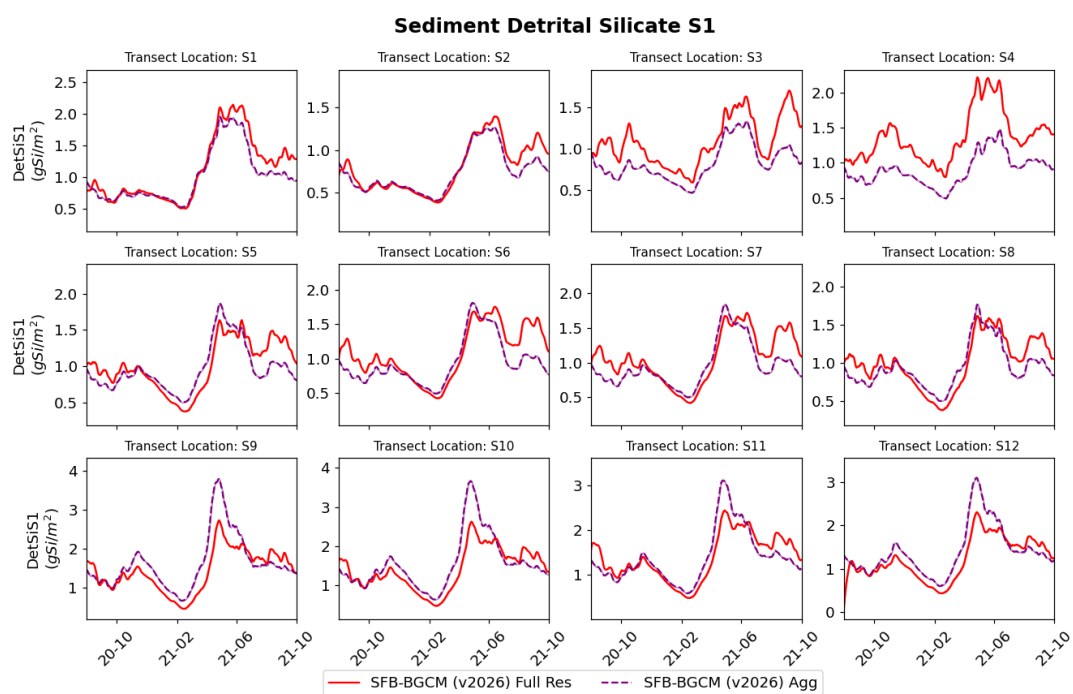


Figure A25: Figures showing the sediment organic phosphorus in the top (S1) layer from the full resolution (red) and aggregated grid (purple) model runs.

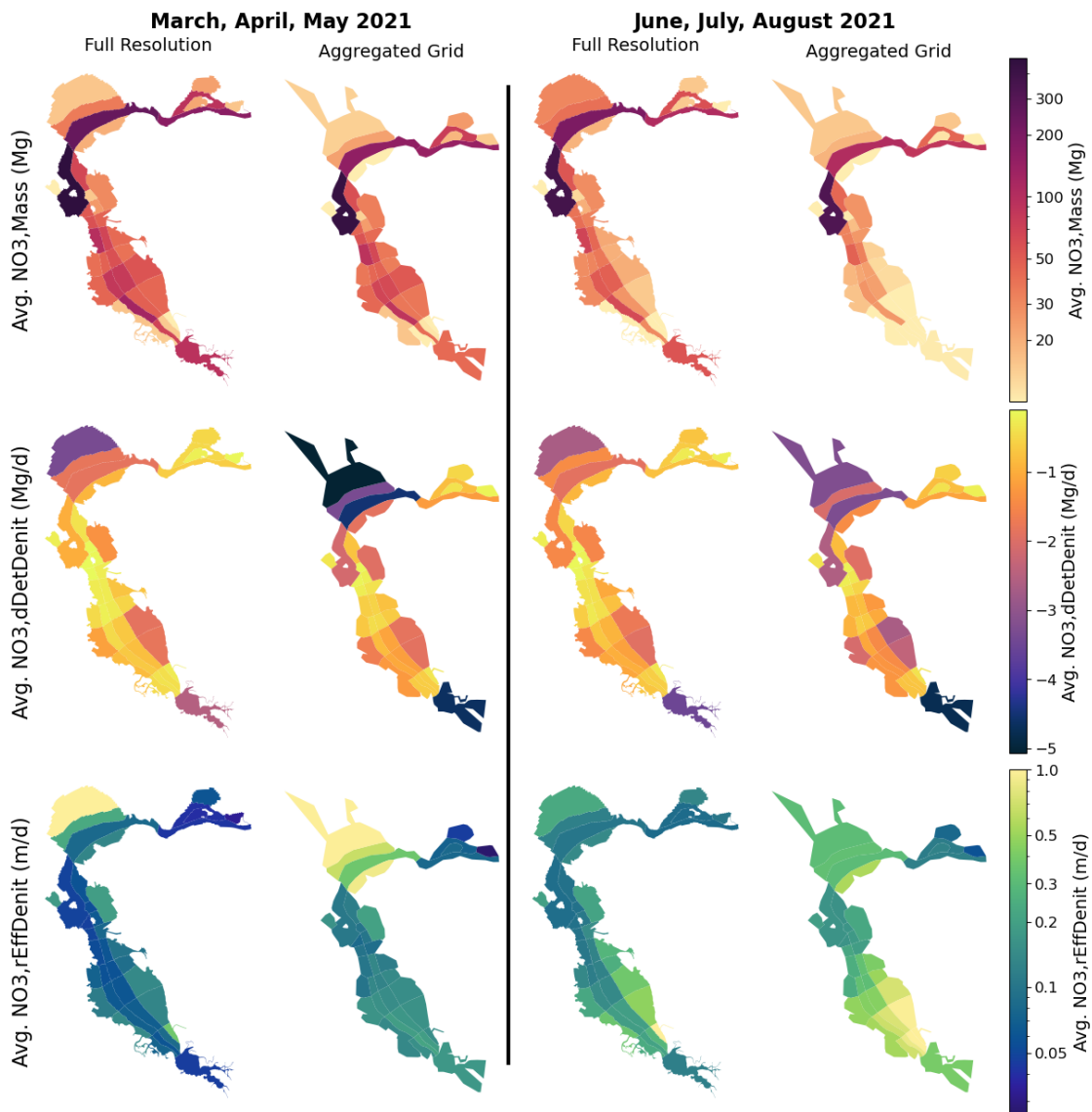


Figure A26: Figure showing control volumes for the Whole Bay showing the average nitrate mass in each volume (top row), average modeled denitrification flux in each volume (center row), and the effective denitrification rate (bottom row). The left major column shows the model output for the full resolution (left) and aggregated grid (right) for March, April, May 2021. The right major column shows the model output for June, July, August 2021.

D Supplemental Results for Sensitivity Analysis

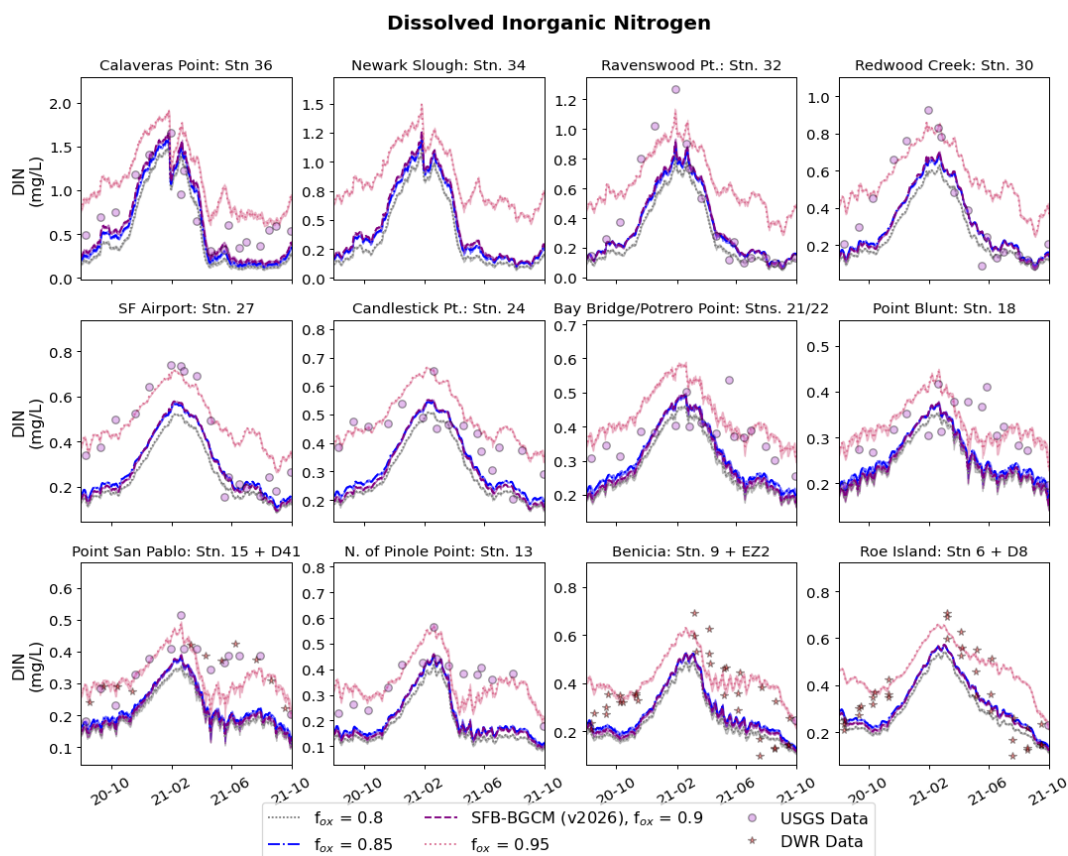


Figure A27: Comparison of dissolved inorganic nitrate from the reference modeled DIN from the aggregated grid (purple) to experiments focusing on the fraction of aerobic remineralization in the sediments. The reference value in v2026 is 0.9, experiments show values 0.8 (gray), 0.85 (blue), and 0.95 (pink). Model output is shown with the USGS Peterson Cruise Data (circles) and the DWR Environmental Monitoring Program (stars). The shaded areas show the daily minimum and maximum values.

