JGR Biogeosciences

RESEARCH ARTICLE
10.1029/2023JG007536

Predicting Carbonate Chemistry on the Northwest Atlantic Shelf Using Neural Networks

Ivan D. Lima[1], Zhaohui A. Wang[1], Louise P. Cameron[1], Jonathan H. Grabowski[2], and Jennie E. Rheuban[1]

[1]Department of Marine Chemistry and Geochemistry, Woods Hole Oceanographic Institution, Woods Hole, MA, USA, [2]Marine Science Center, Northeastern University, Nahant, MA, USA

Key Points:

- Three neural network models were developed to predict carbonate chemistry for the Northwest Atlantic Shelf with high precision and accuracy
- Distinct models have increased numbers of predictors that improve model performance and allow for a wide range of different applications
- One model was applied to modern-day hydrographic and satellite data to generate seasonal maps of surface and bottom water conditions

Correspondence to:
J. E. Rheuban,
jrheuban@whoi.edu

Citation:
Lima, I. D., Wang, Z. A., Cameron, L. P., Grabowski, J. H., & Rheuban, J. E. (2023). Predicting carbonate chemistry on the Northwest Atlantic Shelf using neural networks. *Journal of Geophysical Research: Biogeosciences*, *128*, e2023JG007536. https://doi.org/10.1029/2023JG007536

Received 20 APR 2023
Accepted 12 JUN 2023

Author Contributions:
Conceptualization: Ivan D. Lima, Jennie E. Rheuban
Data curation: Louise P. Cameron, Jennie E. Rheuban
Formal analysis: Ivan D. Lima, Jennie E. Rheuban
Funding acquisition: Ivan D. Lima, Zhaohui A. Wang, Jonathan H. Grabowski, Jennie E. Rheuban
Investigation: Ivan D. Lima, Louise P. Cameron, Jennie E. Rheuban
Methodology: Ivan D. Lima
Project Administration: Zhaohui A. Wang, Jennie E. Rheuban
Resources: Zhaohui A. Wang
Software: Ivan D. Lima

Abstract The Northwest Atlantic Shelf (NAS) region has experienced accelerated warming, heatwaves, and is susceptible to ocean acidification, yet also suffers from a paucity of carbonate chemistry observations, particularly at depth. We address this critical data gap by developing three different neural network models to predict dissolved inorganic carbon (DIC) and total alkalinity (TA) in the NAS region from more readily available hydrographic and satellite data. The models predicted DIC with r^2 between 0.913–0.963 and root mean square errors (RMSE) between 15.4–23.7 (μmol kg^{-1}) and TA with r^2 between 0.986–0.983 and RMSE between 9.0–10.4 (μmol kg^{-1}) on an unseen test data set that was not used in training the models. Cross-validation analysis revealed that all models were insensitive to the choice of training data and had good generalization performance on unseen data. Uncertainty in DIC and TA were low (coefficients of variation 0.1%–1%). Compared with other predictive models of carbonate system variables in this region, a larger and more diverse data set with full seasonal coverage and a more sophisticated model architecture resulted in a robust predictive model with higher accuracy and precision across all seasons. We used one of the models to generate a reconstructed seasonal distribution of carbonate chemistry fields based on DIC and TA predictions that shows a clear seasonal progression and large spatial gradients consistent with observations. The distinct models will allow for a range of applications based on the predictor variables available and will be useful to understand and address ocean sustainability challenges.

Plain Language Summary The U.S. northeast coast is particularly susceptible to climate change and ocean acidification. However, the lack of observations on seawater carbonate chemistry makes it difficult to assess the impacts of ocean acidification on the region. We address this information gap by developing three different machine learning models to predict carbonate system parameters from more readily available field and satellite data. The models predicted carbonate system parameters with high accuracy and good precision. Compared with other models of carbonate chemistry variables for this region, a larger data set with full seasonal and vertical coverage of the water column and a more complex model architecture resulted in a robust model with low error and uncertainty across all four seasons as well as in surface and subsurface waters. The reconstructed distributions of carbonate chemistry fields on U.S. northeast coast based on one of the models were consistent with observations. We anticipate that the distinct versions of the model will allow for a wide range of different applications based on the predictor variables that are available.

1. Introduction

The Northwest Atlantic Shelf Region (NAS) is a pronounced western boundary current margin that receives significant influences from the Gulf Stream on the ocean side, while experiencing the cumulative impacts of terrestrial inputs from small-to-medium size rivers, saltmarshes, and groundwater on the landward side (Cai et al., 2020; Fennel et al., 2019; Najjar et al., 2018; Wang et al., 2013). The region has been the location of dramatic shifts in climate indicators over the last century, such as accelerated warming and heatwaves (Chen et al., 2014; Forsyth et al., 2015; Gawarkiewicz et al., 2019; Mills et al., 2013; Perez et al., 2021; Pershing et al., 2015; Schlegel et al., 2021), rapid sea-level rise (Piecuch et al., 2018; Sallenger et al., 2012), and the increasing direct and indirect influences from the Gulf Stream as a result of changes in its meanders and interaction with the shelf (Andres, 2016; Gangopadhyay et al., 2019; Gawarkiewicz et al., 2018; Monim, 2017). This region is also potentially more vulnerable to ocean acidification due to naturally low pH and calcium carbonate saturation state (Gledhill et al., 2015; Siedlecki et al., 2021; Wang et al., 2013, 2017; Wanninkhof et al., 2015). Various anthropogenic and natural drivers, such as riverine inputs, warming, eutrophication, and changes in circulation, also

Visualization: Ivan D. Lima
Writing – original draft: Ivan D. Lima, Zhaohui A. Wang, Jennie E. Rheuban
Writing – review & editing: Ivan D. Lima, Zhaohui A. Wang, Louise P. Cameron, Jonathan H. Grabowski, Jennie E. Rheuban

contribute to the complexity of coastal acidification in the region (Rheuban et al., 2019; Salisbury et al., 2008; Salisbury & Jönsson, 2018; Siedlecki et al., 2021; Wallace et al., 2014; Wang et al., 2013).

These drivers potentially have significant impacts on carbonate chemistry as well as CO_2 and other inorganic carbon fluxes over a wide range of spatial and temporal scales in this region. However, the existing observational data of the full carbonate system are still mostly limited to sporadic seasonal cruises over the last three decades, and there are large data gaps in many locations even in the most extensively sampled seasons (Fennel et al., 2019; Pousse et al., 2022). Despite improved data collection on surface pCO_2 through ships of opportunity (e.g., Bakker et al., 2016; Chen et al., 2019; Signorini et al., 2013), other carbonate parameters and subsurface carbonate chemistry are poorly sampled. Only in the recent decade, with the emergence of ocean acidification as a potential threat to marine ecosystems, has there been a concerted effort to expand water sampling programs to include regular sampling of subsurface carbonate chemistry. However, the current observing programs still have a seasonal coverage in their designs, all with limited spatial coverages. These limitations make it difficult to evaluate the impacts of regional short-term (e.g., meso-scale and Gulf Stream eddies) and long-term (e.g., interannual and decadal scales) climate-related changes on water-column carbonate chemistry.

This study serves to address this critical data gap by developing a neural network model to predict regional carbonate chemistry from more readily available hydrographic and satellite data. A number of other studies have successfully developed empirical predictions of carbonate chemistry variables using methods that vary from relatively simple to significantly more complex and range from regionally specific algorithms to global assessments. The most common strategy for predicting carbonate chemistry variables is a multi-linear regression approach, either predicting dissolved inorganic carbon (DIC) and/or total alkalinity (TA) and calculating the carbonate chemistry variable of interest (e.g., pH, saturation state, pCO_2), or directly predicting variable of interest (e.g., Alin et al., 2012; Bostock et al., 2013; Carter et al., 2018, 2021; Davis et al., 2018; Evans et al., 2013; Hales et al., 2012; Juranek et al., 2009; Kim et al., 2015; Lee et al., 2006; McGarry et al., 2021; Millero et al., 1998; Turk et al., 2017; Vance et al., 2022; Velo et al., 2013; Williams et al., 2016). More recently, some studies have incorporated machine learning techniques such as neural networks or random forest regression (Bittig et al., 2018; Broullon et al., 2019; Chen et al., 2019; Fourrier et al., 2020; Li, Bellerby, Ge, et al., 2020; Li, Bellerby, Wallhead, et al., 2020; Lohrenz et al., 2018; McNeil & Sasse, 2016; Sasse et al., 2013; Sauzède et al., 2017; Velo et al., 2013; Xu et al., 2020). Using both approaches, these empirical predictions have been quite successful at reproducing the variables of interest with consistently low errors reporting RMSE less than 20 μmol kg^{-1} for DIC and/or TA, and often less than 10 μmol kg^{-1} with high r^2 values (e.g., $r^2 > 0.82$ and often >0.93).

However, many of these studies are not appropriate for use on the NAS. Regional algorithms developed for other areas, for example, the North Pacific (e.g., Alin et al., 2012; Davis et al., 2018; Hales et al., 2012; Juranek et al., 2011; Kim et al., 2015), Southern Ocean (Bostock et al., 2013; Williams et al., 2016), Gulf of Alaska (Evans et al., 2013), Mediterranean Sea (Fourrier et al., 2020), East China Sea (Li, Bellerby, Ge, et al., 2020; Li, Bellerby, Wallhead, et al., 2020), Gulf of Mexico (Chen et al., 2019; Lohrenz et al., 2018), and the North Atlantic (McGarry et al., 2021; Turk et al., 2017; Xu et al., 2020), may inherently incorporate drivers or relationships that are inconsistent with those in this study region or predict different variables than those of interest for this study. For example, salinity is generally a strong indicator of seawater total alkalinity as it reflects mixing of different water masses (Li et al., 2022; Rheuban et al., 2019; Wang et al., 2017), but salinity-alkalinity relationships may differ significantly by region if the underlying carbonate chemistry of the mixing endmembers vary (e.g., Lee et al., 2006). In fact, McGarry et al. (2021) found that regionally specific algorithms perform poorly when used outside of their original study area. Published global approaches are also not appropriate for the goals of this study, as many globally developed algorithms focus on data from the open ocean and selectively exclude data from nearshore waters with depth <200 m, exclude data from surface waters altogether, or report poor performance and caution against using those algorithms for surface waters (Bittig et al., 2018; Broullon et al., 2019; Carter et al., 2018, 2021). Indeed, Fourrier et al. (2020) found that two global neural network models were unable to predict nutrient and carbonate system variables in the Mediterranean Sea with a high degree of accuracy, and thus developed a regional neural network model based on the architecture of the global approaches.

In this study, we develop neural network models to predict DIC and TA in the NAS region using a combination of in situ hydrographic observations and satellite remote sensing data. To illustrate a use case, we use one of the models to reconstruct water-column DIC and TA, from which we estimate pH and aragonite saturation state (Ω_A) as relevant ecosystem indicators of potential impacts of ocean and coastal acidification on key sensitive shellfish

species. Finally, the reconstructed fields are used to explore modern (2013–2019) spatial and seasonal trends in surface and bottom water carbonate chemistry.

2. Materials and Methods

2.1. Field and Satellite Data

The neural network models use a combination of hydrographic observations and satellite data as input variables. Bottle measurements for the study region were obtained from several sources. The main source of data for this analysis is from Coastal Ocean Data Analysis Product in North America (CODAP-NA, Jiang et al., 2021). This data product compiles carbonate and nutrient chemistry from cruises that sampled the North American coastline—including the NAS region—and has undergone a rigorous quality control and quality assurance process. Additional data sources include the turnover cruises for the Ocean Observatories Initiative (OOI) Pioneer Array (Northeast Shelf, NES, data from https://alfresco.oceanobservatories.org), located south of Martha's Vineyard, Massachusetts, as well as samples collected during sampling cruises for the Northeast Shelf Long Term Ecological Research site, found in the same region as the NES Pioneer Array. Data with questionable or bad quality control flags were excluded from the analysis, and the combined data set includes 4350 DIC and 4151 and TA measurements and the corresponding observations of temperature, salinity, dissolved oxygen, and sample depth for the period 2003–2018 (Figure 1a). The compiled bottle data included all four seasons (Figure 1b) and spatial coverage throughout the region (Figure 2).

AVISO daily absolute dynamic topography (ADT) data with a resolution of 0.25° for the period 1993–2020 were downloaded from Copernicus Marine and Environment Monitoring Service (CMEMS, https://marine.copernicus.eu/). High resolution (0.01°) daily sea surface temperature (SST) data for the period 2002–2020 were obtained from NASA's Physical Oceanography Distributed Active Archive Center (PO.DAAC, https://podaac.jpl.nasa.gov). Daily surface chlorophyll (Chl) and diffuse attenuation coefficient at 490 nm (KD490) data with a resolution of 4 km for the period 1997–2020 were downloaded from NASA's Goddard Space Flight Center (https://oceancolor.gsfc.nasa.gov). These products contain merged data from the SeaWiFS and MODIS missions. Annual atmospheric xCO_2 measurements in Mauna Loa, Hawaii, were downloaded from NOAA's Global Monitoring Laboratory Mauna Loa CO_2 records (Tans & Keeling, 2021).

