

# Appendix B:

## Model Evaluation of Temperature and Nitrogen

### 1 Data Sources and Methods

In situ data for model evaluation were downloaded from the Chesapeake Bay Program Water Quality Database ([https://www.chesapeakebay.net/what/downloads/cbp\\_water\\_quality\\_database\\_1984\\_present](https://www.chesapeakebay.net/what/downloads/cbp_water_quality_database_1984_present)). Observation times were matched up with hourly model data output at the stations shown in Fig. 2 of the manuscript. Measurements at depths corresponding to model depths from the free surface were used in the comparison. In this appendix, we present model evaluation of water temperature for the years 2001-2005.

### 2 Validation statistics

For  $N$  observations, let  $y_i$  in  $y = [y_1, \dots, y_N]$  be the in situ measurements, and  $\hat{y}_i$  in  $\hat{y} = [\hat{y}_1, \dots, \hat{y}_N]$  be the corresponding model output for either simulation in this study. The average of all observations is  $\bar{y}$ , and the average of all corresponding modeled data is  $\bar{\hat{y}}$ . The validation statistics are calculated as follows:

Linear correlation coefficient

$$R^2 = \left( \frac{\frac{1}{N} \sum_{i=1}^N (\hat{y}_i - \bar{\hat{y}})(y_i - \bar{y})}{\sigma_{\hat{y}} \sigma_y} \right)^2$$

Mean Bias

$$\text{mean bias} = \frac{1}{N} \sum_{i=1}^N (\hat{y}_i - y_i)$$

Root Mean Square Error

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{y}_i - y_i)^2}$$

Absolute Mean Relative Error

$$absMRE = 100 * \frac{|\text{mean bias}|}{\bar{y}}$$

Table B.1: Observed vs. modeled temperatures from 2001 to 2005 for mainstem stations shown in Fig. 2 for coincident times and depths

| Region           | N      | Observed vs. standard |           |      |        | Observed vs. bgc-light |           |      |        |
|------------------|--------|-----------------------|-----------|------|--------|------------------------|-----------|------|--------|
|                  |        | $R^2$                 | mean bias | RMSE | absMRE | $R^2$                  | mean bias | RMSE | absMRE |
| Oligohaline      | 1,373  | 0.98                  | -0.38     | 1.44 | 2.3    | 0.98                   | -0.41     | 1.46 | 2.6    |
| Upper Mesohaline | 6,147  | 0.98                  | 0.22      | 1.17 | 1.4    | 0.98                   | 0.14      | 1.18 | 0.9    |
| Lower Mesohaline | 6,743  | 0.99                  | 0.06      | 1.01 | 0.4    | 0.99                   | 0.03      | 1.03 | 0.2    |
| Polyhaline       | 3,508  | 0.97                  | -0.21     | 1.45 | 1.2    | 0.97                   | -0.21     | 1.46 | 1.2    |
| All              | 17,771 | 0.98                  | 0.03      | 1.19 | 0.2    | 0.98                   | -0.01     | 1.21 | 0.1    |

Table B.2: In situ nitrite+nitrate [ $mmolN\ m^{-3}$ ] vs. modeled nitrate [ $mmolNO_3^-\ m^{-3}$ ] from 2001 to 2005 for mainstem stations shown in Fig. 2 for coincident times and depths

| Region           | N     | Observed vs. standard |           |      |        | Observed vs. bgc-light |           |      |        |
|------------------|-------|-----------------------|-----------|------|--------|------------------------|-----------|------|--------|
|                  |       | $R^2$                 | mean bias | RMSE | absMRE | $R^2$                  | mean bias | RMSE | absMRE |
| Oligohaline      | 498   | 0.18                  | -8.2      | 30.8 | 13     | 0.18                   | -8.7      | 30.9 | 13     |
| Upper Mesohaline | 1,219 | 0.62                  | 0.03      | 14.4 | 0.14   | 0.62                   | -0.21     | 14.4 | 0.93   |
| Lower Mesohaline | 1,202 | 0.48                  | 1.2       | 9.75 | 16     | 0.47                   | 1.2       | 9.75 | 15     |
| Polyhaline       | 1,203 | 0.26                  | 1.5       | 7.17 | 56     | 0.25                   | 1.4       | 7.13 | 54     |
| All              | 4,122 | 0.68                  | -0.19     | 14.8 | 1.1    | 0.68                   | -0.36     | 14.8 | 2.0    |