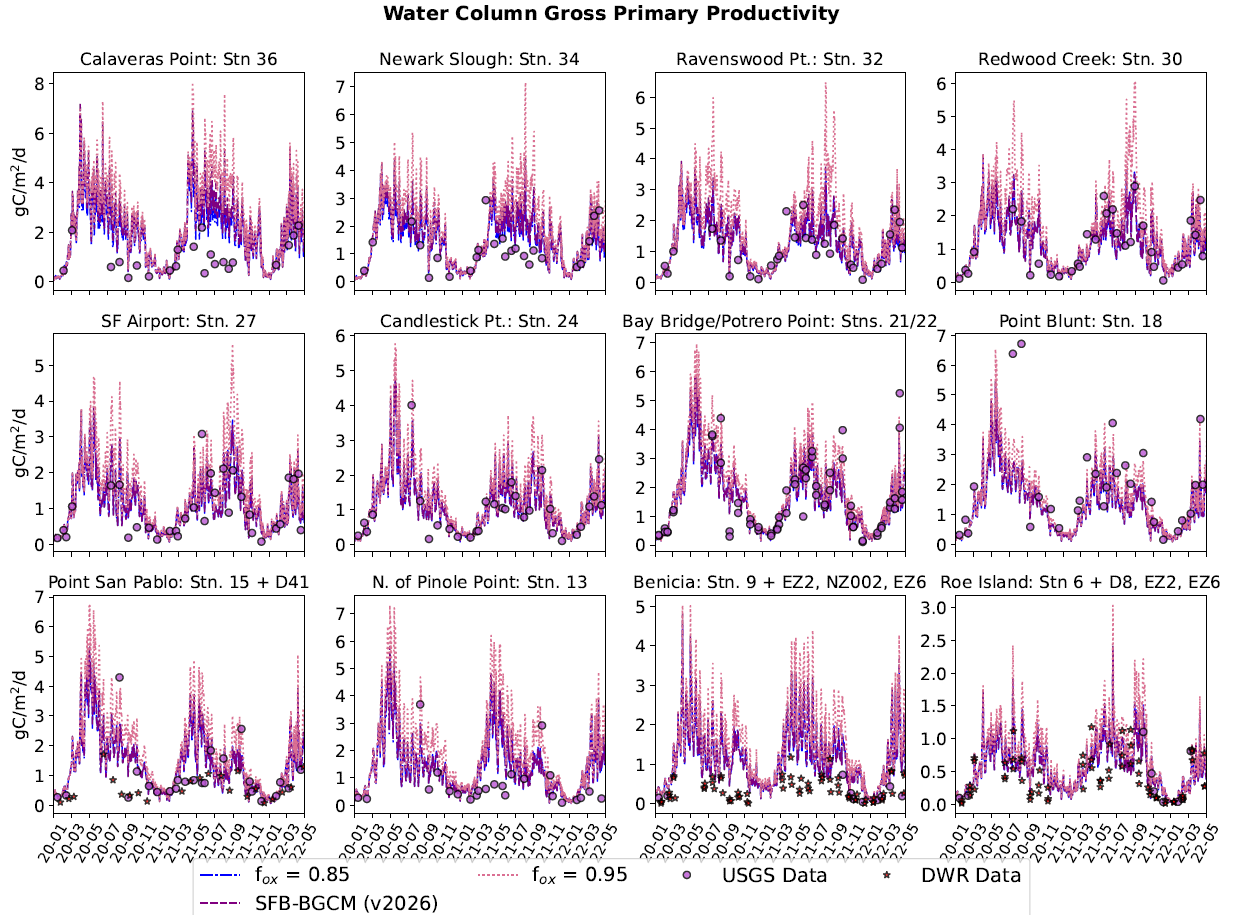


Figure A28: Comparison of water column gross primary productivity from the reference modeled GPP from the aggregated grid (purple) to experiments focusing on the fraction of aerobic remineralization in the sediments. The reference value in v2026 is 0.9, experiments show values 0.85 (blue), and 0.95 (pink). Model output is shown with with the USGS Peterson Cruise Data (circles) and the DWR Environmental Monitoring Program (stars). The shaded areas show the daily minimum and maximum values.

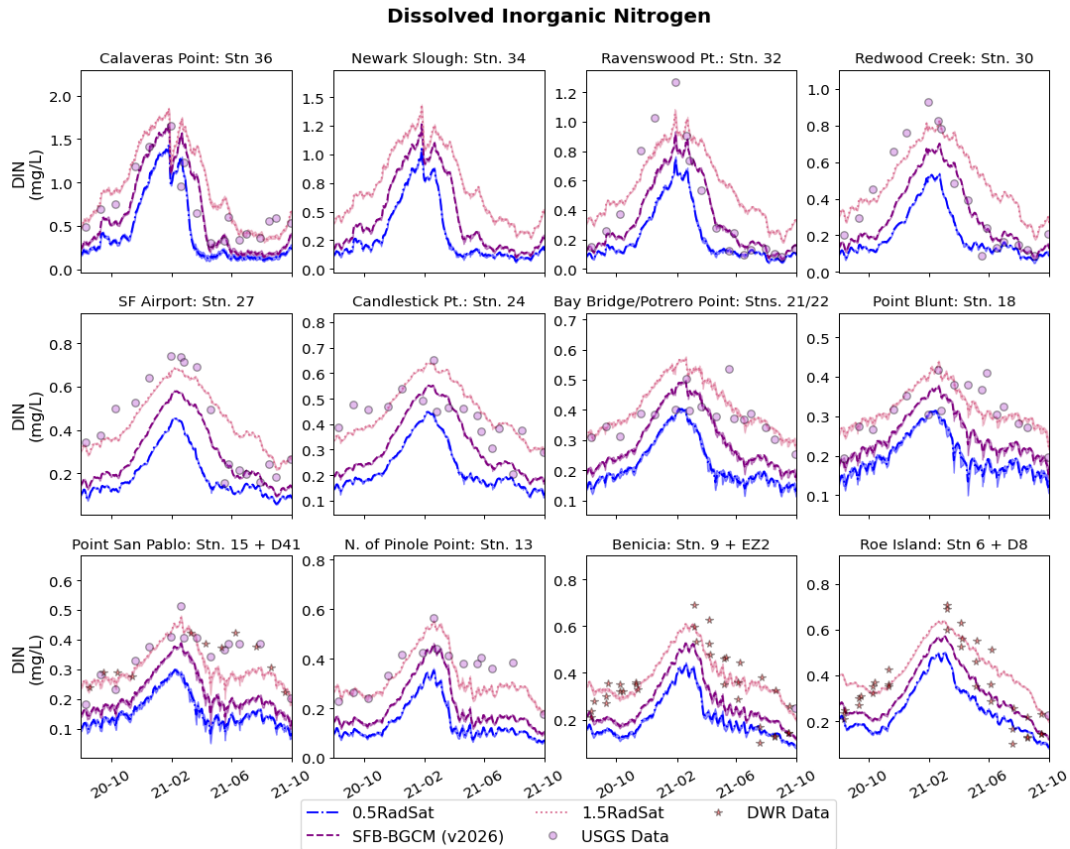


Figure A29: Comparison of dissolved inorganic nitrate from the reference modeled DIN from the aggregated grid (purple) to experiments focusing on the half saturation of radiation. Experiments varying the half saturation of radiation is shown with a reduction of 50% (blue), and an increase of 50% (pink). Model output is shown with the USGS Peterson Cruise Data (circles) and the DWR Environmental Monitoring Program (stars). The shaded areas show the daily minimum and maximum values.

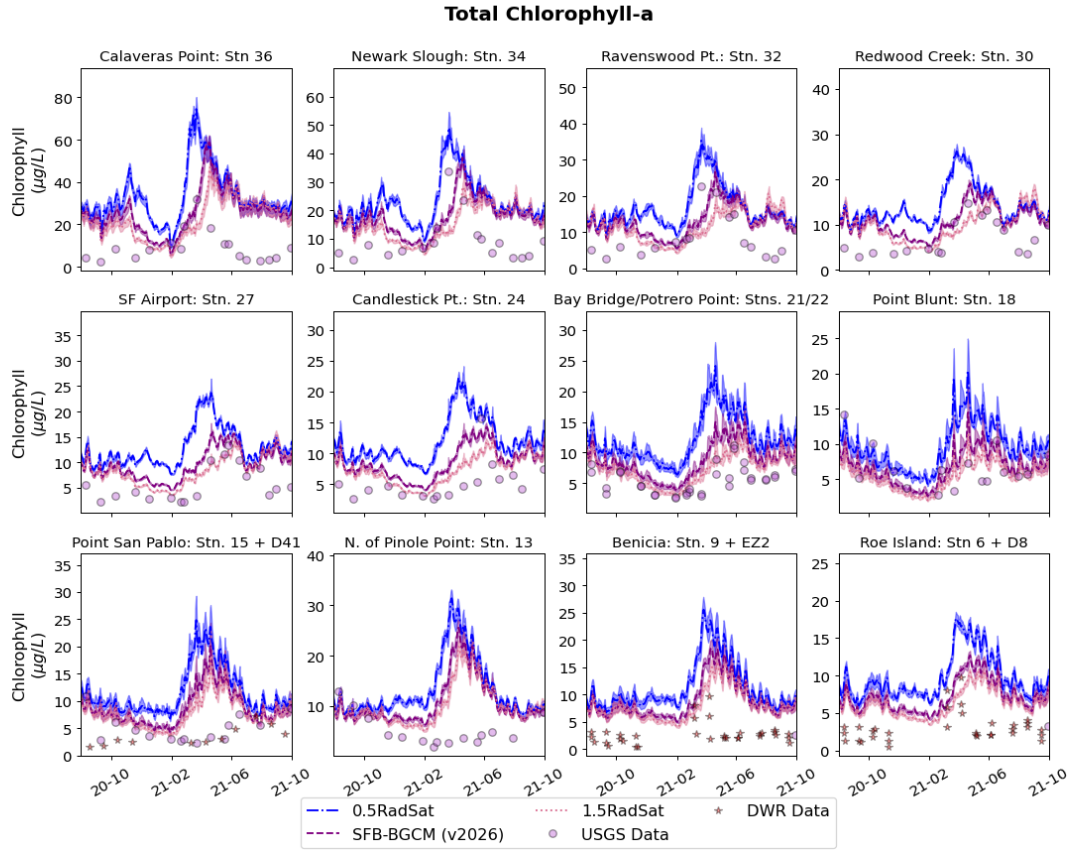


Figure A30: Comparison of total chlorophyll from the aggregated grid (purple) to experiments focusing on the half saturation of radiation. Experiments varying the half saturation of radiation is shown with a reduction of 50% (blue), and an increase of 50% (pink). Model output is shown with the USGS Peterson Cruise Data (circles) and the DWR Environmental Monitoring Program (stars). The shaded areas show the daily minimum and maximum values.