Satellite observations were extracted at the times and locations of bottle stations by matching the dates of the bottle stations and satellite images and extracting the satellite image value at the sample location using bilinear interpolation. Atmospheric xCO_2 was also extracted for the dates of bottle stations. The extracted atmospheric xCO_2 and satellite data were then merged with the bottle data to produce a final data set containing temperature (degrees Celsius), salinity (PSU), dissolved oxygen (μmol kg^{-1}), sample depth (meters), atmospheric xCO_2 (ppmv), ADT (meters), SST (degrees Celsius), Chl (mg m^{-3}), KD490 (m^{-1}), DIC (μmol kg^{-1}), and TA (μmol kg^{-1}).

2.2. Model Development

The neural network models used in this study consist of multiple layers of interconnected nodes also known as multi-layer perceptrons (MLP, Bishop, 1995). MLPs map the relationship between input variables and dependent output variables by minimizing the mean squared error (MSE) between the predicted and observed values of the output variables. This minimization is done through an iterative optimization method (back-propagation) that adjusts the node weights in the direction in which the loss function is decreasing most rapidly. Neural networks can accurately capture complex nonlinear relationships and interaction effects without making assumptions about the form of the true function, and therefore can be described as universal approximators (Hornik et al., 1989; Marzban, 2009). Neural networks have been successfully used to estimate carbonate chemistry variables (pCO_2, pH, DIC and TA) at regional (Fourrier et al., 2020; Landschutzer et al., 2013; Li, Bellerby, Ge, et al., 2020) and global scales (Bittig et al., 2018; Broullon et al., 2019; Laruelle et al., 2017; Sauzède et al., 2017).

Carbonate system variability on the NAS is controlled by thermodynamic processes, biological activities, physical circulation and mixing, and air-sea exchange, with each varying across different spatial and temporal scales (Cai et al., 2020; Gledhill et al., 2015; Siedlecki et al., 2021; Wang et al., 2013). In the neural network models, temperature and salinity represent the effects of thermodynamic processes (CO_2 solubility and dissociation). These variables are also good indicators of water masses and therefore contain information related to physical

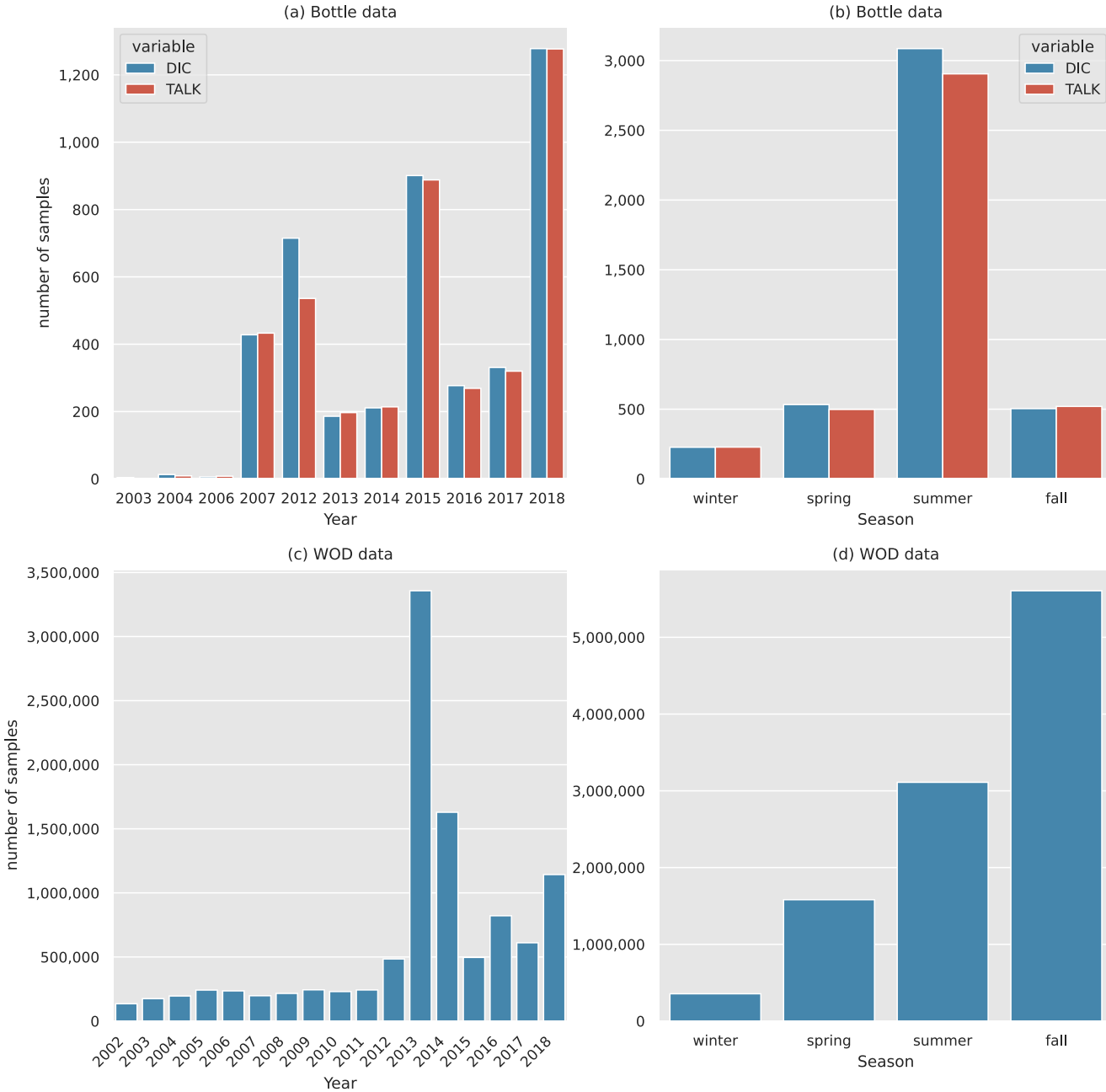

Figure 1. Frequency distribution of bottle dissolved inorganic carbon (DIC) and total alkalinity (TALK) samples (a, b) and data from the World Ocean Database (WOD, measured via CTD and autonomous floats) (c, d) data by year and season.

circulation and mixing as well. Dissolved oxygen, Chl and KD490 are strongly associated with photosynthesis and respiration and are good proxies for biological metabolism. Satellite ADT and SST include information related to circulation patterns and mixing processes, while depth is a proxy for variations in physical and biogeochemical processes related to vertical environmental gradients. Atmospheric xCO_2 accounts for the effect of rising atmospheric CO_2 concentrations on air-sea exchange, and thus potential DIC increase due to anthropogenic CO_2 invasion. In a neural network, the combination of linear transformations and a non-linear activation function allows it to capture complex interactions between features. Thus, the model is able to use the information provided by the satellite data on surface processes in conjunction with sample depth and other in situ observations to learn relationships between the different input variables and estimate DIC and TA at specific depths. Many studies that use neural networks to predict carbonate chemistry in the ocean include geographical

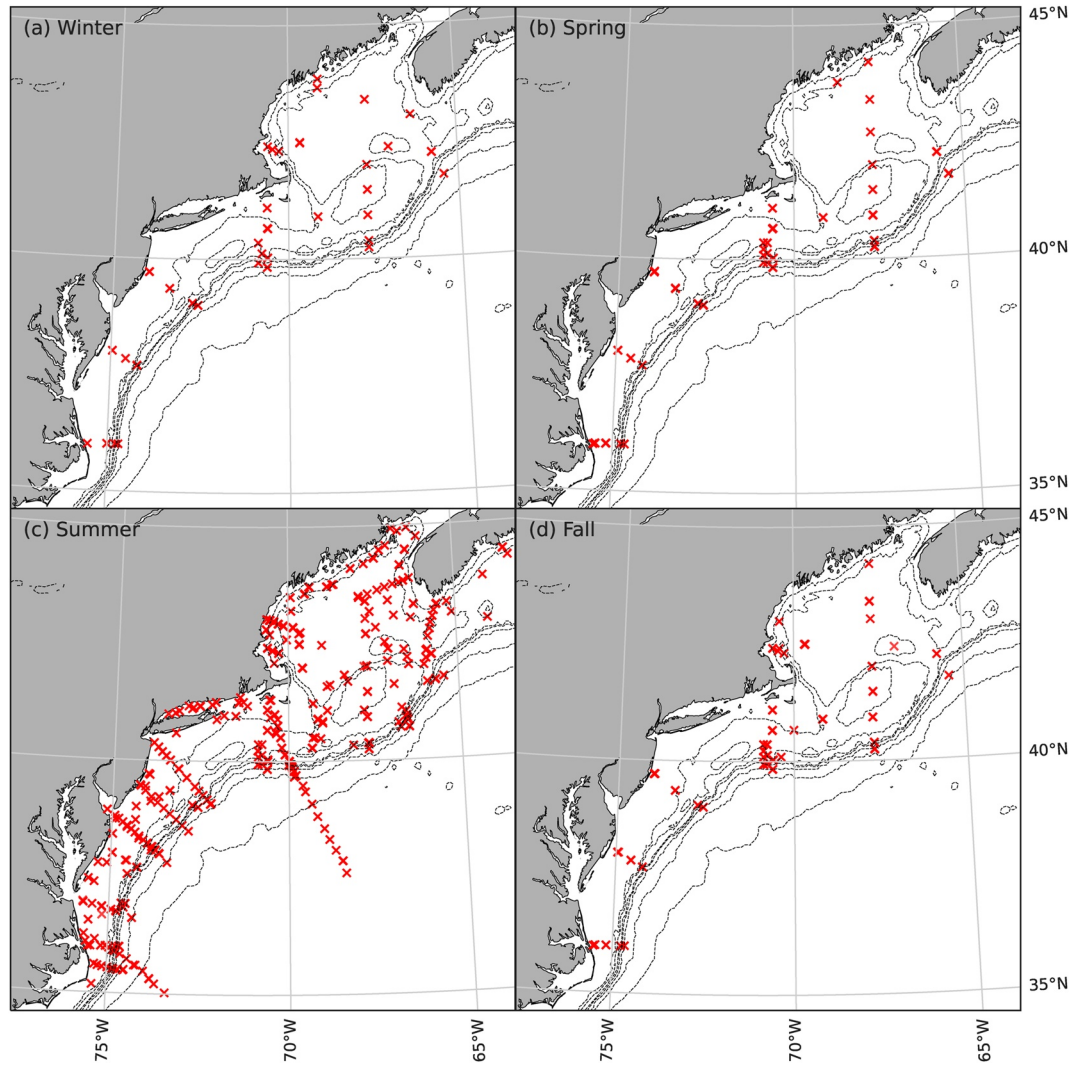

Figure 2. Spatial distribution of bottle data by season.

coordinates (latitude, longitude) and time as input variables (Bittig et al., 2018; Li, Bellerby, Ge, et al., 2020; Sasse et al., 2013; Sauzède et al., 2017; Zeng et al., 2014). Doing so makes the neural network more of a mapping or interpolation tool as it is able to learn the spatial-temporal patterns in the distribution of the data. In this study, we use a more mechanistic approach and only include predictor variables that directly affect or are proxies for processes that impact DIC and TA.

Prior to doing any model fitting, the merged bottle and satellite data set was split via a stratified random sampling approach into "training" (75%) and "test" (25%) datasets. The data set was stratified by season to ensure that both the training and testing data sets contained similar seasonal distributions of data. The "training" data set was used to train and tune the neural network models while the "test" data set was used to evaluate the accuracy of the models on unseen data. The "test" data set was not used at all during the training of the models, and serves as a completely independent data set for evaluation of the model's performance. We performed a 5-fold cross-validation analysis to assess how sensitive the models are to the choice of training data set and how well they generalize to new or unseen data. In this procedure, the merged bottle and satellite data set was randomly split into 5 sets of (approximately) equal size called folds, and a series of models was trained using each one of the folds as the test set and the remaining folds as the training set. Accuracy was computed for each model on the corresponding test set, producing a series of (5) accuracy estimates. The obtained range of accuracies provided information on the sensitivity of the model to changes in the training data set and how well we can expect the model to perform when applied to new data. Model accuracy was quantified using the coefficient of determination (r^2) and the root mean squared error (RMSE). We

defined the overall error in the model output as the combination of the measurement error for the input variables and the output data as well as the model error as in Sauzède et al. (2017). Measurement error for temperature, salinity, dissolved oxygen, atmospheric xCO_2, DIC, and TA were 0.06°C, 0.007 PSU, 3.4 µmol kg^{-1}, 0.07 ppm, 2.4 µmol kg^{-1}, and 5 µmol kg^{-1}, respectively (Jiang et al., 2021; Zhao & Tans, 2006). Measurement errors for ADT, KD490, and SST were 0.026 m (Pujol et al., 2023), 2.1 m^{-1} (calculated from data from Seegers et al., 2018, accessed via SeaBASS for this study region, https://seabass.gsfc.nasa.gov/search#val), and 0.6°C (Martin et al., 2012).