### 3 Time series

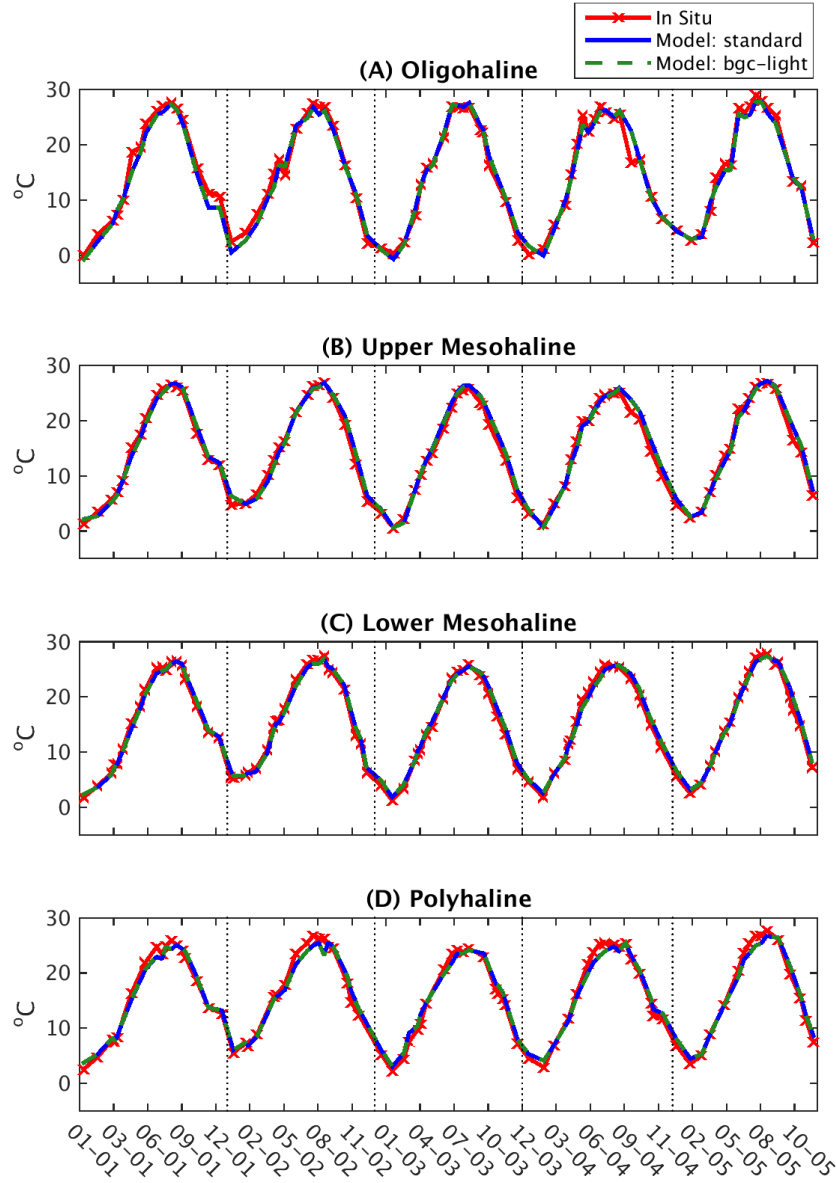


Figure B.1: For each region, water temperature observations and corresponding model output for every 7-day period was averaged over all matching depths. Panels A-D show the temperature for the mainstem stations in regions shown in the manuscript Fig. 2. The red line indicates temperature observations, with "x"'s marking times when observations are available. Blue and dashed green lines are data from the two model runs described in the manuscript, "standard" and "bgc-light".