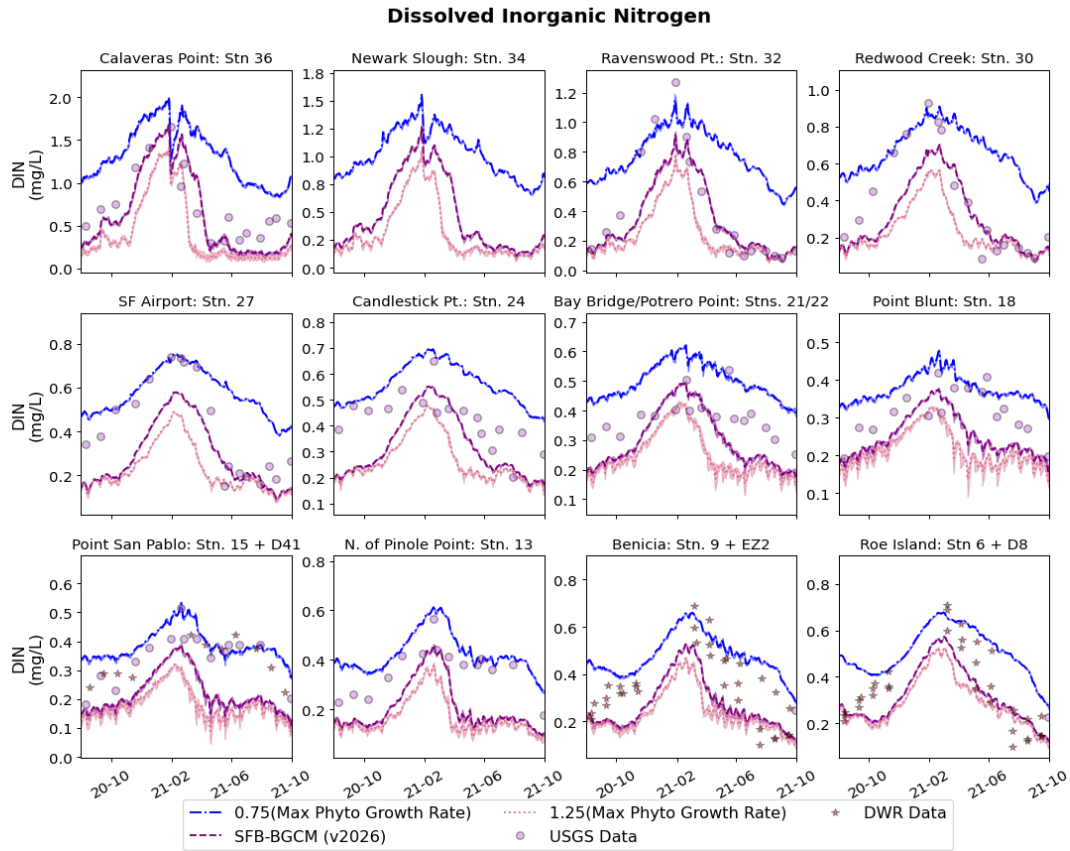


Figure A31: Comparison of dissolved inorganic nitrogen from the aggregated grid (purple) to experiments focusing on the maximum phytoplankton growth rate. Experiments varying the the maximum phytoplankton growth rate is shown with a reduction of 25% (blue), and an increase of 25% (pink). Model output is shown with with the USGS Peterson Cruise Data (circles) and the DWR Environmental Monitoring Program (stars). The shaded areas show the daily minimum and maximum values.

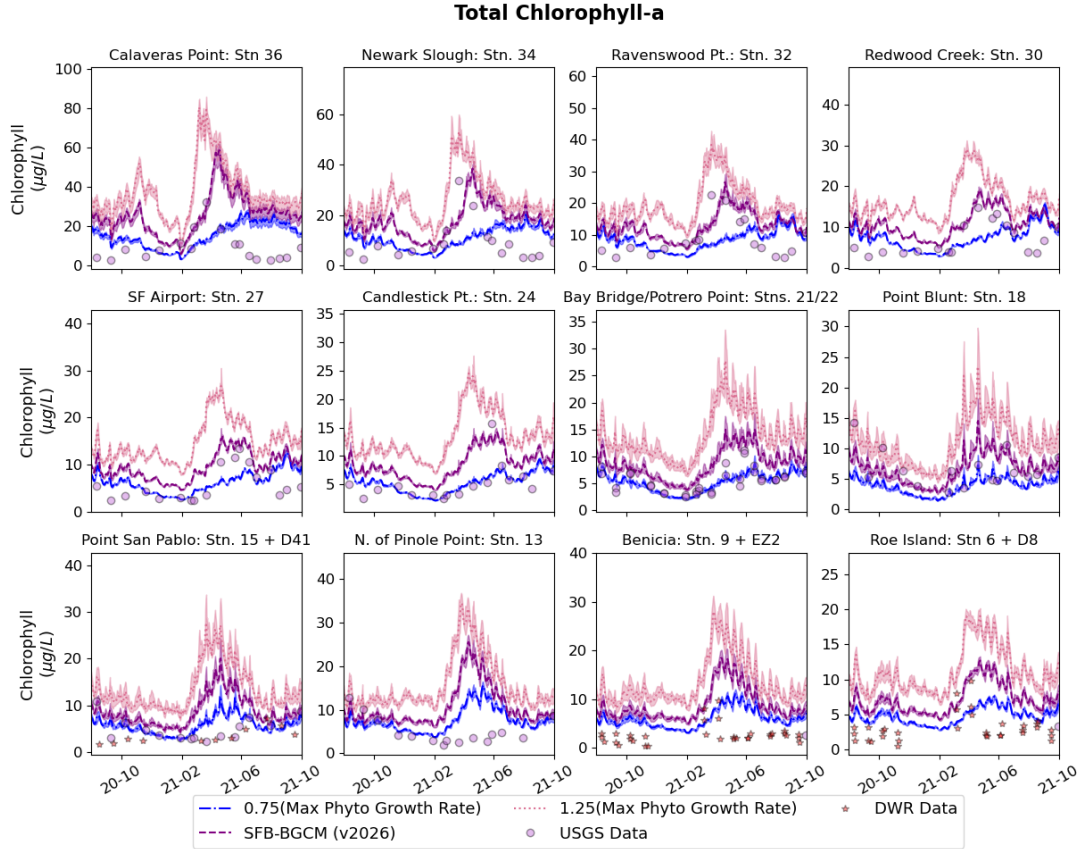


Figure A32: Comparison of total chlorophyll from the aggregated grid (purple) to experiments focusing on the maximum phytoplankton growth rate. Experiments varying the the maximum phytoplankton growth rate is shown with a reduction of 25% (blue), and an increase of 25% (pink). Model output is shown with with the USGS Peterson Cruise Data (circles) and the DWR Environmental Monitoring Program (stars). The shaded areas show the daily minimum and maximum values.