Neural networks are sensitive to the scaling and distribution of the input variables. Therefore, the satellite Chl and KD490 were log-transformed to account for their log-normal distribution. The diffuse attenuation coefficient at 490 nm (KD490) indicates how strongly visible light in the blue to green region of the spectrum is attenuated within the water column. This parameter is related to the presence of scattering particles in the water column, either organic or inorganic, and thus is an indicator of water turbidity. KD490 was highly correlated with Chl ($r = 1$), and these two variables provided very similar information on water quality to the model. Preliminary testing showed that including both KD490 and Chl as input variables did not improve the accuracy of the DIC and TA predictions significantly and that using either KD490 or Chl provided very similar results. Thus, we opted for a simpler model including only KD490 instead of both KD490 and Chl.

As part of the model development, we experimented with different neural network layouts, activation functions and parameter values (learning rate). For this procedure, the training set was further subdivided into a second training set and a validation set. The different model configurations were trained using the second training set and evaluated on the validation set. The model configuration that provided the best results consists of an input layer corresponding to the predictor variables, two fully connected hidden layers with 256 nodes each and the output layer corresponding to the predicted variable (DIC and TA). For the activation function we used the Leaky Rectified Linear Unit, or Leaky ReLU. This activation function is based on the Rectified Linear Unit (ReLU), but it modifies the function to allow small negative values (instead of zero) when the input is negative. The learning rate parameter determines the amount by which the neural network node weights are updated during training and was set to 0.01. We used a Batch Normalization layer before each hidden layer and the output layer. Batch Normalization normalizes, re-centers and re-scales the inputs (Ioffe & Szegedy, 2015). It provides built-in standardization of the input variables when included before the first hidden layer in the neural network, and significantly speeds up the training process by allowing us to use higher learning rates. Separate models were created for estimating DIC and TA. For the final model, part of the training set (20%) was used as a validation set during model training. The learning curves (Figure 3) revealed that the models converged quickly and did not overfit the data.

Our main goal in this study was to produce a model that can predict seawater carbonate chemistry from more readily available data so that we can reduce the sparseness of carbonate system observations in the NAS region. Dissolved oxygen measurements are relatively scarce in CTD, glider and autonomous float data, and satellite observations only became available in the late 1970s. This significantly limits the usefulness of a model that uses these variables as predictors. Therefore, we developed three versions of the neural network models for DIC and TA. In the first version (Model 1), we used all available variables as predictors (temperature, salinity, dissolved oxygen, sample depth, atmospheric xCO_2, ADT, SST and KD490). In the second version (Model 2), we excluded dissolved oxygen from the input variables, and in the third version (Model 3) we excluded dissolved oxygen and satellite data (ADT, SST and KD490) from the predictor variables. Model 2 allowed the use of a substantially larger fraction of the available field data, significantly increasing spatial and temporal coverage, while Model 3 would allow a user to extend an analysis to include earlier time periods, before the satellite era.

To evaluate the uncertainty in the models' estimates of DIC and TA, we trained an ensemble (one for each model version) of 100 models on the training data set and used the model ensembles to make predictions on the test data set. Within each ensemble, the models' node weights were randomly initialized, thereby producing distinct models, and DIC and TA were calculated for the test set using each of the models within the ensemble. Standard deviations for DIC and TA estimates from each model version were computed using the corresponding model ensemble's predictions on the test data set. We used the median standard deviations for DIC and TA in each ensemble as a metric for the uncertainty in the model predictions.

2.3. Reconstruction of Carbonate Chemistry Fields

To reconstruct the distribution of carbonate chemistry fields, we compiled a data set of hydrographic observations from CTDs and autonomous floats for the NAS region using data from NOAA's World Ocean Database

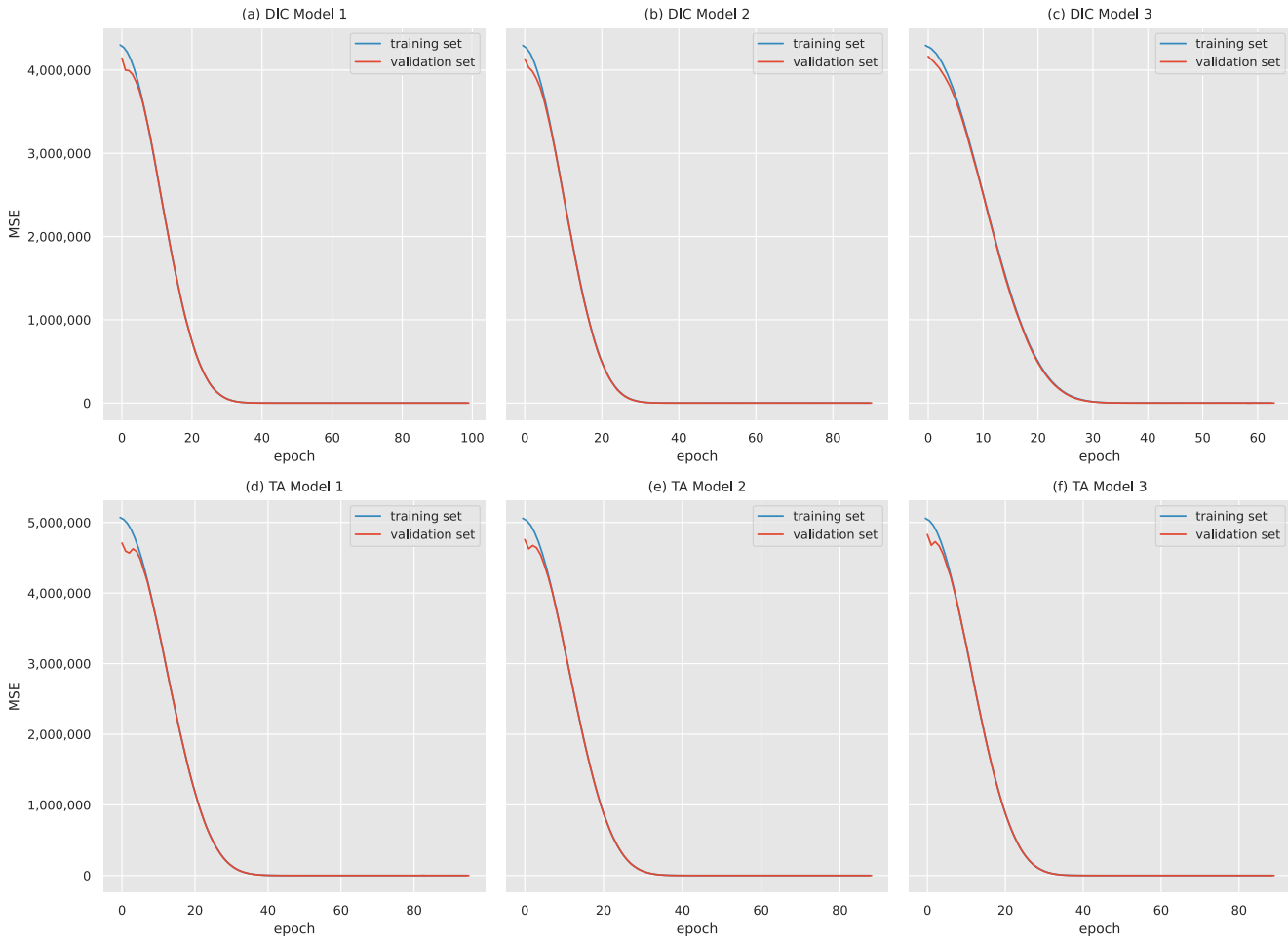

Figure 3. Neural network loss function (MSE) on the training and validation sets as a function of the training interaction (epoch) for the different model versions.

(WOD) for the period 2002–2018 (Figure 1c). The compiled data set provided good spatial coverage in all four seasons (Figure 4), although observations were more abundant in the Summer and Fall months (Figure 1d). Atmospheric xCO_2 and satellite data were extracted for the dates and locations of field observations, applying the same procedure used for the bottle data. Most of the compiled hydrographic observations did not include dissolved oxygen measurements. Thus, we applied Model 2 to the merged data set of hydrographic and satellite observations to estimate DIC and TA so that we could use all available hydrographic data. Total pH and aragonite saturation state (Ω_A) were computed from the obtained DIC and TA estimates (with in situ temperature, salinity, and pressure) using PyCO2SYS (https://pyco2sys.readthedocs.io/en/latest/, Humpheries et al., 2021). Carbonic acid dissociation coefficients K1 and K2 from Lueker et al., 2000, sulfate dissociation constant KSO4 from Dickson (1990), total borate concentration from Lee et al. (2010), and HF dissociation coefficient KHF from Perez and Fraga (1987) were used in carbonate system calculations. Estimates of surface and near bottom temperature, salinity, DIC, TA, pH and Ω_A were obtained by averaging the data in the top 5 m and extracting the deepest value from each cast, respectively. The extracted data were used to generate seasonal maps of surface and near bottom temperature, salinity, DIC, TA, pH and Ω_A by grouping the data by season and remapping them into a regular grid with 1/8-degree resolution encompassing the NAS region using inverse distance weighted interpolation with a radius of 100 km.

3. Results and Discussion

The model including all predictor variables (Model 1) fitted the data remarkably well with $r^2 = 0.963$ and RMSE = 15.4 μmol kg^{-1} for DIC and $r^2 = 0.986$ and RMSE = 9.0 μmol kg^{-1} for TA predictions on the test data

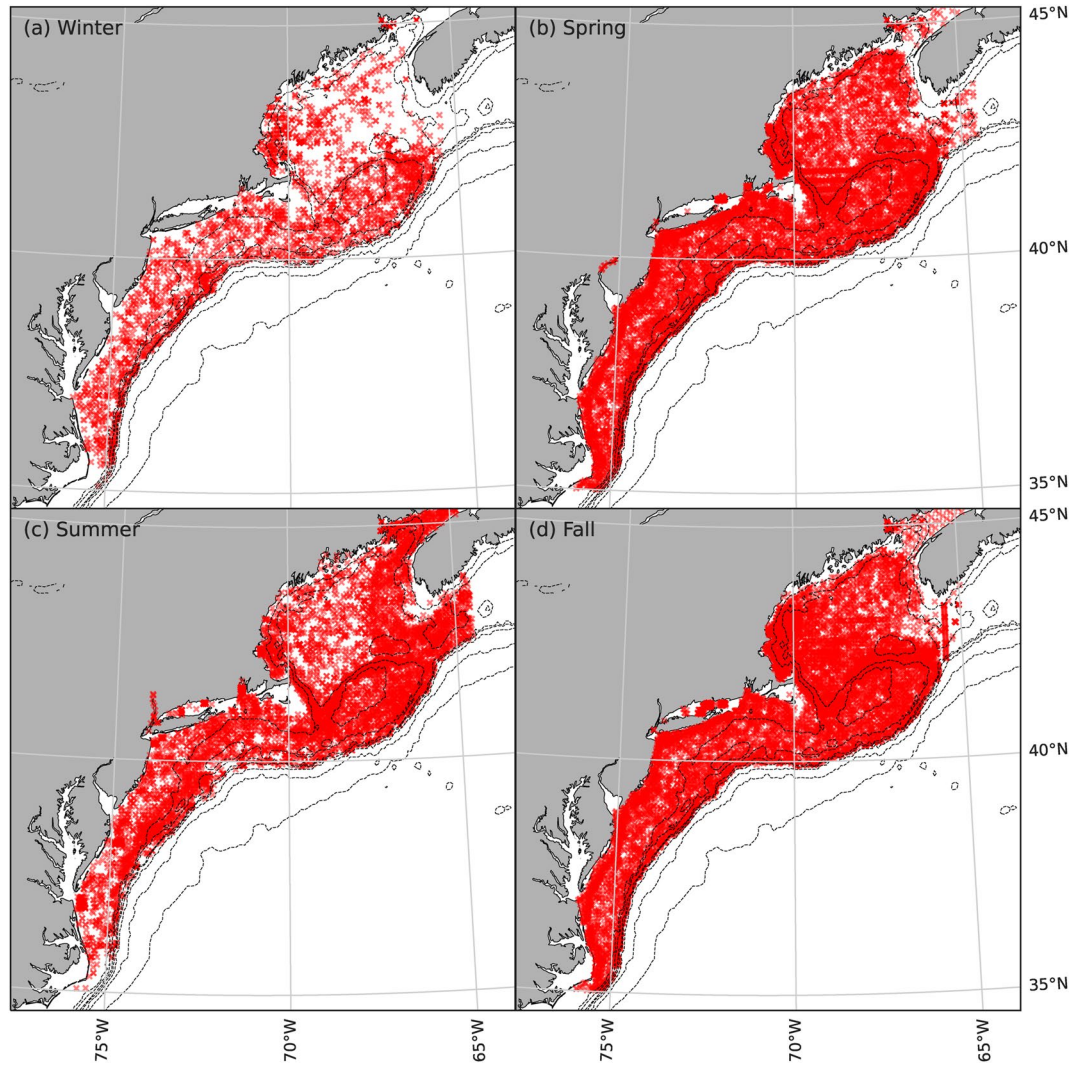

Figure 4. Spatial distribution of compiled CTD and autonomous float data by season.