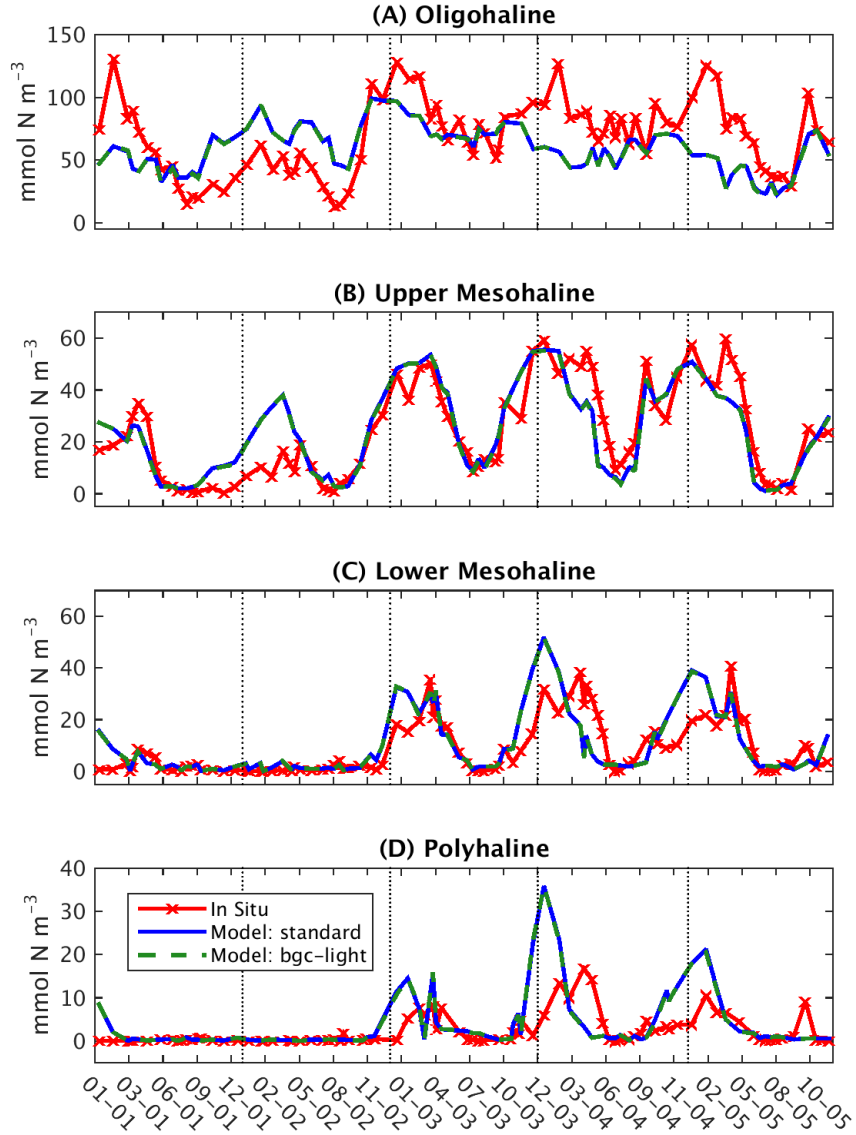


Figure B.2: For each region,  $[NO_2^- + NO_3^-]$  observations and corresponding  $[NO_3^-]$  model output for every 7-day period was averaged over all matching depths. Panels A-D show the nutrients for the mainstem stations in regions shown in the manuscript Fig. 2. The red line indicates nutrient observations, with "x"s marking times when observations are available. Blue and dashed green lines are data from the two model runs described in the manuscript, "standard" and "bgc-light".