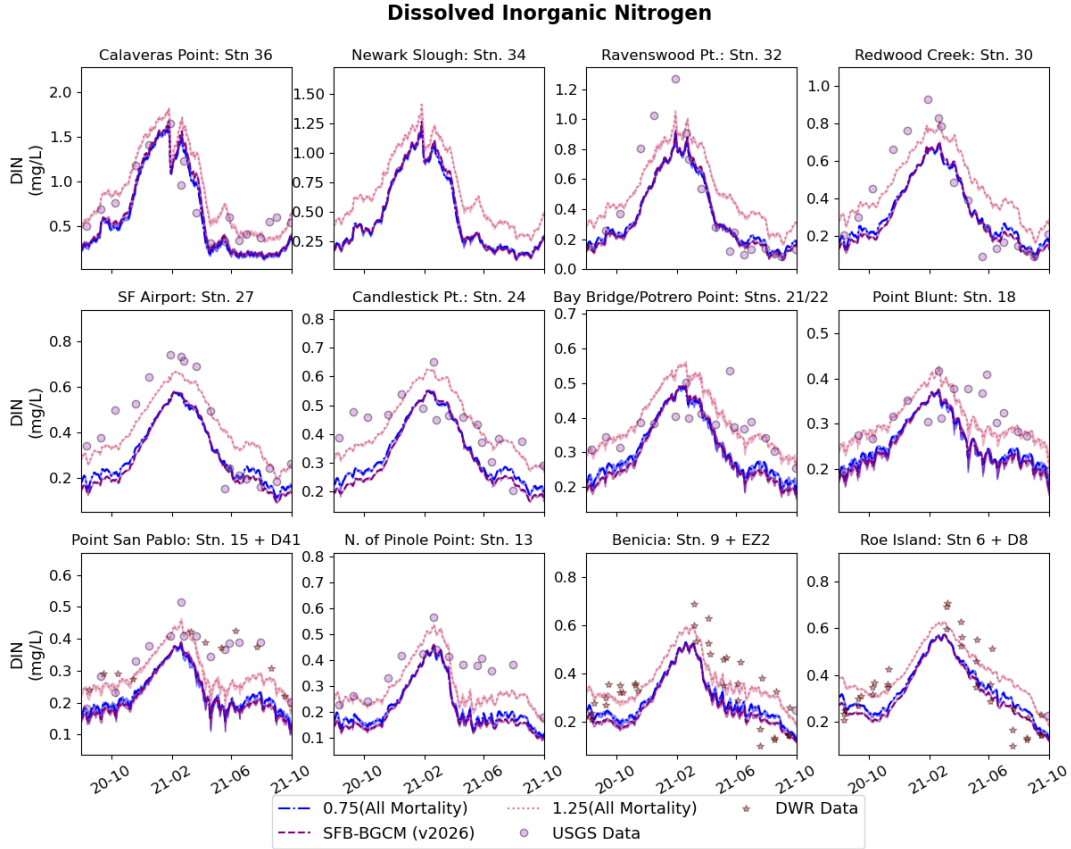


Figure A33: Comparison of dissolved inorganic nitrogen from the aggregated grid (purple) to experiments focusing on perturbations to both the linear and quadratic mortality rate. Experiments varying the mortality rate is shown with a reduction of 25% (blue), and an increase of 25% (pink). Model output is shown with the USGS Peterson Cruise Data (circles) and the DWR Environmental Monitoring Program (stars). The shaded areas show the daily minimum and maximum values.

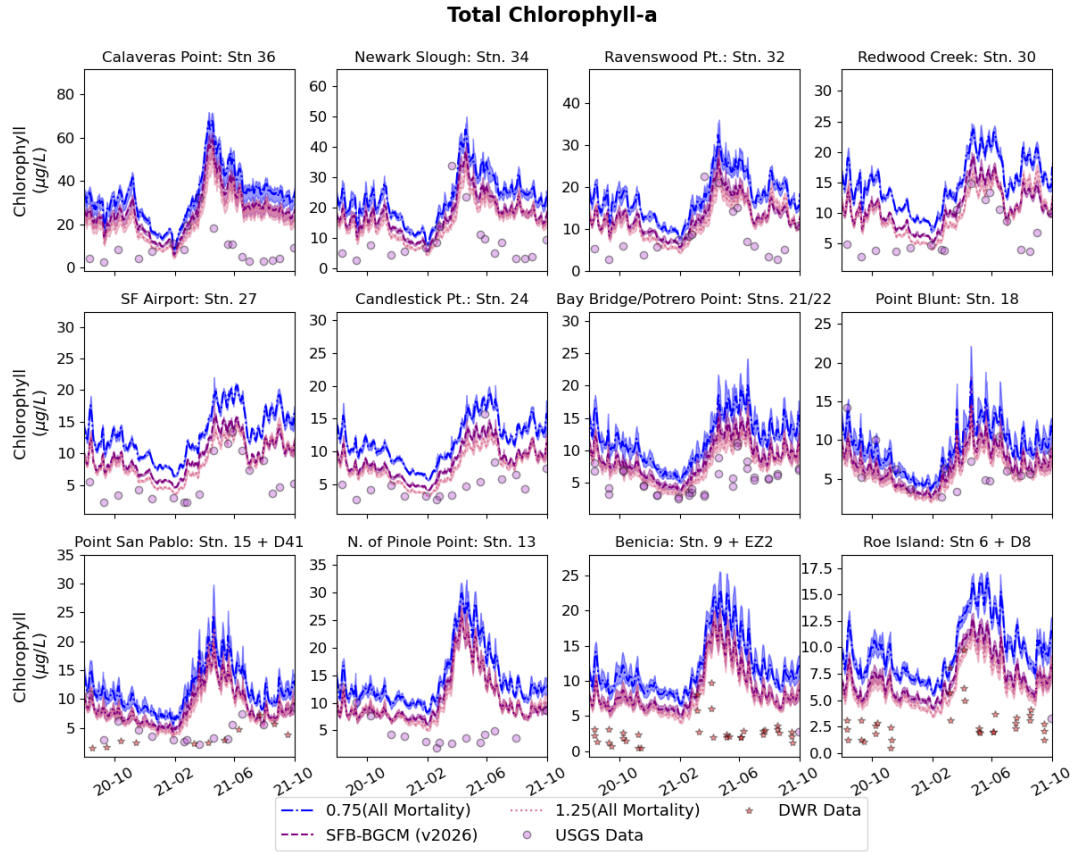


Figure A34: Comparison of total chlorophyll from the aggregated grid (purple) to experiments focusing on perturbations to both the linear and quadratic mortality rate. Experiments varying the mortality rate is shown with a reduction of 25% (blue), and an increase of 25% (pink). Model output is shown with the USGS Peterson Cruise Data (circles) and the DWR Environmental Monitoring Program (stars). The shaded areas show the daily minimum and maximum values.

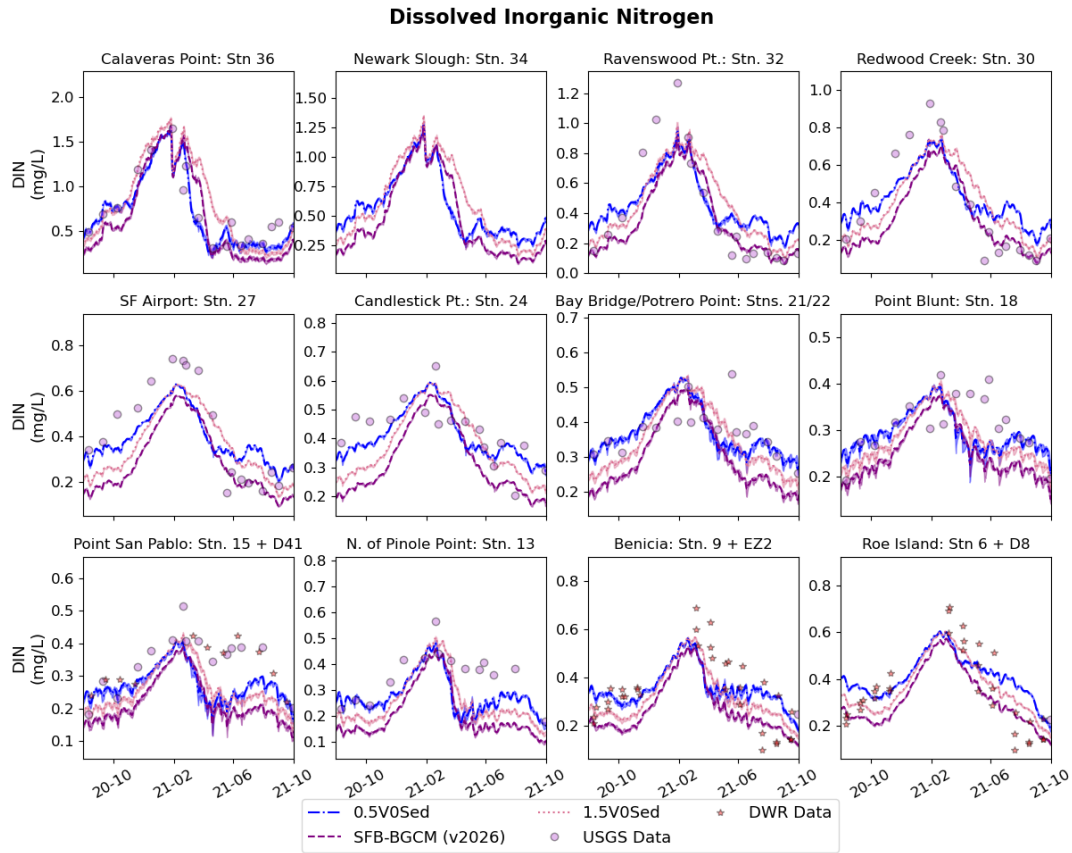


Figure A35: Comparison of dissolved inorganic nitrogen from the aggregated grid (purple) to experiments focusing on perturbations to the settling rate of sinking phytoplankton and particulate organic matter. Experiments varying the settling rate is shown with a reduction of 50% (blue), and an increase of 50% (pink). Model output is shown with the USGS Peterson Cruise Data (circles) and the DWR Environmental Monitoring Program (stars). The shaded areas show the daily minimum and maximum values.