set. The 5-fold cross-validation analysis produced r^2 values in the range 0.955–0.968 with a mean $r^2 = 0.961$ for DIC and 0.981–0.989 with a mean $r^2 = 0.986$ for TA, indicating low sensitivity to the choice of training data set and good generalization performance on new data. Removing dissolved oxygen as a predictor (Model 2) results in a small increase in model error with $r^2 = 0.944$ and RMSE = 19.1 μmol kg^{-1} for DIC and $r^2 = 0.985$ and RMSE = 9.7 μmol kg^{-1} for TA predictions on the test data set. 5-fold cross-validation shows r^2 values between 0.931 and 0.951 (mean $r^2 = 0.943$) and 0.975 and 0.986 (mean $r^2 = 0.983$) for DIC and TA, respectively, indicating good generalization performance for Model 2 as well. The model excluding dissolved oxygen and satellite data as predictors (Model 3) has higher error than the previous two models, but it still is a good predictor of DIC and TA with $r^2 = 0.913$ and RMSE = 23.9 μmol kg^{-1} for DIC and $r^2 = 0.983$ and RMSE = 10.4 μmol kg^{-1} for TA on the test data set. The r^2 values obtained in the 5-fold cross-validation analysis range between 0.899 and 0.914 (mean $r^2 = 0.904$) for DIC and 0.973 and 0.983 for TA, also denoting low sensitivity to changes in the training data set and low error when predicting DIC and TA from unseen data (Figure 5, Table 1). Comparison of the r^2 and RMSE values for DIC and TA (Table 1) revealed that the model error for TA was very similar in all three versions of the NN models, as TA was strongly correlated with salinity, largely reflects mixing of different water masses, and is significantly less affected by the other variables. Model error for DIC decreased with the inclusion of dissolved oxygen and satellite data as predictors, which improved the neural network model performance by incorporating information on biogeochemical processes in addition to physical oceanography. We anticipate that model performance could be further improved with the inclusion of additional indicators of biogeochemical

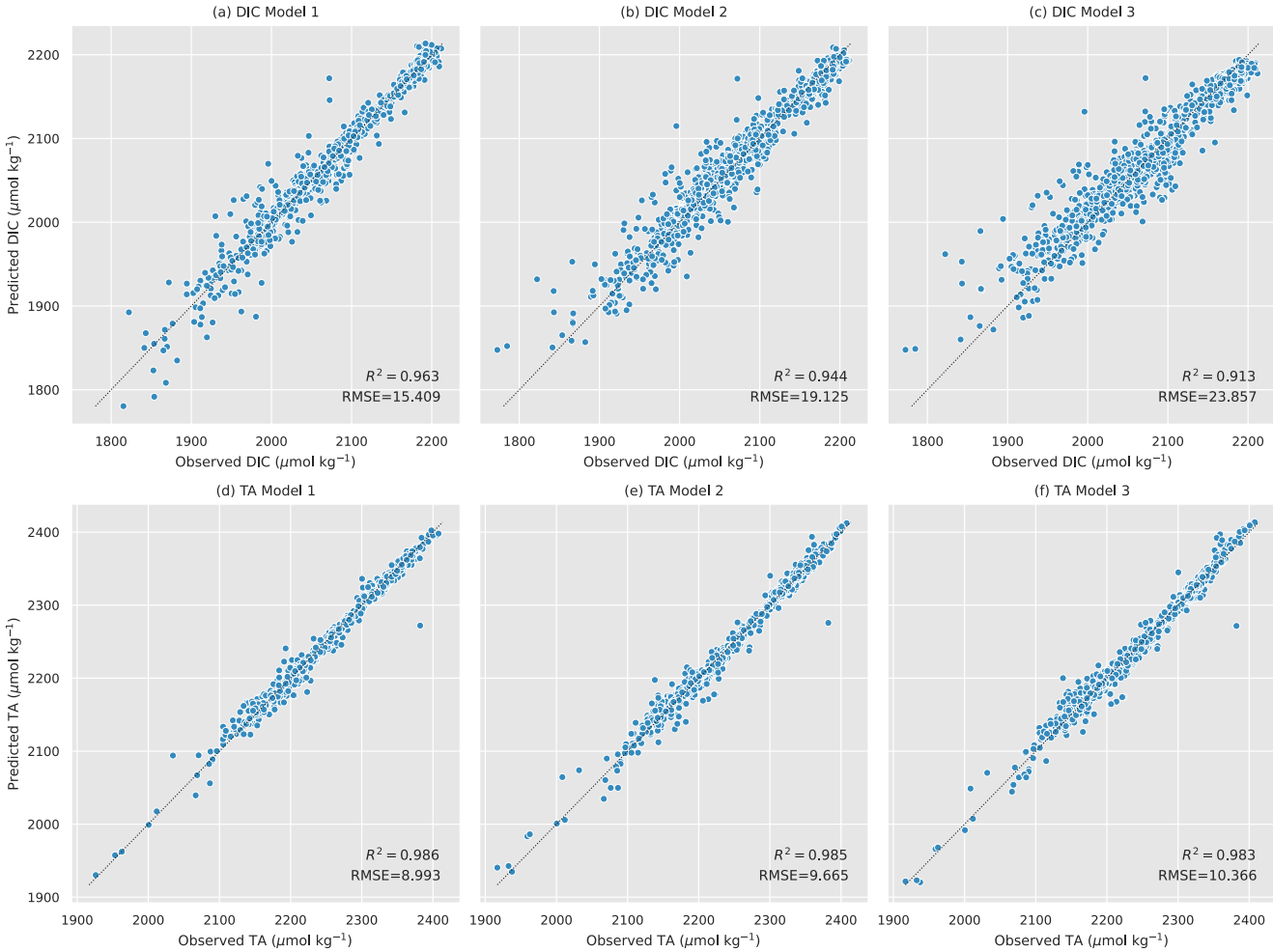

Figure 5. Predicted versus observed dissolved inorganic carbon (DIC) and total alkalinity (TA) values in the test set for the different model versions.

processes, for example, nitrate, phosphate, or net community productivity, among others (Fourrier et al., 2020; Sauzède et al., 2017). However, the differences between the different model versions were relatively small (Tables 1 and 2) and all three models predict DIC and TA in the NAS region with low overall error.

The frequency distributions of standard deviations for the ensemble model predictions showed that the uncertainty in TA estimates also remains relatively unchanged in the different model versions (Figure 6b, Table 3). The median standard deviation values for TA estimates in all three model versions were very low (2.85–3.11 μmol kg^{-1}) and very close to each other (Table 3). The uncertainty in DIC estimates differed among model versions (Figure 6a and Table 3), with standard deviations that were generally higher in Models 2 and 3 compared to Model 1 (Figure 6a). Nevertheless, the uncertainty from the distribution of standard deviations in DIC estimates from all three model versions was still very low with coefficients of variation between 0.1% and 1%. The greater variability in the uncertainty of DIC estimates compared to TA between the different model versions is consistent with the idea that TA is mainly driven by physical mixing that can be predicted by variations in salinity while DIC is a function of many other factors, and the neural network models have more difficulty learning and accurately predicting these relationships.

Despite the somewhat limited availability of discrete carbonate system observations in both space and time in this region, the neural network model still

Table 1
r^2 and RMSE Values for the Different Versions of the Neural Network Models on the Test Data Set

	DIC			TA		
	r^2	RMSE	RMSE final	r^2	RMSE	RMSE final
Model 1	0.963	15.4	16.7	0.986	9.0	11.1
Model 2	0.944	19.1	19.9	0.985	9.7	11.2
Model 3	0.913	23.9	24.5	0.983	10.4	11.8

Note. RMSE values are in μmol kg^{-1}. RMSE final shows the overall error including measurement and model error.

Table 2
Best, Worst, and Mean r^2 Values From the Cross-Validation Analysis for the Different Versions of the Neural Network Model

	DIC			TA		
	Best r^2	Worst r^2	Mean r^2	Best r^2	Worst r^2	Mean r^2
Model 1	0.968	0.955	0.961	0.989	0.981	0.986
Model 2	0.951	0.931	0.943	0.986	0.975	0.983
Model 3	0.914	0.899	0.904	0.983	0.973	0.980

performed well in predicting unseen data in the test set (Table 1). Both model residuals and absolute errors for DIC and TA were uncorrelated with any of the predictor variables ($p > 0.1$ for all predictors, data not shown), indicating high predictive capabilities throughout the water column, and under all environmental conditions tested. The underlying data also included observations from all four seasons, and when subsetting the test set into seasonal data, the models also performed well. For both DIC and TA, model RMSE values for the test set were consistent across all four seasons (Table 4), and consistent with the overall RMSE for the entire test set, supporting the neural network's ability to learn the underlying physical and biogeochemical processes driving the carbonate system in this region through the proxy variables chosen as predictors. RMSE for winter data was slightly higher in Model 3, the model with the fewest predictor variables, but this time period also has the fewest observational data available.

3.1. Use Case: Seasonal Climatology of the Carbonate System

In this example, we provide a use-case of the neural network model output as a tool to interpolate and fill significant observational data gaps. This use case serves to assess both the seasonality of regional carbonate chemistry and also is a test as to whether our model output is consistent with previous studies. Discrete observations of carbonate chemistry variables (i.e., DIC, TA, pH, pCO_2, and aragonite saturation state) are very sparse both spatially and temporally (see Section 1—Introduction, and Figures 1 and 2), while hydrographic CTD data (e.g., temperature, salinity, and depth), and dissolved oxygen sensor data along with satellite observations are much more frequently measured due to the logistical differences in sample collection and analysis. This is particularly evident in subsurface waters and at depth, where observations are even more sparse than in surface waters, as ships of opportunity with underway pCO_2 systems have significantly increased the observational capability in recent decades. These sparse observations at depth make it particularly difficult to evaluate the full extent of seasonal variation in seawater carbonate chemistry, carbon cycling and benthic or midwater ecosystem or habitat characteristics with respect to pH or calcium carbonate saturation state. Here, we examine the climatological mean conditions of physical parameters (salinity and temperature) and the neural network-generated carbonate parameters at the sea surface and bottom, which showed seasonal progression and large spatial gradients on the Northwest Atlantic Shelf (Figures 7 and 8).

3.1.1. Surface

Wintertime had the lowest temperature but relatively high salinity among the four seasons (Figure 7). The latitudinal gradients of surface temperature and salinity were relatively weak in winter but their cross-shelf gradients were

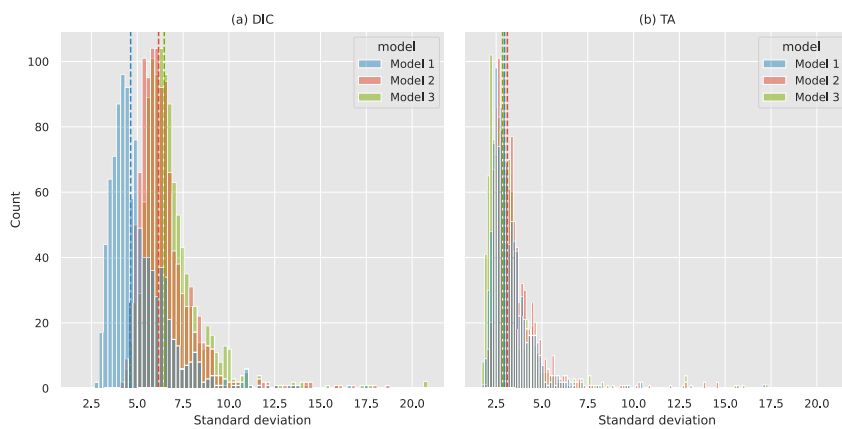

Figure 6. Frequency distributions of standard deviations (μmol kg^{-1}) values for the different model ensembles' dissolved inorganic carbon (DIC) and total alkalinity (TA) predictions on the test data set (a and b). Each model ensemble corresponds to a different model version. The blue, red, and green vertical dashed lines indicate the median standard deviation values for DIC and TA, respectively.

Table 3
Standard Deviation for Different Model Versions Estimated From the Model Ensembles on the Test Data Set

		Model 1	Model 2	Model 3
DIC std	Min	2.703	4.199	4.342
	Max	17.416	18.542	20.846
	Median	4.640	6.164	6.473
TA std	Min	1.816	1.989	1.726
	Max	17.167	14.577	17.335
	Median	2.957	3.112	2.853

Note. Values are in μmol kg^{-1}.

clearly visible in the Southern New England (SNE) shelf and the Mid-Atlantic Bight (MAB), which reflects the shelfbreak frontal exchange between low temperature, low salinity shelf water and high temperature, high salinity slope water that receives influences from the Gulf Stream (Gangopadhyay et al., 2019; Gawarkiewicz et al., 2012; Lentz, 2008; Zhang et al., 2011). The southern end of the MAB had the highest temperature and salinity in winter, which is consistent with the area as the exit point where shelf water mixes with the Gulf Stream (Biscaye et al., 1994; Lentz, 2008; Vlahos et al., 2002; Wang et al., 2013). By spring, surface temperature is slightly elevated across the region, while its latitudinal and cross-shelf gradients are still not well developed. Surface salinity was overall lower than winter values across the region, especially along the coastline, reflecting freshwater inputs during the spring freshet. In summer, a strong latitudinal gradient of surface temperature developed with the high values of >20°C across the shelf in the southern part of the region, while the low values (~10–15°C) occur in much of the Gulf of Maine. Summer had the lowest surface salinity among the four seasons, and its across-shelf gradient is relatively weak. The gradients of temperature and salinity across the shelfbreak front in summer seem to be less developed than winter and spring. By fall, surface water in much of the region was cooling, and the latitudinal gradient of surface temperature started to weaken. Surface salinity in fall started to increase from its summer values, and a strong cross-shelf gradient and a shelfbreak frontal gradient developed, especially in the middle part of the region.