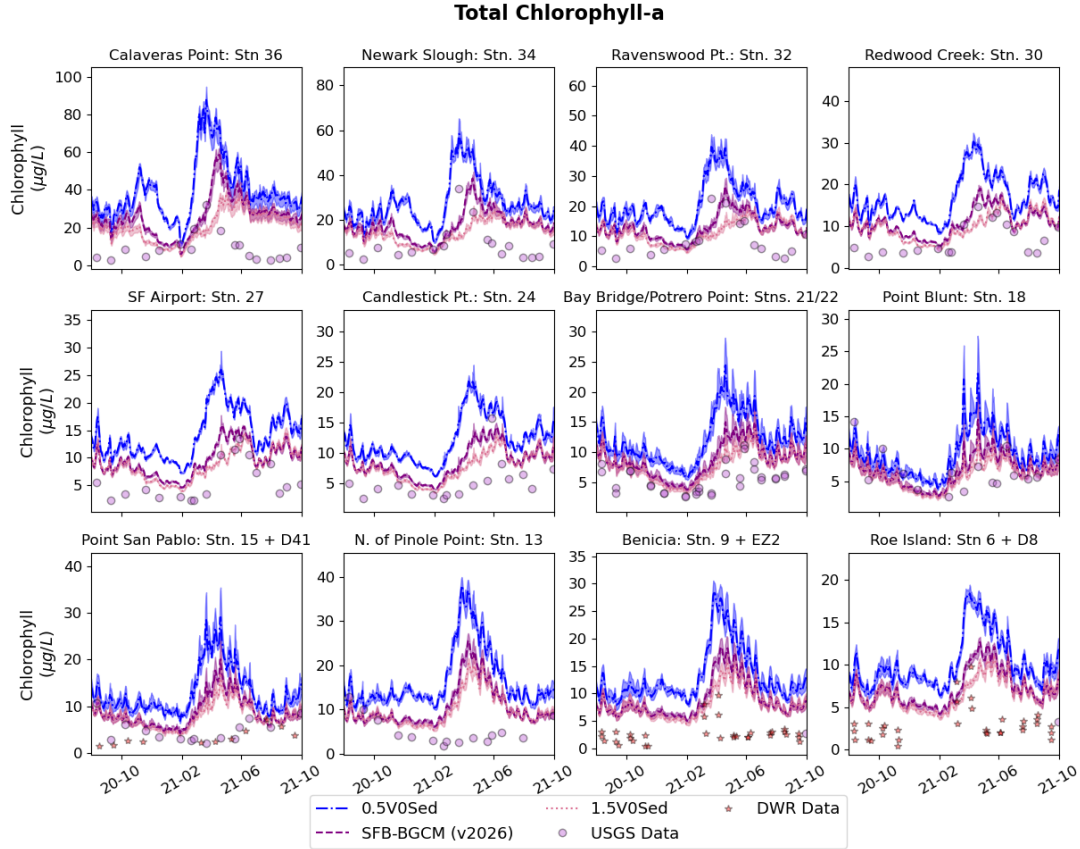


Figure A36: Comparison of total chlorophyll from the aggregated grid (purple) to experiments focusing on perturbations to both the settling rate of sinking phytoplankton and particulate organic matter. Experiments varying the settling rate is shown with a reduction of 50% (blue), and an increase of 50% (pink). Model output is shown with the USGS Peterson Cruise Data (circles) and the DWR Environmental Monitoring Program (stars). The shaded areas show the daily minimum and maximum values.

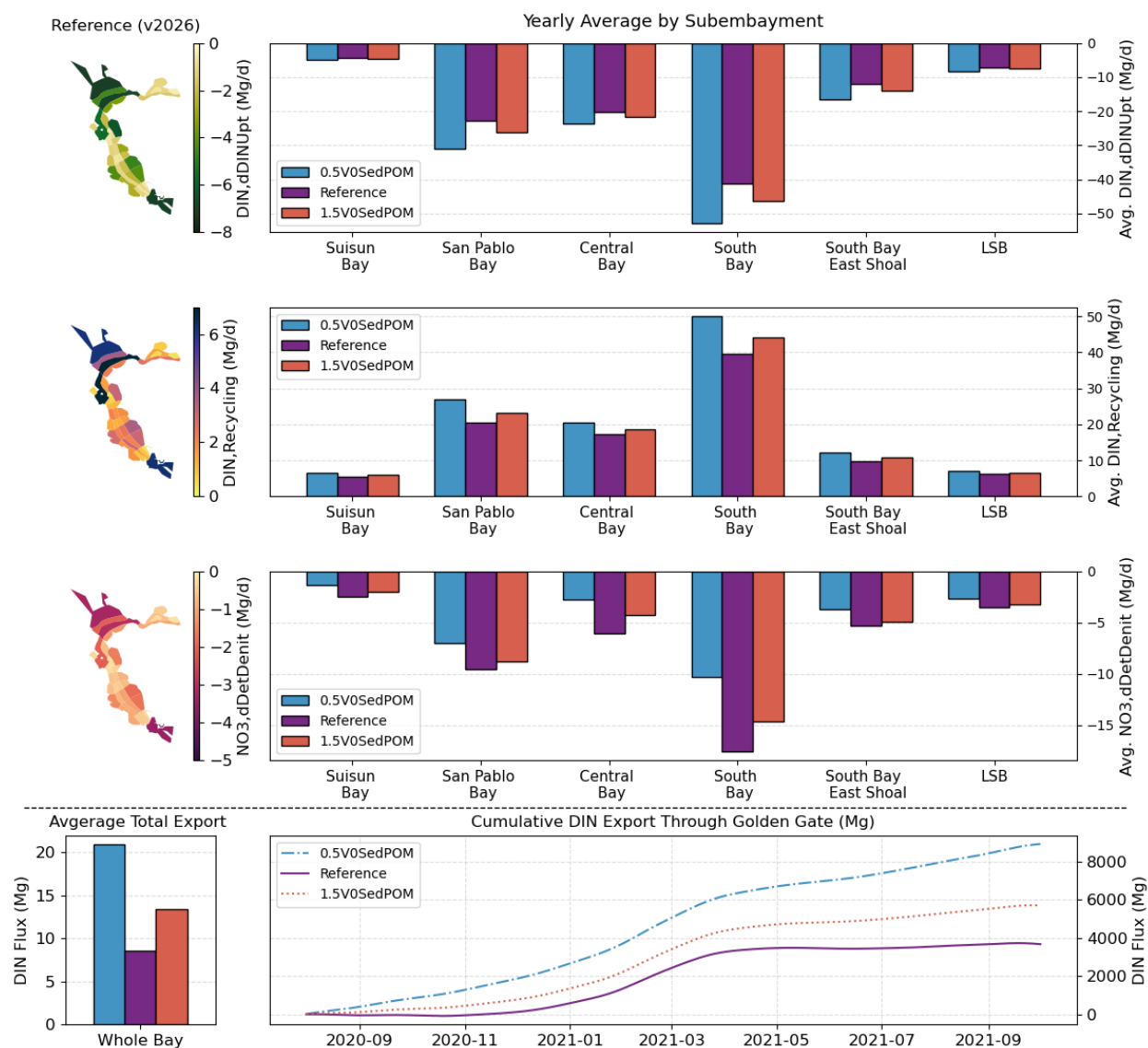


Figure A37: Figure showing control volume analysis for sensitivity to settling particulate organic matter rates for the whole bay. Parameters shown are: average dissolved inorganic nitrogen uptake rate in each volume (top row), percent difference in modeled recycling in each volume (second row), and the percent difference in modeled denitrification (third row). The left column shows a map of the average model output for the entire water year. The right column shows a bar chart comparing the reference solution to two representative model sensitivity experiments. The bottom row shows the average total export (left) and cumulative DIN export (right) through the Golden Gate Strait.

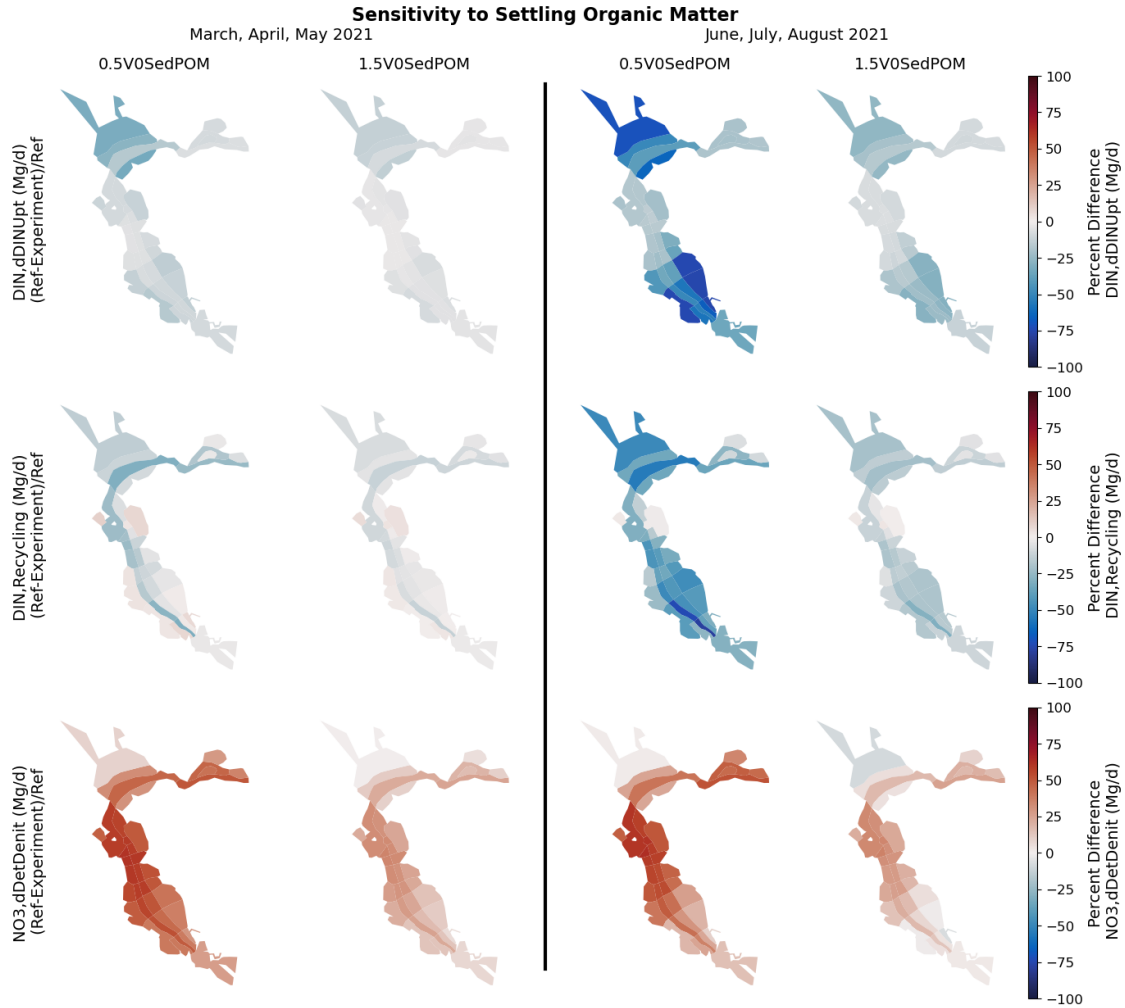


Figure A38: Figure showing control volumes for the whole Bay for an experiment focusing on the sensitivity to settling particulate organic matter rates. The parameters shown in this figure are the percent difference in average dissolved inorganic nitrogen uptake in each subembayment (top row), percent difference in DIN recycling in each volume (center row), and the percent difference in denitrification (bottom row). The left major column shows the model output for the $f_{ox}=0.85$ (left) and $f_{ox}=0.95$ (right) for March, April, May 2021. The right major column shows the model output for June, July, August 2021. Positive values (red) indicate regions of the domain where the model reference state has larger magnitude than the results in the experiment. Negative values (blue) indicate regions of the domain where the process in the experiment has larger magnitude than the model.

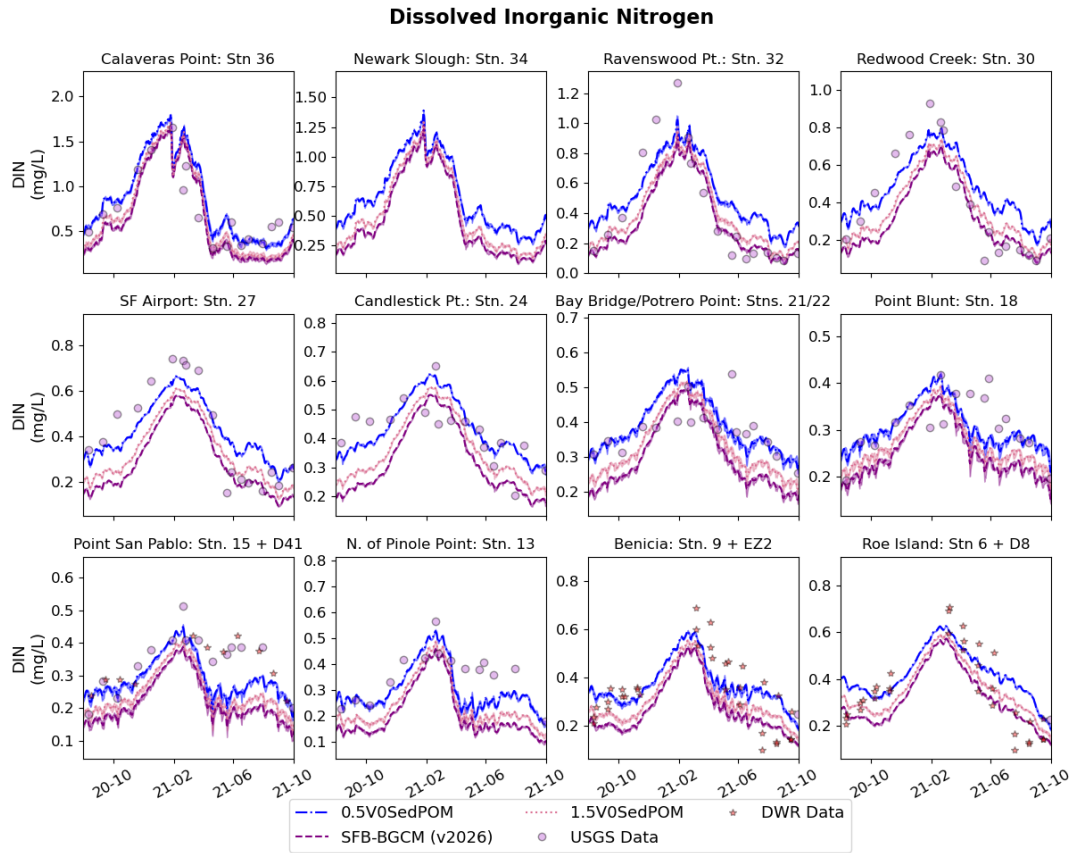


Figure A39: Comparison of dissolved inorganic nitrogen from the aggregated grid (purple) to experiments focusing on perturbations to the settling rate of sinking particulate organic matter. Experiments varying the settling rate is shown with a reduction of 50% (blue), and an increase of 50% (pink). Model output is shown with the USGS Peterson Cruise Data (circles) and the DWR Environmental Monitoring Program (stars). The shaded areas show the daily minimum and maximum values.

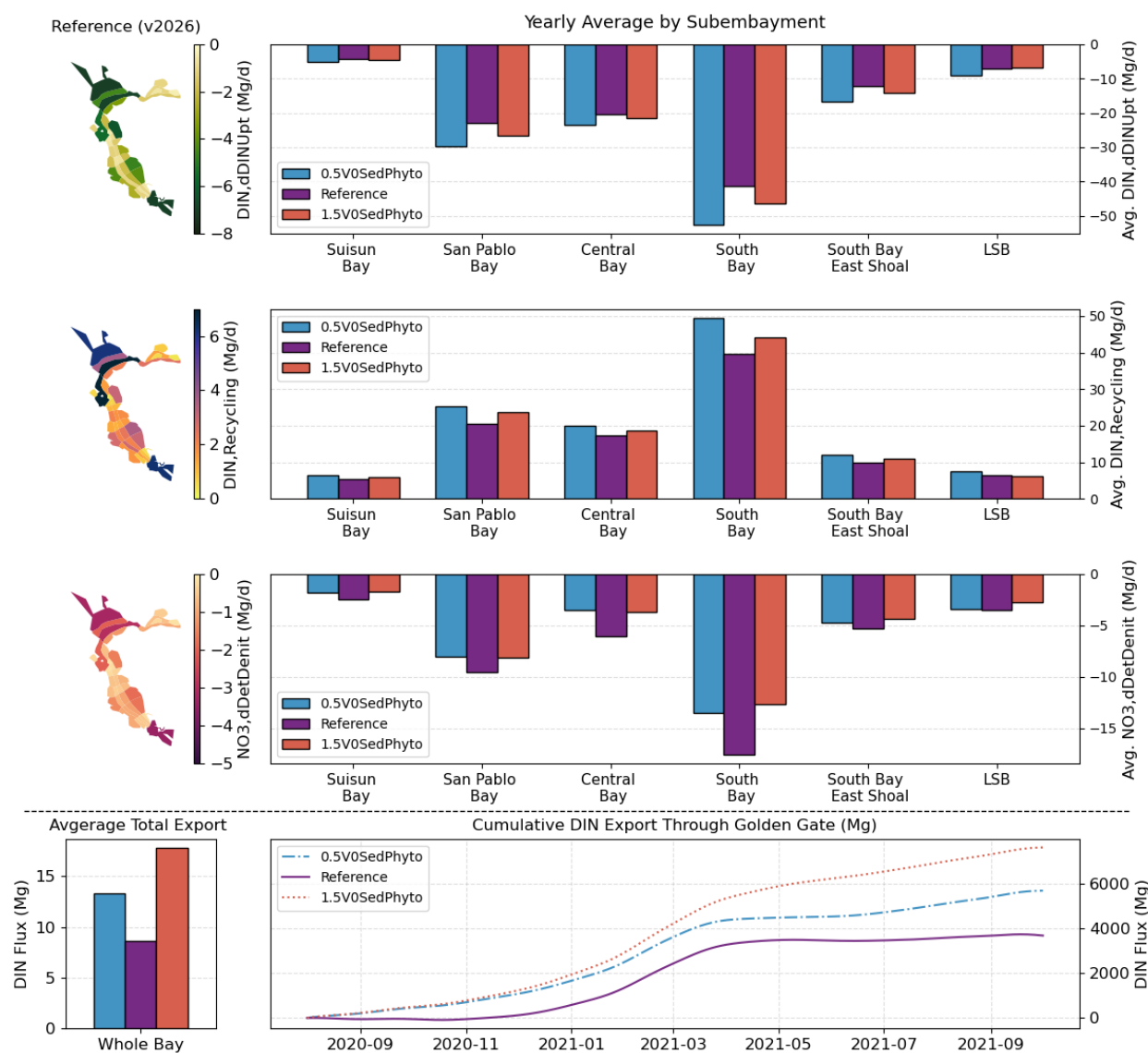


Figure A40: Figure showing control volume analysis for sensitivity to settling phytoplankton rates for the whole bay. Parameters shown are: average dissolved inorganic nitrogen uptake rate in each volume (top row), percent difference in modeled recycling in each volume (second row), and the percent difference in modeled denitrification (third row). The left column shows a map of the average model output for the entire water year. The right column shows a bar chart comparing the reference solution to two representative model sensitivity experiments. The bottom row shows the average total export (left) and cumulative DIN export (right) through the Golden Gate Strait.