The climatological surface TA field closely followed the distribution of salinity in all four seasons, which is indicative of its mostly conservative nature, a finding that has been widely documented in previous studies (e.g., Cai et al., 2010; McGarry et al., 2021; Wang et al., 2013, 2017). The TA gradients across the shelfbreak in the SNE shelf and the MAB were more pronounced in winter and fall, while such a feature was weak in summertime. Winter generally had the highest TA values at the surface among the seasons, whereas they were lowest in summer. The latitudinal gradients of TA at the surface were not as clear as the cross-shelf gradients for most seasons, except in summer, when the TA values of surface water was generally lower inthe middle and southern parts of the MAB than most of the Gulf of Maine, which may be driven by low salinity and low alkalinity inputs from a few mid-size rivers in the area (i.e., Hudson River, Delaware River, and Chesapeake Bay).

The surface DIC distribution, however, was negatively correlated with that of surface temperature, likely reflecting that the solubility pump of CO_2 plays a role in shaping the DIC distribution in the region. Winter had the highest surface DIC values among the seasons, followed by spring, corresponding to the annual peak of the solubility pump when the biological pump is relatively weak. The cross-shelf and shelfbreak frontal gradients of DIC were weaker than those observed for surface TA for all seasons. The latitudinal gradient was largely unclear in winter and spring, while the DIC concentration of surface water was lower in the MAB than the Gulf of Maine in summer and fall. Again, this was likely associated with more seasonal freshwater inputs in the MAB. DIC concentrations at the surface were relatively low in summer and fall compared to winter and spring, corresponding to the seasonal high and low rates of the biological pump. The seasonal DIC distribution at the surface reflected a combination of the effects from physical processes (e.g., riverine inputs and shelfbreak frontal

Table 4
Seasonal RMSE From the Test Set for Each Model Version for DIC and TA

	Model 1		Model 2		Model 3	
	DIC	TA	DIC	TA	DIC	TA
DJF	13.6 (49)	11.3 (50)	23.6 (57)	13.8 (57)	25.6 (57)	12.6 (57)
MAM	12.8 (99)	8.8 (97)	19.0 (133)	9.9 (125)	19.0 (133)	9.3 (125)
JJA	15.6 (726)	7.5 (681)	19.3 (772)	8.4 (726)	25.1 (772)	9.6 (726)
SON	16.6 (119)	14.1 (122)	15.6 (126)	13.3 (130)	19.5 (126)	13.9 (130)

Note. Units are in μmol kg^{-1} and number of observations are shown in parentheses. Seasons have been split into winter (December, January, February—DJF), spring (March, April, May—MAM), summer (June, July, August—JJA), and fall (September, October, November—SON).

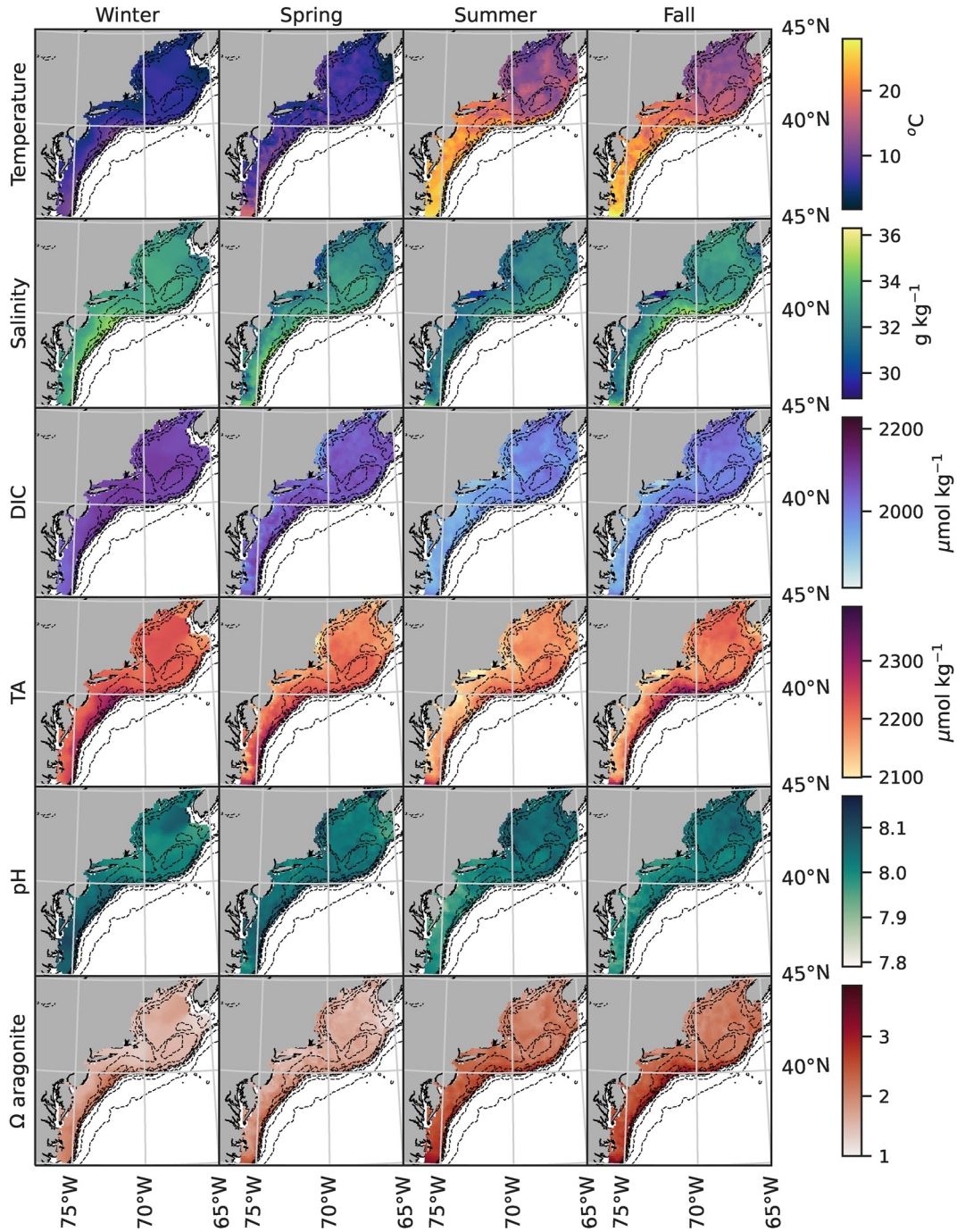

Figure 7. Seasonal distributions of surface temperature, salinity, dissolved inorganic carbon (DIC), total alkalinity (TA), pH and aragonite saturation state (Ω) reconstructed from World Ocean Database (WOD) data. DIC and TA were estimated by using Model 2 on the merged data set of hydrographic and satellite observations. Total pH and Ω_A were computed from the DIC and TA estimates plus in situ temperature, salinity, and pressure using PyCO2SYS (Humpheries et al., 2021).

exchange), the solubility pump, and biogeochemical processes (e.g., seasonal changes in the biological pump) in the region.

The surface pH in the region exhibited seasonal features observed in the surface DIC and TA distributions, but their correlations were complex and reflected more factors that control pH seasonality and its spatial distribution compared to DIC and TA. For example, cross-shelf and shelfbreak frontal gradients of surface pH were

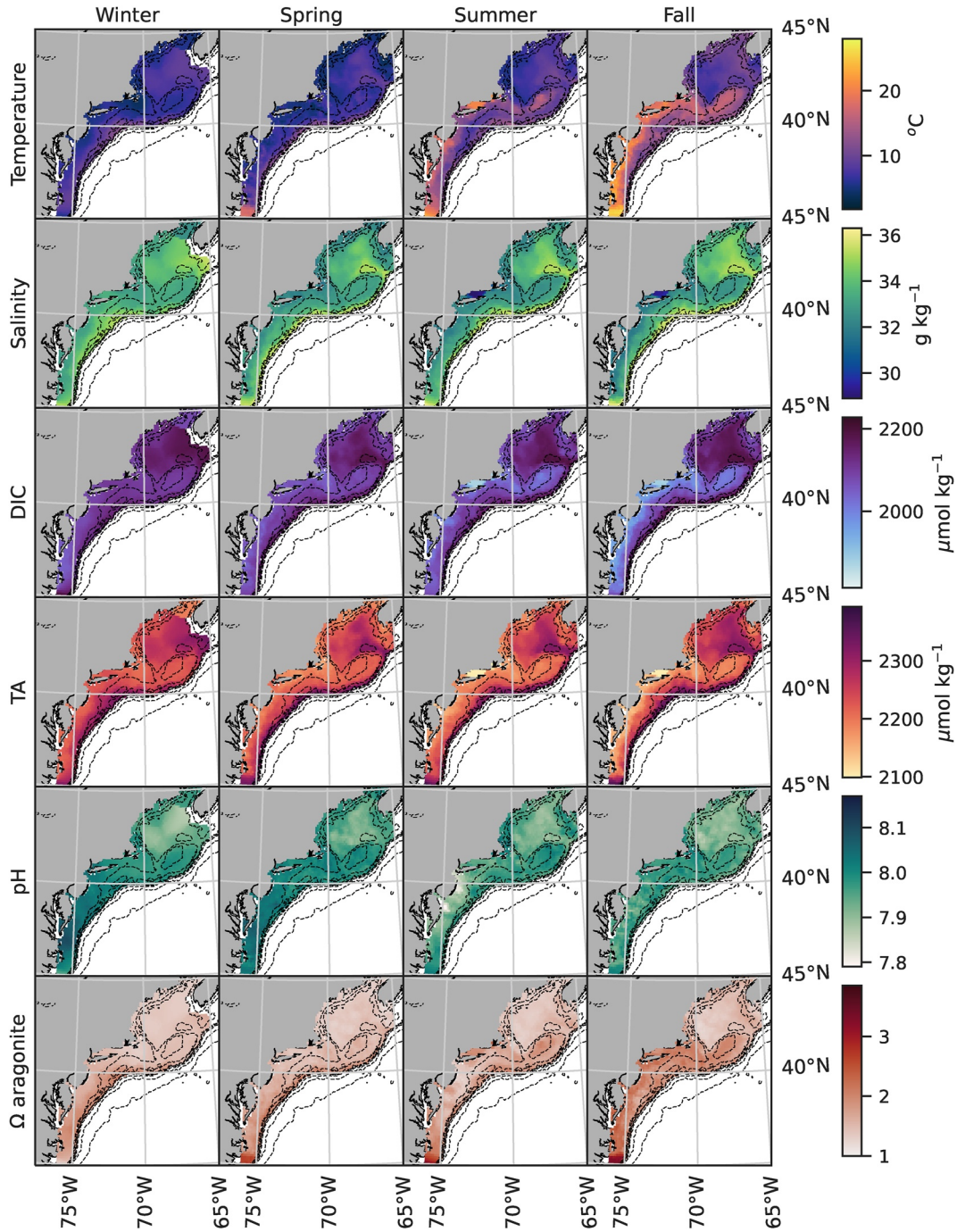

Figure 8. Seasonal distributions of bottom temperature, salinity, dissolved inorganic carbon (DIC), total alkalinity (TA), pH and aragonite saturation state (Ω) reconstructed from World Ocean Database (WOD) data. DIC and TA were estimated by using Model 2 on the merged data set of hydrographic and satellite observations. Total pH and aragonite saturation state were computed from the DIC and TA estimates plus in situ temperature, salinity, and pressure using PyCO2SYS (Humpheries et al., 2021).

pronounced, which were similar to that of surface TA. Winter and spring had higher surface pH values than summer and fall, corresponding to the seasonality of surface DIC, which bears physical (e.g., mixing and freshwater inputs), thermal (solubility) and biogeochemical imprints. Heating and cooling have two distinct effects on pH: (a) changes due to carbonate equilibria and (b) through the solubility of CO_2. The two processes, however, have the opposite effects on pH values, and as such, the pH distribution responds to the thermal effect in

a more complex way compared to DIC, which is affected by only the solubility of CO_2 during heating or cooling. On the other hand, the surface distribution of aragonite saturation states was largely negatively correlated with that of surface DIC, highlighting that the variability in surface DIC plays a major role in controlling the variability of aragonite saturation states. For example, the sources and sinks of DIC (e.g., respiration, photosynthesis, air-sea CO_2 exchange and upwelling of high CO_2 water) mainly affected the distribution of aragonite saturation states at the surface, a finding that is consistent with the previous studies (Cai et al., 2020; Wang et al., 2017). Interestingly, seasonal highs (summer and fall) and lows (winter and spring) of Ω_A at the surface were the opposite of the seasonality of surface pH and TA, highlighting the decoupling effects on Ω_A and pH from physical and biogeochemical processes.