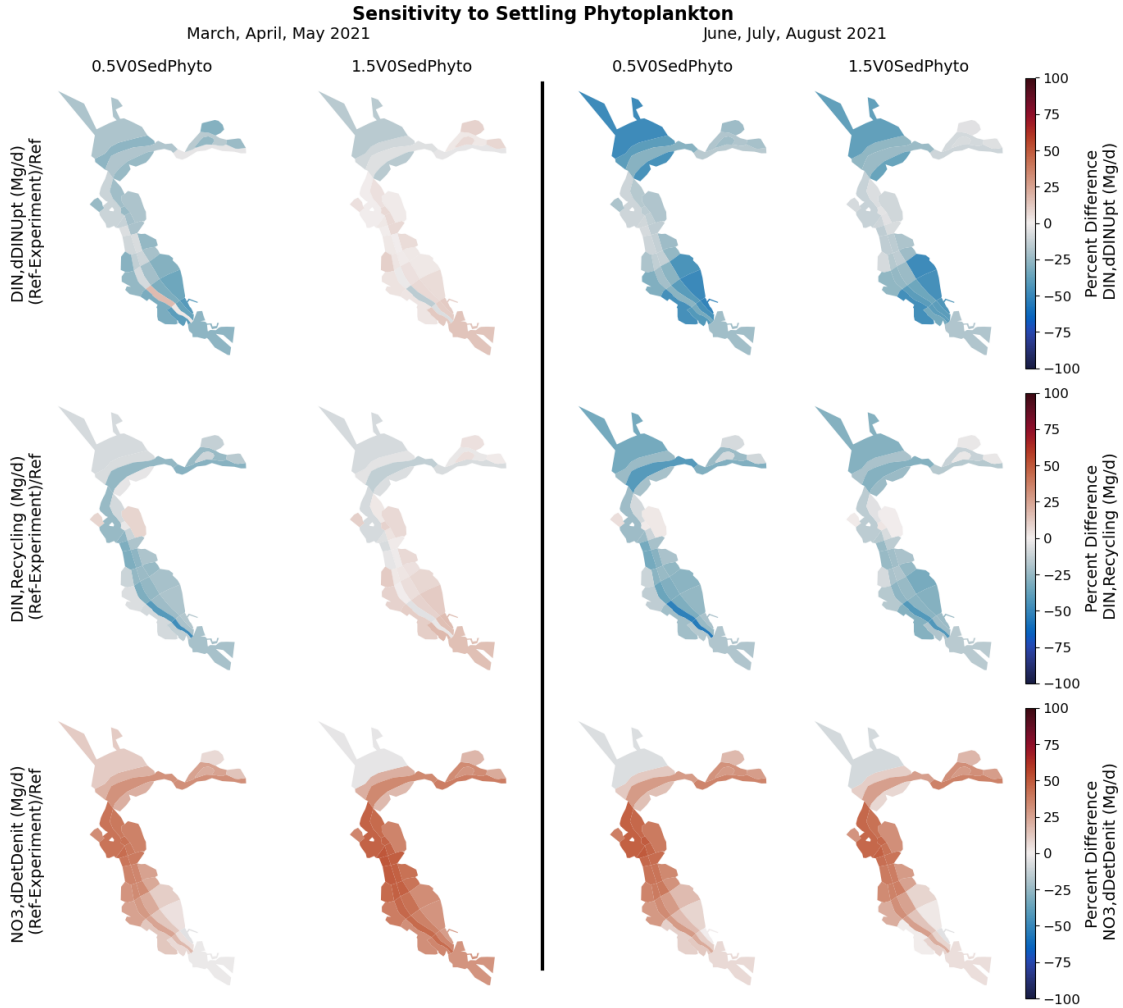


Figure A41: Figure showing control volumes for the whole Bay for an experiment focusing on the sensitivity to settling phytoplankton rates. The parameters shown in this figure are the percent difference in average dissolved inorganic nitrogen uptake in each subembayment (top row), percent difference in DIN recycling in each volume (center row), and the percent difference in denitrification (bottom row). The left major column shows the model output for the $f_{ox}=0.85$ (left) and $f_{ox}=0.95$ (right) for March, April, May 2021. The right major column shows the model output for June, July, August 2021. Positive values (red) indicate regions of the domain where the model reference state has larger magnitude than the results in the experiment. Negative values (blue) indicate regions of the domain where the process in the experiment has larger magnitude than the model.

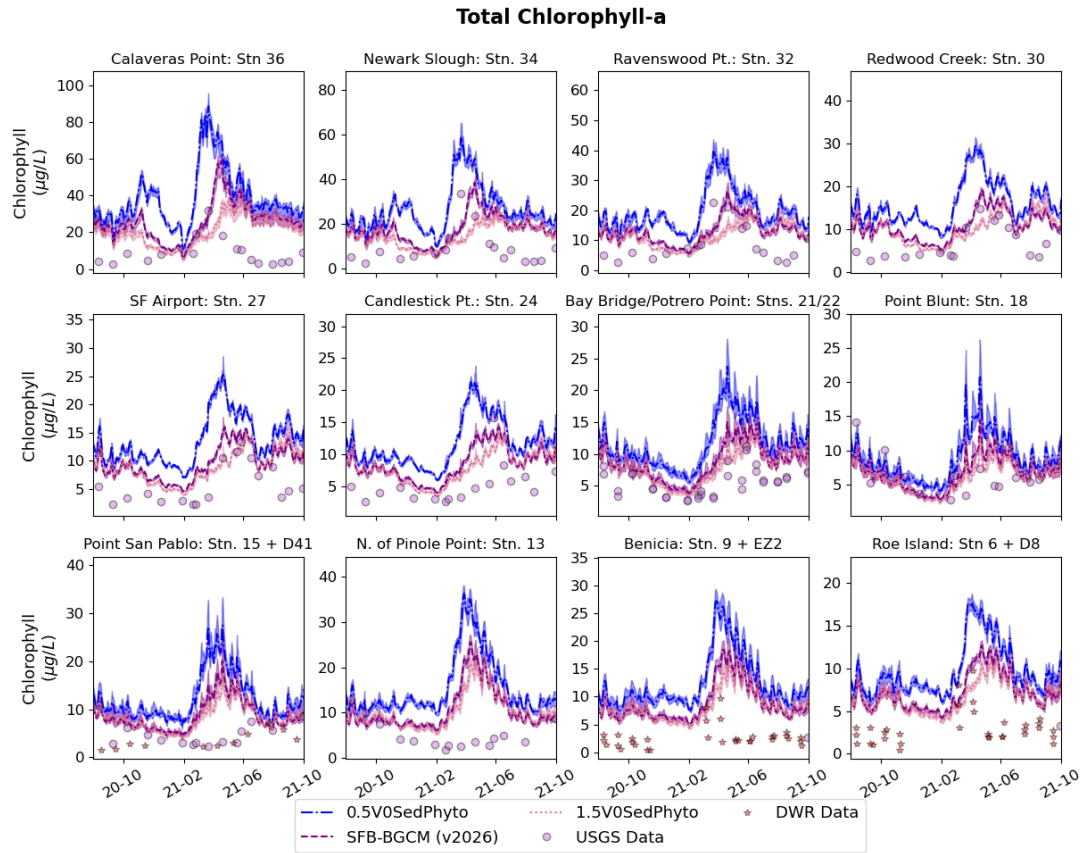


Figure A42: Comparison of total chlorophyll from the aggregated grid (purple) to experiments focusing on perturbations to the settling rate of sinking phytoplankton. Experiments varying the settling rate of sinking phytoplankton is shown with a reduction of 50% (blue), and an increase of 50% (pink). Model output is shown with the USGS Peterson Cruise Data (circles) and the DWR Environmental Monitoring Program (stars). The shaded areas show the daily minimum and maximum values.

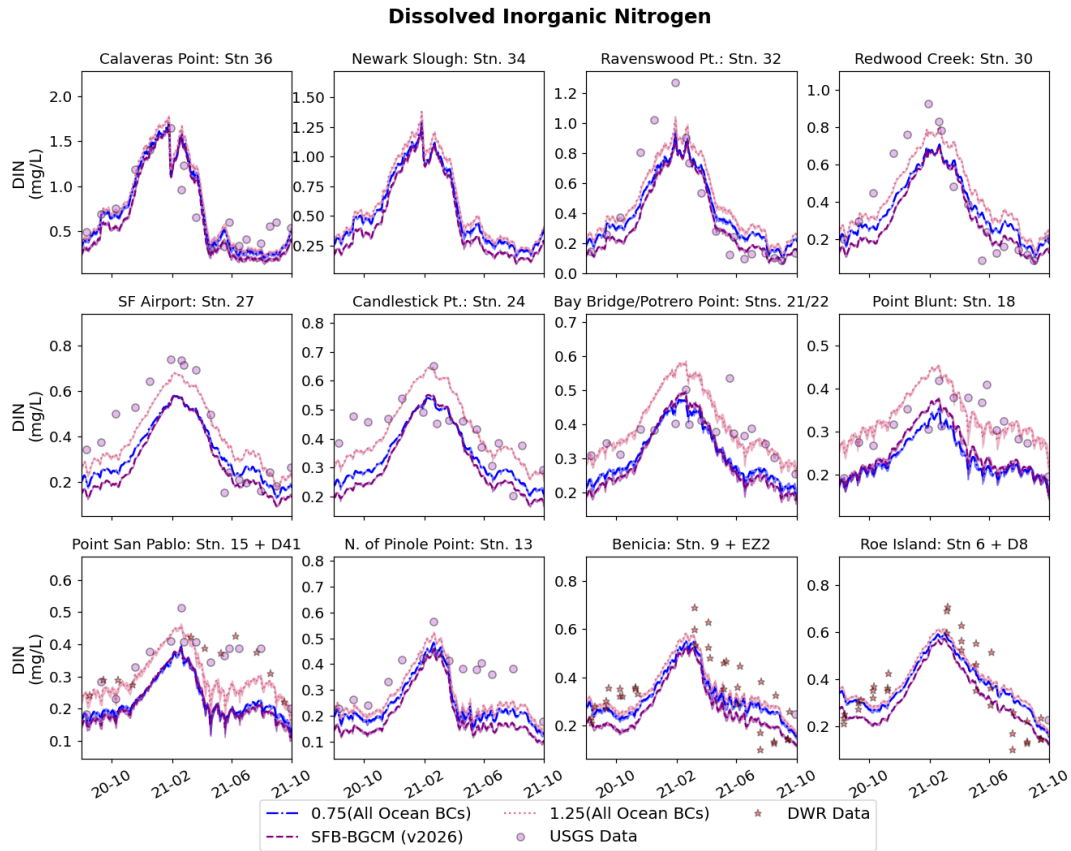


Figure A43: Comparison of dissolved inorganic nitrogen concentration from the aggregated grid (purple) to experiments focusing on perturbations to the ocean boundary conditions. Experiments varying the ocean boundary is shown with a reduction of 25% (blue), and an increase of 25% (pink). Model output is shown with the USGS Peterson Cruise Data (circles) and the DWR Environmental Monitoring Program (stars). The shaded areas in show the daily minimum and maximum values.