3.1.2. Bottom

Generally cooler and saltier, the distributions of salinity and temperature varied less spatially (cross-shelf and latitudinal) and seasonally compared to the surface distributions (Figure 8). The correlations between physical (salinity and temperature) and carbonate parameters, as well as among carbonate parameters, were similar to surface conditions. One major feature that was distinct from the surface is that there was consistently high salinity bottom water across all seasons in the deep basins of the Gulf of Maine, which originated from the slope water of the Gulf Stream (Siedlecki et al., 2021; Townsend et al., 2010). This bottom water DIC and TA concentrations were consistently high across all seasons. Accumulation of respired carbon from surface production may further elevate bottom water DIC in these basins but would likely have a limited effect on the TA signature since respiration does not strongly influence TA. These conditions resulted in consistent low pH (mostly 7.8–7.9) and low Ω_A (mostly <1.5) across all seasons in the deep bottom water of the Gulf of Maine.

In the other parts of the NAS, the seasonality of DIC, TA, pH and Ω_A in bottom water was also less pronounced than their surface distributions. Except for bottom water in the Gulf of Maine, the latitudinal gradients of these parameters were less clear, but the cross-shelf gradients were more pronounced compared to those at the surface. The shelfbreak frontal gradient was clearly visible across all seasons for DIC, TA, and pH, highlighting the significant impacts of the frontal exchange on the carbonate system at depth, which is consistent with previous transects during cruises in the region (e.g., Wang et al., 2013; Wright-Fairbanks et al., 2020). The consistent gradient across all four seasons suggests previous observations of this feature are likely representative of a consistent pattern occurring throughout the year. One striking feature is that most of the bottom water has a Ω_A value less than 1.5, except the southern tip of the region, and some nearshore and mid-shelf water of the MAB in fall. This finding suggests bottom water in the region likely experiences episodic aragonite undersaturation as a result of high-CO_2 events, such as accumulation of respired surface blooms and high CO_2 water intrusion from upwelling of subsurface slope water.

3.2. Comparison With Similar Studies

Although a number of other studies have developed predictive models of carbonate system variables from more readily available measurements (see Introduction), it was necessary to complete this regional assessment in order to accurately predict conditions in the NAS region. Our results are directly comparable to, and consistent with, two regional studies that also focused on the NAS, McGarry et al. (2021), and Li et al. (2022). Similar to this study, McGarry et al. (2021) developed multiple models using different sets of predictor variables with increasing model complexity, and Li et al. (2022) developed separate models for different regions with slightly different water mass characteristics. Both used a multi-linear regression (MLR) approach that incorporated the predictors temperature, salinity, dissolved oxygen, and/or nitrate concentration. Both studies also found that models that incorporated proxies for biogeochemical drivers performed better than those that only included physical mixing. The predictability of their models was generally high, achieving an r^2 and RMSE of 0.86–0.98 and 5.9–22.3 μmol kg^{-1} for DIC and 0.956–0.976 and 6.6–10.9 μmol $kg^{-1}$for TA, respectively, depending on model choice.

Our assessment differs from both studies in several ways. First, the model architecture differs. Linear regression models can be very powerful with high-dimensional data but can only generate linear functions, which may limit their ability to predict complex, non-linear systems. Second, our input data set includes more than 4.5-fold more observations for training the models (n = 3,262 for DIC and n = 3,113 for TA vs. 599 and 529, 777 for this study vs. McGarry et al. (2021) or Li et al. (2022), respectively), and more independent data for evaluating

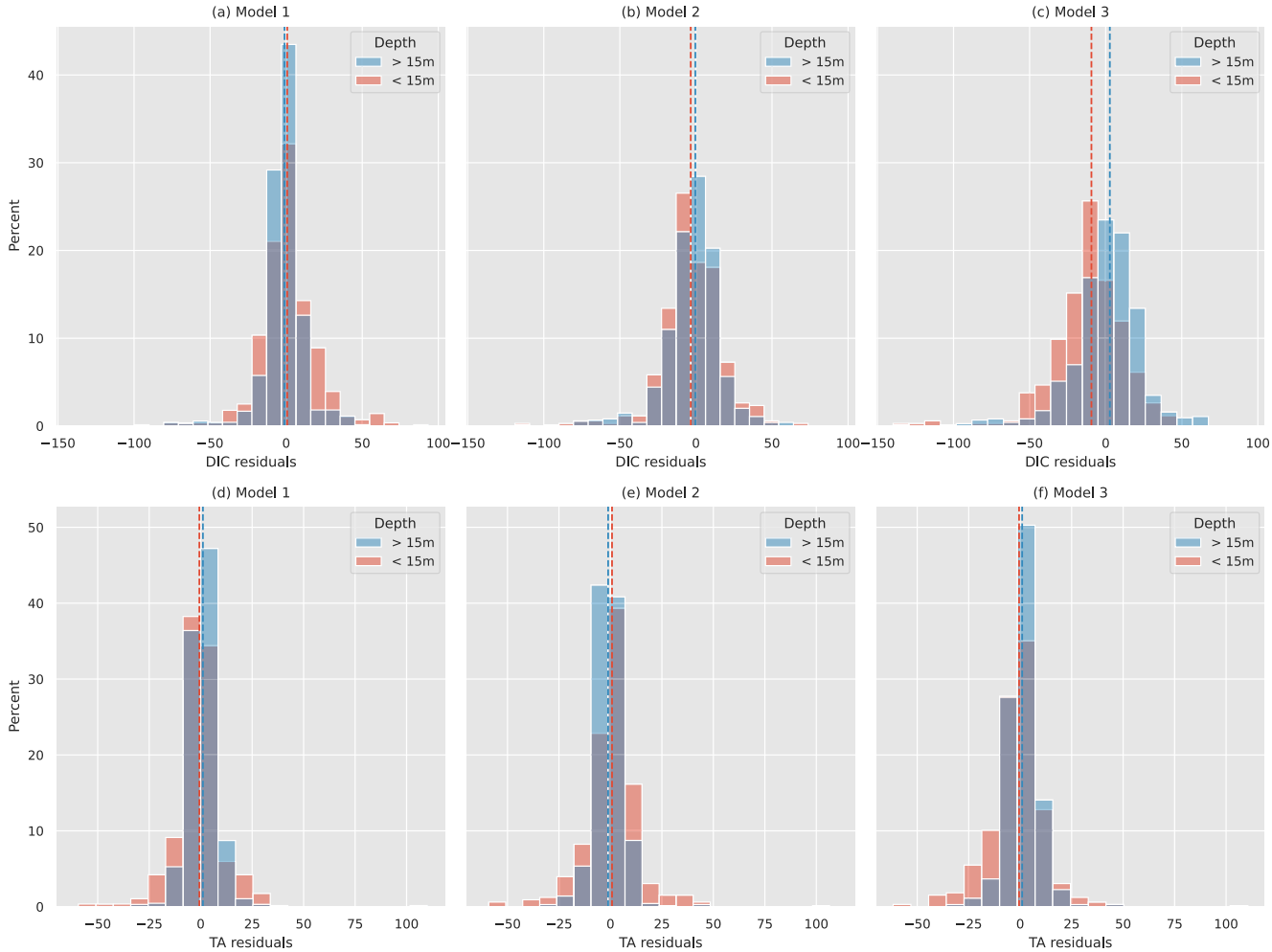

Figure 9. Observed dissolved inorganic carbon (DIC, a, b, c) and total alkalinity (TA, d, e, f) minus predicted DIC and TA from the test set for model 1 (a, d), model 2 (b, e), and model 3 (c, f). Data have been split into surface (depth <15 m) and subsurface (depth ≥15 m) layers. Vertical lines are the medians for each distribution.

the performance of the models (n = 1,088 for DIC and n = 1,038 for TA vs. 741 and 204, 208, for this study and McGarry et al. (2021) or Li et al. (2022), respectively). Third, although the majority of the observational data used in this study were collected during summer months, our data compilation includes observations from all four seasons (Figure 1) that are incorporated into both the training and testing datasets. The use of a data set with full seasonal coverage that captured more of the variability in the carbonate system in this region, in combination with a model architecture that can capture complex, non-linear systems resulted in a predictive model with high accuracy and precision across all four seasons. Indeed, McGarry et al. (2021) and Li et al. (2022) report significantly worse model performance on observations from outside of the summer season for which their model was calibrated. Fourth, the input variables, while similar between the two approaches, represent proxies for slightly different processes; for example, in this study, we included the additional proxies for mesoscale and submesoscale variability (e.g., ADT and high-resolution SST) and air-sea CO_2 flux (e.g., atmospheric xCO_2) such that the neural network may be able to better learn the complex drivers of the carbonate system in this region.

Finally, this analysis includes data from throughout the water column, while both McGarry and Li's models predict carbonate chemistry for only subsurface waters, focusing on depths greater than 15 and 20 m for McGarry and Li, respectively. The neural networks described in this study have between 68,353 (Model 1) and 67,329 (Model 3) trainable parameters that are optimized during the fitting process. As such, a single neural network model can be trained to predict DIC or TA throughout the water column. Indeed, for all three models, model residuals and RMSE on the test set were similar in surface (depth <15 m) and subsurface waters (depth ≥15 m) for both DIC and TA (Figure 9, Table 5), although for surface conditions, RMSE was slightly higher in both

Table 5
Median Model Residual and RMSE for the Test Set for DIC and TA for Surface and Subsurface Waters

	DIC		TA	
	<15 m	≥15 m	<15 m	≥15 m
Median Residual (μmol kg^{-1})				
Model 1	0.7	−1.1	−0.6	1.1
Model 2	−3.4	−0.4	0.8	−1.1
Model 3	−9.4	2.7	−0.5	1.0
RMSE (μmol kg^{-1})				
Model 1	18.4	14.0	11.0	8.0
Model 2	20.3	18.5	12.5	8.1
Model 3	25.9	22.9	13.0	8.9

parameters. Consistent with other metrics of model performance, Model 3 (hydrographic data only) showed the largest deviation in median residual for DIC when comparing surface and subsurface waters (−9.4 vs. 2.7 μmol kg^{-1}, respectively). Both metrics are consistent with the overall observations of model performance: increasing model complexity improved the models' abilities to predict carbonate system parameters, specifically by incorporating features that were relevant proxies for biogeochemical processes, which may be more important predictors in surface waters than at depth. When compared to McGarry et al. (2021) and Li et al. (2022), our neural network models achieved modest improvements in RMSE for predicted DIC and TA on the test set for summer, subsurface waters (Table 6) when compared to models with similar predictor variables. Here, we note that we found slightly higher RMSE values than McGarry et al. (2021) and Li et al. (2022) did for their evaluation data when applying their models to our carbonate chemistry data set because the underlying data sets we use are different. McGarry and Li report RMSE for a single cruise, while this comparison was made on a subset of the test data set, and includes data from 14 different cruises spanning 2007–2018. Additionally, the neural network models provide DIC and TA estimates for the full water column and all four seasons in the NAS region, while McGarry et al. (2021)'s and Li et al. (2022)'s MLR models are restricted to subsurface waters in the summer months. This makes the neural network models applicable to a much wider range of data, as illustrated in our use case (Section 3.1). A significant advantage of the MLR models is that they allow for relatively straightforward and interpretable inference from the model's coefficients. For example, the MLR approach provides context for the relative importance of each predictor variable, which can give indications of the strength of the underlying drivers for which that variable is a proxy. A similar assessment is more difficult with the neural network described here due to the complexity of the network's architecture.

3.3. Potential Applications of Different Model Versions & Future Work

In this study, we developed three different versions of the neural network model to predict carbonate chemistry variables from more readily available hydrographic and satellite data. Each model version has a different utility–here we describe a single use-case of one of the model versions, Model 2, to generate seasonal and spatially explicit estimates of modern-day surface and bottom water conditions. In this case, we chose Model 2 due to the lack of co-located oxygen data with many of the CTD casts obtained from the World Ocean Database but the availability of satellite observations. Additionally, given the strongly regional nature of the underlying data set and this algorithm development, we recommend that these algorithms only be used in the Northwest Atlantic region. This framework would likely be applicable in other regions, but users may need to develop their own regional algorithms trained on data collected in their region of interest.

Table 6
Model RMSE Between Predicted and Observed Dissolved Inorganic Carbon (DIC) and Total Alkalinity (TA) for the Test Set, Focusing on Subsurface Conditions for Observations During Summertime Compared to McGarry et al. (2021) and Li et al. (2022)

	DIC	TA	
Model type	RMSE (μmol kg^{-1})		References
Model 1 (>15 m)	14.2 (530)	5.9 (485)	This study
DIC II, TA V	16.9 (530)	7.9 (485)	McGarry et al. (2021)
Model 1 (>20 m)	12.9 (489)	5.9 (452)	This study
MAB (>20 m)	16.0 (530)	—	Li et al. (2022)
Model 2 (>15 m)	18.6 (534)	6.3 (505)	This study
Model 3 (>15 m)	23.5 (534)	7.5 (505)	This study
DIC I, TA IV	33.4 (534)	7.7 (505)	McGarry et al. (2021)

Note. We compare models II and V (DIC and TA, respectively) from McGarry et al. (2021) and the Mid-Atlantic Bight subregion for DIC from Li et al. (2022) to Model 1 in this study, and we compare models I and IV (DIC and TA, respectively) from McGarry et al. (2021) to Models 2 and 3 in this study. The number of samples used to estimate RMSE is given in parentheses.

Other versions of the neural network models could be useful in different contexts. For example, Model 1, which has the lowest error and predicts the carbonate system variables with the highest accuracy, should be used when all predictors are available and input data are of high quality. This model is likely to be increasingly useful in the near future as observational platforms such as the OOI or IOOS programs or other coastal observing sensor arrays become more readily available in this region and produce high quality biogeochemical data outputs. With inputs that reflect both short and longer timescale physical (e.g., temperature and salinity vs. ADT and atmospheric xCO_2, respectively) and biogeochemical processes (e.g., dissolved oxygen vs. KD490, respectively), Model 1 may also be useful for both geospatial and time series analyses, such as long-term trends in surface pCO_2 or air-sea exchange. Such an analysis would also benefit from comparison of the neural network model output to other databases such as the Surface Ocean CO_2 Atlas (SOCAT), which is beyond the scope of this study. Model 3, the

simplest version presented here, which includes only hydrographic variables as predictors along with water depth and atmospheric xCO_2, also has significant potential utility. One use case for this model might be to investigate changes in historical carbonate chemistry, as observational data in this region from the World Ocean Database have reasonably good coverage dating back before the satellite era to the mid-twentieth century. Similar to other studies that empirically model carbonate system variables (e.g., a suggested use case in Carter et al., 2018, 2021), any of the three models may also be useful to generate a second carbonate system predictor variable for use with underway pCO_2 data, or to quantify pH for comparison and quality control of observational platforms.

Data Availability Statement

All data used in this analysis are available in public data repositories, and no login is required for access. The CODAP-NA data set can be found at the following link: https://doi.org/10.5194/essd-13-2777-2021. All data from the OOI Coastal Pioneer Array NES can be found on https://alfresco.oceanobservatories.org/. Samples were used from cruises Pioneer 1–11 (Ocean Observatories Initiative, 2015, 2016, 2019a, 2019b, 2019c, 2019d, 2019e, 2020a, 2020b, 2020c, 2021a, 2021b), references and links to each cruise data set can be found in the reference list in this manuscript. Data from the NES-LTER are available at: https://www.ncei.noaa.gov/archive/accession/0278969 (Wang et al., 2023). All data analysis and model code can be found at https://doi.org/10.5281/zenodo.8018245 (Lima, 2023).

Acknowledgments
This work was supported by NOAA Award Numbers NA20OAR4310448 to JER, IDL, and ZAW and NA20NMF4540028 to ZAW, JER, and JHG.

References

Alin, S. R., Feely, R. A., Dickson, A. G., Hernández-Ayón, J. M., Juranek, L. W., Ohman, M. D., & Goericke, R. (2012). Robust empirical relationships for estimating the carbonate system in the southern California Current System and application to CalCOFI hydrographic cruise data (2005–2011). *Journal of Geophysical Research*, *117*(C5). https://doi.org/10.1029/2011JC007511

Andres, M. (2016). On the recent destabilization of the Gulf Stream path downstream of Cape Hatteras. *Geophysical Research Letters*, *43*(18), 9836–9842. https://doi.org/10.1002/2016GL069966

Bakker, D. C. E., Pfeil, B., Landa, C. S., Metzl, N., O'Brien, K. M., Olsen, A., et al. (2016). A multi-decade record of high-quality fCO_2 data in version 3 of the Surface Ocean CO_2 Atlas (SOCAT). *Earth System Science Data*, *8*(2), 383–413. https://doi.org/10.5194/essd-8-383-2016

Biscaye, P. E., Flagg, C. N., & Falkowski, P. G. (1994). The shelf edge exchange processes experiment, SEEP-II: An introduction to hypotheses, results and conclusions. *Deep Sea Research Part II: Topical Studies in Oceanography*, *41*(2), 231–252. https://doi.org/10.1016/0967-0645(94)90022-1

Bishop, C. M. (1995). *Neural networks for pattern recognition*. Clarendon Press.

Bittig, H. C., Steinhoff, T., Claustre, H., Fiedler, B., Williams, N. L., Sauzède, R., et al. (2018). An alternative to static climatologies: Robust estimation of open ocean CO_2 variables and nutrient concentrations from T, S, and O_2 data using Bayesian neural networks. *Frontiers in Marine Science*, *5*. https://doi.org/10.3389/fmars.2018.00328

Bostock, H. C., Mikaloff Fletcher, S. E., & Williams, M. J. M. (2013). Estimating carbonate parameters from hydrographic data for the intermediate and deep waters of the Southern Hemisphere oceans. *Biogeosciences*, *10*(10), 6199–6213. https://doi.org/10.5194/bg-10-6199-2013

Broullón, D., Pérez, F. F., Velo, A., Hoppema, M., Olsen, A., Takahashi, T., et al. (2019). A global monthly climatology of total alkalinity: A neural network approach. *Earth System Science Data*, *11*(3), 1109–1127. https://doi.org/10.5194/essd-11-1109-2019

Cai, W.-J., Hu, X., Huang, W.-J., Jiang, L.-Q., Wang, Y., Peng, T.-H., & Zhang, X. (2010). Alkalinity distribution in the Western North Atlantic Ocean margins. *Journal of Geophysical Research*, *115*(C8), C08014. https://doi.org/10.1029/2009JC005482

Cai, W.-J., Xu, Y.-Y., Feely, R. A., Wanninkhof, R., Jönsson, B., Alin, S. R., et al. (2020). Controls on surface water carbonate chemistry along North American ocean margins. *Nature Communications*, *11*(1), 2691. https://doi.org/10.1038/s41467-020-16530-z

Carter, B. R., Bittig, H. C., Fassbender, A. J., Sharp, J. D., Takeshita, Y., Xu, Y.-Y., et al. (2021). New and updated global empirical seawater property estimation routines. *Limnology and Oceanography: Methods*, *19*(12), 785–809. https://doi.org/10.1002/lom3.10461

Carter, B. R., Feely, R. A., Williams, N. L., Dickson, A. G., Fong, M. B., & Takeshita, Y. (2018). Updated methods for global locally interpolated estimation of alkalinity, pH, and nitrate. *Limnology and Oceanography: Methods*, *16*(2), 119–131. https://doi.org/10.1002/lom3.10232

Chen, K., Gawarkiewicz, G. G., Lentz, S. J., & Bane, J. M. (2014). Diagnosing the warming of the Northeastern U.S. Coastal Ocean in 2012: A linkage between the atmospheric jet stream variability and ocean response. *Journal of Geophysical Research: Oceans*, *119*(1), 218–227. https://doi.org/10.1002/2013JC009393

Chen, S., Hu, C., Barnes, B. B., Wanninkhof, R., Cai, W.-J., Barbero, L., & Pierrot, D. (2019). A machine learning approach to estimate surface ocean pCO_2 from satellite measurements. *Remote Sensing of Environment*, *228*, 203–226. https://doi.org/10.1016/j.rse.2019.04.019

Davis, C. V., Hewett, K., Hill, T. M., Largier, J. L., Gaylord, B., & Jahncke, J. (2018). Reconstructing aragonite saturation state based on an empirical relationship for Northern California. *Estuaries and Coasts*, *41*(7), 2056–2069. https://doi.org/10.1007/s12237-018-0372-0

Dickson, A. G. (1990). Standard potential of the reaction: AgCl(s) + 12H2(g) = Ag(s) + HCl(aq), and and the standard acidity constant of the ion HSO4- in synthetic sea water from 273.15 to 318.15 K. *The Journal of Chemical Thermodynamics*, *22*(2), 113–127. https://doi.org/10.1016/0021-9614(90)90074-Z

Evans, W., Mathis, J. T., Winsor, P., Statscewich, H., & Whitledge, T. E. (2013). A regression modeling approach for studying carbonate system variability in the northern Gulf of Alaska. *Journal of Geophysical Research: Oceans*, *118*(1), 476–489. https://doi.org/10.1029/2012JC008246

Fennel, K., Alin, S., Barbero, L., Evans, W., Bourgeois, T., Cooley, S., et al. (2019). Carbon cycling in the North American Coastal Ocean: A synthesis. *Biogeosciences*, *16*(6), 1281–1304. https://doi.org/10.5194/bg-16-1281-2019

Forsyth, J. S. T., Andres, M., & Gawarkiewicz, G. G. (2015). Recent accelerated warming of the continental shelf off New Jersey: Observations from the CMV Oleander expendable bathythermograph line. *Journal of Geophysical Research: Oceans*, *120*(3), 2370–2384. https://doi.org/10.1002/2014JC010516

Fourrier, M., Coppola, L., Claustre, H., D'Ortenzio, F., Sauzède, R., & Gattuso, J.-P. (2020). A regional neural network approach to estimate water-column nutrient concentrations and carbonate system variables in the Mediterranean Sea: CANYON-MED. *Frontiers in Marine Science*, *7*. https://doi.org/10.3389/fmars.2020.00620

Gangopadhyay, A., Gawarkiewicz, G., Silva, E. N. S., Monim, M., & Clark, J. (2019). An observed regime shift in the formation of warm core rings from the Gulf Stream. *Scientific Reports*, *9*(1), 12319. https://doi.org/10.1038/s41598-019-48661-9

Gawarkiewicz, G., Chen, K., Forsyth, J., Bahr, F., Mercer, A. M., Ellertson, A., et al. (2019). Characteristics of an advective marine heatwave in the middle Atlantic Bight in early 2017. *Frontiers in Marine Science*, *6*. https://doi.org/10.3389/fmars.2019.00712

Gawarkiewicz, G., Todd, R. E., Zhang, W., Partida, J., Gangopadhyay, A., Monim, M.-U.-H., et al. (2018). The changing nature of shelf-break exchange revealed by the OOI pioneer array. *Oceanography*, *31*(1), 60–70. https://doi.org/10.5670/oceanog.2018.110

Gawarkiewicz, G. G., Todd, R. E., Plueddemann, A. J., Andres, M., & Manning, J. P. (2012). Direct interaction between the Gulf Stream and the shelfbreak south of New England. *Scientific Reports*, *2*(1), 553. https://doi.org/10.1038/srep00553

Gledhill, D. K., White, M. M., Salisbury, J., Thomas, H., Mlsna, I., Liebman, M., et al. (2015). Ocean and Coastal acidification off New England and Nova Scotia. *Oceanography*, *28*(2), 182–197. https://doi.org/10.5670/oceanog.2015.41

Hales, B., Strutton, P. G., Saraceno, M., Letelier, R., Takahashi, T., Feely, R., et al. (2012). Satellite-based prediction of pCO_2 in coastal waters of the eastern North Pacific. *Progress in Oceanography*, *103*, 1–15. https://doi.org/10.1016/j.pocean.2012.03.001

Hornik, K., Stinchcombe, M., & White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks*, *2*(5), 359–366. https://doi.org/10.1016/0893-6080(89)90020-8

Humpheries, M. P., Lewis, E. R., Sharp, J. D., & Pierrot, D. (2021). PyCO2SYS v1.8: Marine carbonate system calculations in Python. *Geoscientific Model Development*, *15*(1), 15–43. https://doi.org/10.5194/gmd-15-15-2022

Ioffe, S., & Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. arXiv:1502.03167[cs.LG] Retrieved from https://arxiv.org/abs/1502.03167

Jiang, L.-Q., Feely, R. A., Wanninkhof, R., Greeley, D., Barbero, L., Alin, S., et al. (2021). Coastal Ocean Data Analysis Product in North America (CODAP-NA) – an internally consistent data product for discrete inorganic carbon, oxygen, and nutrients on the North American ocean margins. *Earth System Science Data*, *13*(6), 2777–2799. 10.5194/essd-13-2777-2021

Juranek, L. W., Feely, R. A., Gilbert, D., Freeland, H., & Miller, L. A. (2011). Real-time estimation of pH and aragonite saturation state from Argo profiling floats: Prospects for an autonomous carbon observing strategy. *Geophysical Research Letters*, *38*(17). https://doi.org/10.1029/2011GL048580

Juranek, L. W., Feely, R. A., Peterson, W. T., Alin, S. R., Hales, B., Lee, K., et al. (2009). A novel method for determination of aragonite saturation state on the continental shelf of central Oregon using multi-parameter relationships with hydrographic data. *Geophysical Research Letters*, *36*(24), L24601. https://doi.org/10.1029/2009GL040778

Kim, T.-W., Park, G.-H., Kim, D., Lee, K., Feely, R. A., & Millero, F. J. (2015). Seasonal variations in the aragonite saturation state in the upper open-ocean waters of the North Pacific Ocean. *Geophysical Research Letters*, *42*(11), 4498–4506. https://doi.org/10.1002/2015GL063602

Landschützer, P., Gruber, N., Bakker, D. C. E., Schuster, U., Nakaoka, S., Payne, M. R., et al. (2013). A neural network-based estimate of the seasonal to inter-annual variability of the Atlantic Ocean carbon sink. *Biogeosciences*, *10*(11), 7793–7815. https://doi.org/10.5194/bg-10-7793-2013