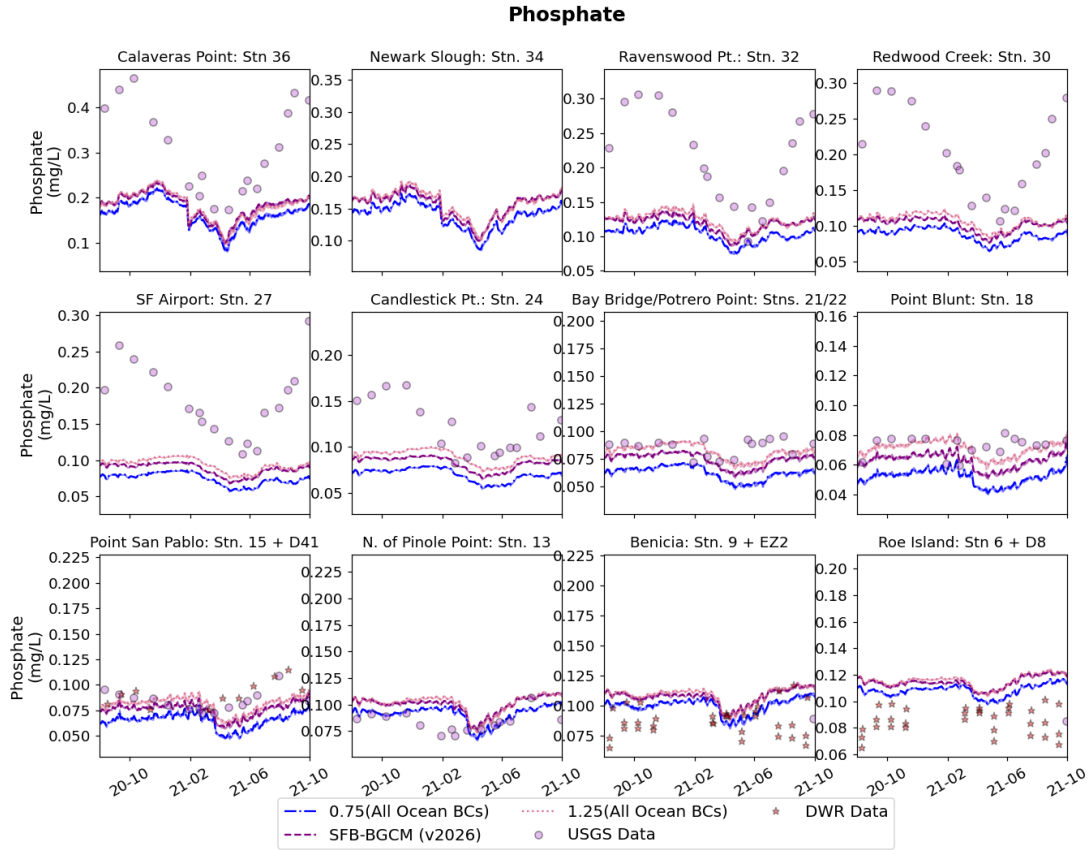


Figure A44: Comparison of phosphorus concentration from the aggregated grid (purple) to experiments focusing on perturbations to the ocean boundary conditions. Experiments varying the ocean boundary is shown with a reduction of 25% (blue), and an increase of 25% (pink). Model output is shown with the USGS Peterson Cruise Data (circles) and the DWR Environmental Monitoring Program (stars). The shaded areas in show the daily minimum and maximum values.

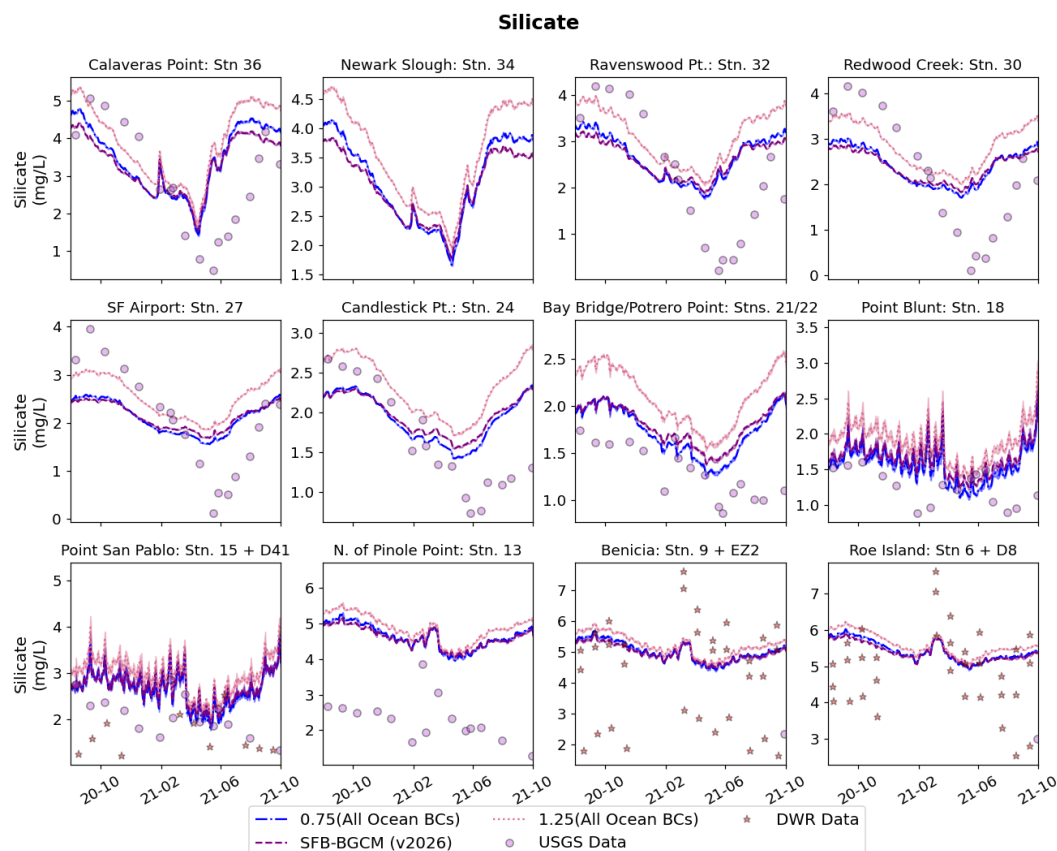


Figure A45: Comparison of silicate concentration from the aggregated grid (purple) to experiments focusing on perturbations to the ocean boundary conditions. Experiments varying the ocean boundary is shown with a reduction of 25% (blue), and an increase of 25% (pink). Model output is shown with the USGS Peterson Cruise Data (circles) and the DWR Environmental Monitoring Program (stars). The shaded areas in show the daily minimum and maximum values.

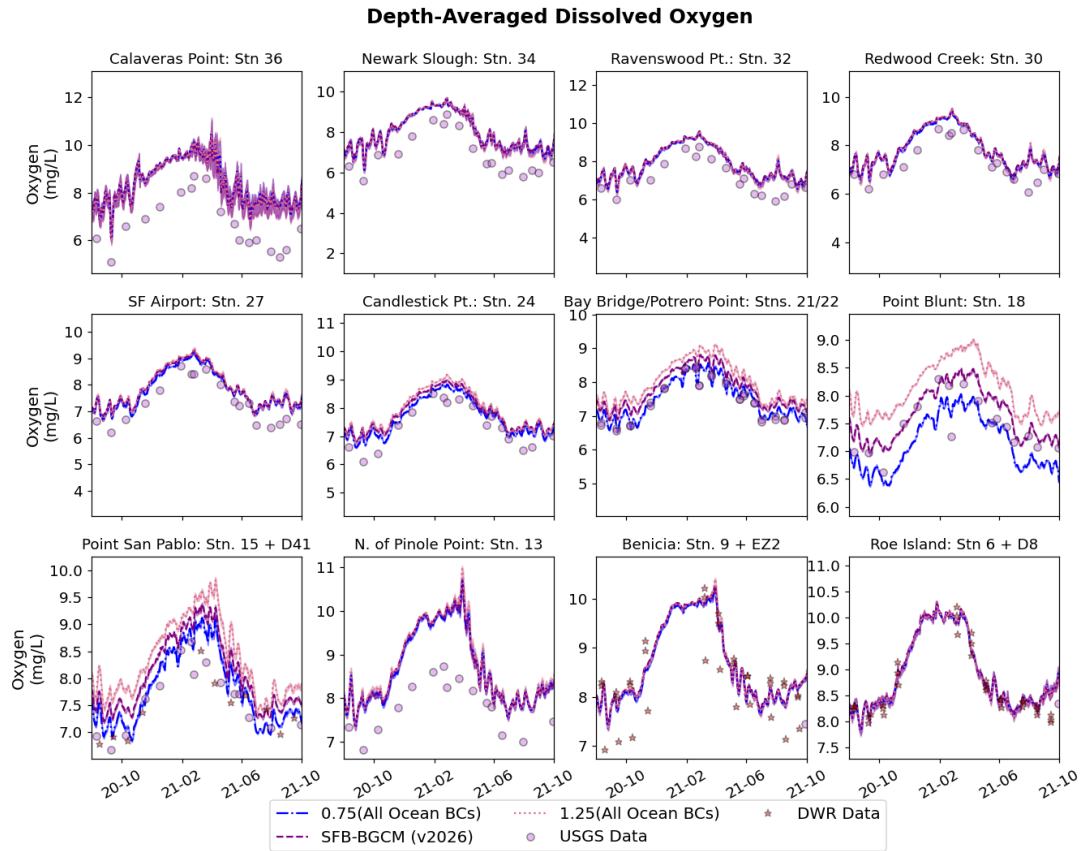


Figure A46: Comparison of depth-averaged dissolved oxygen concentration from the aggregated grid (purple) to experiments focusing on perturbations to the ocean boundary conditions. Experiments varying the ocean boundary is shown with a reduction of 25% (blue), and an increase of 25% (pink). Model output is shown with the USGS Peterson Cruise Data (circles) and the DWR Environmental Monitoring Program (stars). The shaded areas in show the daily minimum and maximum values.