Laruelle, G. G., Landschützer, P., Gruber, N., Tison, J.-L., Delille, B., & Regnier, P. (2017). Global high-resolution monthly pCO_2 climatology for the coastal ocean derived from neural network interpolation. *Biogeosciences*, *14*(19), 4545–4561. https://doi.org/10.5194/bg-14-4545-2017

Lee, K., Kim, T.-W., Byrne, R. H., Millero, F. J., Feely, R. A., & Liu, Y.-M. (2010). The universal ratio of boron to chlorinity for the North Pacific and North Atlantic oceans. *Geochimica et Cosmochimica Acta*, *74*(6), 1801–1811. https://doi.org/10.1016/j.gca.2009.12.027

Lee, K., Tong, L. T., Millero, F. J., Sabine, C. L., Dickson, A. G., Goyet, C., et al. (2006). Global relationships of total alkalinity with salinity and temperature in surface waters of the world's oceans. *Geophysical Research Letters*, *33*(19), L19605. https://doi.org/10.1029/2006GL027207

Lentz, S. J. (2008). Observations and a model of the mean circulation over the middle Atlantic Bight continental shelf. *Journal of Physical Oceanography*, *38*(6), 1203–1221. https://doi.org/10.1175/2007JPO3768.1

Lueker, T. J., Dickson, A. G., & Keeling, C. D. (2000). Ocean pCO_2 calculated from dissolved inorganic carbon, alkalinity, and equations for K1 and K2: Validation based on laboratory measurements of CO_2 in gas and seawater at equilibrium. *Marine Chemistry*, *70*(1), 105–119. https://doi.org/10.1016/S0304-4203(00)00022-0

Li, X., Bellerby, R. G. J., Ge, J., Wallhead, P., Liu, J., & Yang, A. (2020). Retrieving monthly and interannual total-scale pH (pH *T*) on the East China Sea shelf using an artificial neural network: ANN-pH *T*-v1. *Geoscientific Model Development*, *13*(10), 5103–5117. https://doi.org/10.5194/gmd-13-5103-2020

Li, X., Bellerby, R. G. J., Wallhead, P., Ge, J., Liu, J., Liu, J., & Yang, A. (2020). A Neural Network-Based Analysis of the Seasonal Variability of Surface Total Alkalinity on the East China Sea Shelf. *Frontiers in Marine Science*, *7*. https://doi.org/10.3389/fmars.2020.00219

Li, X., Xu, Y.-Y., Kirchman, D. L., & Cai, W.-J. (2022). Carbonate Parameter Estimation and Its Application in Revealing Temporal and Spatial Variation in the South and Mid-Atlantic Bight, USA. *Journal of Geophysical Research: Oceans*, *127*(7), e2022JC018811. https://doi.org/10.1029/2022JC018811

Lima, I. D. (2023). Code release for AGU Biogeosciences publication. [Software]. Zenodo. https://doi.org/10.5281/zenodo.8018245

Lohrenz, S. E., Cai, W. J., Chakraborty, S., Huang, W. J., Guo, X., He, R., et al. (2018). Satellite estimation of coastal pCO_2 and air-sea flux of carbon dioxide in the northern Gulf of Mexico. *Remote Sensing of Environment*, *207*, 71–83. https://doi.org/10.1016/j.rse.2017.12.039

Martin, M., Dash, P., Ignatov, A., Banzon, V., Beggs, H., Brasnett, B., et al. (2012). Group for high resolution sea surface temperature (GHRSST) analysis fields inter-comparisons. Part 1: A GHRSST multi-product ensemble (GMPE). *Deep Sea Research Part II: Topical Studies in Oceanography*, *77–80*(77–80), 21–30. https://doi.org/10.1016/j.dsr2.2012.04.013s

Marzban, C. (2009). Basic statistics and basic AI: Neural networks. In S. E. Haupt, A. Pasini, & C. Marzban (Eds.), *Artificial intelligence methods in the environmental sciences* (pp. 15–47). Springer. https://doi.org/10.1007/978-1-4020-9119-3_2

McGarry, K., Siedlecki, S. A., Salisbury, J., & Alin, S. R. (2021). Multiple linear regression models for reconstructing and exploring processes controlling the carbonate system of the Northeast US from basic hydrographic data. *Journal of Geophysical Research: Oceans*, *126*(2), e2020JC016480. https://doi.org/10.1029/2020JC016480

McNeil, B. I., & Sasse, T. P. (2016). Future ocean hypercapnia driven by anthropogenic amplification of the natural CO_2 cycle. *Nature*, *529*(7586), 383–386. https://doi.org/10.1038/nature16156

Millero, F. J., Lee, K., & Roche, M. (1998). Distribution of alkalinity in the surface waters of the major oceans. *Marine Chemistry*, *60*(1), 111–130. https://doi.org/10.1016/S0304-4203(97)00084-4

Mills, K. E., Pershing, A. J., Brown, C. J., Chen, Y., Chiang, F.-S., Holland, D. S., et al. (2013). Fisheries management in a changing climate: Lessons from the 2012 ocean heat wave in the Northwest Atlantic. *Oceanography*, *26*(2), 191–195. https://doi.org/10.5670/oceanog.2013.27

Monim, M. (2017). *Seasonal and inter-annual variability of Gulf Stream warm core rings from 2000 to 2016*. (p. 113). University of Massachusetts-Dartmouth.

Najjar, R. G., Herrmann, M., Alexander, R., Boyer, E. W., Burdige, D. J., Butman, D., et al. (2018). Carbon budget of tidal wetlands, estuaries, and shelf waters of Eastern North America. *Global Biogeochemical Cycles*, *32*(3), 389–416. https://doi.org/10.1002/2017GB005790

Ocean Observatories Initiative. (2015). Pioneer 2 carbonate data. [Dataset]. Alfresco OOI Data Portal. Retrieved from https://alfresco.oceanobservatories.org/alfresco/d/d/workspace/SpacesStore/16dca970-6654-4fd4-aa8d-4c0861cbc0a2/Pioneer-02_KN217_DIC_Sample_Data_2015-10-29_ver_1-00.xlsx

Ocean Observatories Initiative. (2016). Pioneer 1 carbonate data. [Dataset]. Alfresco OOI Data Portal. Retrieved from https://alfresco.oceanobservatories.org/alfresco/d/d/workspace/SpacesStore/60f28643-e5d9-4a72-8d8c-34fe7d3c539f/Pioneer-01_KN214_DIC_Sample_Data_2016-06-19_ver_1-01.xlsx

Ocean Observatories Initiative. (2019a). Pioneer 3 carbonate data – Leg 1. [Dataset]. Alfresco OOI Data Portal. Retrieved from https://alfresco.oceanobservatories.org/alfresco/d/d/workspace/SpacesStore/bdc8c958-c7cd-43a8-a740-a218f337a1c1/Pioneer-03_Leg-1_KN222_DIC_Sample_Data_2019-06-19_Ver_1-00.xlsx

Ocean Observatories Initiative. (2019b). Pioneer 3 carbonate data – Leg 2. [Dataset]. Alfresco OOI Data Portal. Retrieved from https://alfresco.oceanobservatories.org/alfresco/d/d/workspace/SpacesStore/e2855ec5-a15f-423d-a8f7-bb02386e84f9/Pioneer-03_Leg-2_KN224_DIC_Sample_Data_2019-06-19_Ver_1-00.xlsx

Ocean Observatories Initiative. (2019c). Pioneer 4 carbonate data. [Dataset]. Alfresco OOI Data Portal. Retrieved from https://alfresco.oceanobservatories.org/alfresco/d/d/workspace/SpacesStore/8fcf4b98-2c28-43d1-be18-714a7f5f88a7/Pioneer-04_AT27_DIC_Sample_Data_2019-06-19_ver_1-00.xlsx

Ocean Observatories Initiative. (2019d). Pioneer 5 carbonate data. [Dataset]. Alfresco OOI Data Portal. Retrieved from https://alfresco.oceanobservatories.org/alfresco/d/d/workspace/SpacesStore/92e7f1c6-5c97-47e7-a312-5414c1fd861e/Pioneer-05_AT31_DIC_Sample_Data_2019-06-19_ver_1-00.xlsx

Ocean Observatories Initiative. (2019e). Pioneer 8 carbonate data. [Dataset]. Alfresco OOI Data Portal. Retrieved from https://alfresco.oceanobservatories.org/alfresco/d/d/workspace/SpacesStore/d4b3fd30-ba0e-47f0-8800-44aa4f1c4123/Pioneer-08_AR18_DIC_Sample_Data_2019-11-06_ver_1-00.xlsx

Ocean Observatories Initiative. (2020a). Pioneer 6 carbonate data. [Dataset]. Alfresco OOI Data Portal. Retrieved from https://alfresco.oceanobservatories.org/alfresco/d/d/workspace/SpacesStore/a25d686c-d258-45e2-9380-88f0c14788f8/Pioneer-06_AR4_DIC_Sample_Data_2020-12-03_ver_1-02.xlsx

Ocean Observatories Initiative. (2020b). Pioneer 7 carbonate data. [Dataset]. Alfresco OOI Data Portal. Retrieved from https://alfresco.oceanobservatories.org/alfresco/d/d/workspace/SpacesStore/5857ea0d-afaf-49d0-a9c0-99342d6ac084/Pioneer-07_AR8_DIC_Sample_Data_2020-03-27_ver_1-00.xlsx

Ocean Observatories Initiative. (2020c). Pioneer 9 carbonate data. [Dataset]. Alfresco OOI Data Portal. Retrieved from https://alfresco.oceanobservatories.org/alfresco/d/d/workspace/SpacesStore/8479660b-87d7-4850-a4be-57ed993609e5/Pioneer-09_AR24_DIC_Sample_Data_2020-05-14_ver_1-00.xlsx

Ocean Observatories Initiative. (2021a). Pioneer 10 carbonate data. [Dataset]. Alfresco OOI Data Portal. Retrieved from https://alfresco.oceanobservatories.org/alfresco/d/d/workspace/SpacesStore/a419a7e5-ab4a-49a8-9651-b98200ea3fcb/Pioneer-10_AR28_DIC_Sample_Data_2021-01-03_Ver_1-00.xlsx

Ocean Observatories Initiative. (2021b). Pioneer 11 carbonate data. [Dataset]. Alfresco OOI Data Portal. Retrieved from https://alfresco.oceanobservatories.org/alfresco/d/d/workspace/SpacesStore/c0460fa7-75c9-44b8-ac05-73c5de27c68a/Pioneer-11_AR31_DIC_Sample_Data_2021-02-09_Ver_1-00.xlsx

Perez, E., Ryan, S., Andres, M., Gawarkiewicz, G., Ummenhofer, C. C., Bane, J., & Haines, S. (2021). Understanding physical drivers of the 2015/16 marine heatwaves in the Northwest Atlantic. *Scientific Reports*, *11*(1), 17623. https://doi.org/10.1038/s41598-021-97012-0

Perez, F. F., & Fraga, F. (1987). Association constant of fluoride and hydrogen ions in seawater. *Marine Chemistry*, *21*(2), 161–168. https://doi.org/10.1016/0304-4203(87)90036-3

Pershing, A. J., Alexander, M. A., Hernandez, C. M., Kerr, L. A., Le Bris, A., Mills, K. E., et al. (2015). Slow adaptation in the face of rapid warming leads to collapse of the Gulf of Maine cod fishery. *Science*, *350*(6262), 809–812. https://doi.org/10.1126/science.aac9819

Piecuch, C. G., Bittermann, K., Kemp, A. C., Ponte, R. M., Little, C. M., Engelhart, S. E., & Lentz, S. J. (2018). River-discharge effects on United States Atlantic and Gulf coast sea-level changes. *Proceedings of the National Academy of Sciences*, *115*(30), 7729–7734. https://doi.org/10.1073/pnas.1805428115

Pousse, É., Munroe, D., Hart, D., Hennen, D., Cameron, L. P., Rheuban, J. E., et al. (2022). Dynamic energy budget modeling of Atlantic surfclam, Spisula solidissima, under future ocean acidification and warming. *Marine Environmental Research*, *177*, 105602. https://doi.org/10.1016/j.marenvres.2022.105602

Pujol, M.-I., & Taburet, G., & SL-TAC Team. (2023). *Quality information document: Sea level TAC - DUACS products*. Copernicus Marine Service. Retrieved from https://catalogue.marine.copernicus.eu/documents/QUID/CMEMS-SL-QUID-008-032-068.pdf

Rheuban, J. E., Doney, S. C., McCorkle, D. C., & Jakuba, R. W. (2019). Quantifying the effects of nutrient enrichment and freshwater mixing on Coastal Ocean acidification. *Journal of Geophysical Research: Oceans*, *124*(12), 9085–9100. https://doi.org/10.1029/2019JC015556

Salisbury, J., Green, M., Hunt, C., & Campbell, J. (2008). Coastal acidification by rivers: A threat to shellfish? *Eos, Transactions American Geophysical Union*, *89*(50), 513. https://doi.org/10.1029/2008EO500001

Salisbury, J. E., & Jönsson, B. F. (2018). Rapid warming and salinity changes in the Gulf of Maine alter surface ocean carbonate parameters and hide ocean acidification. *Biogeochemistry*, *141*(3), 401–418. https://doi.org/10.1007/s10533-018-0505-3

Sallenger, A. H., Doran, K. S., & Howd, P. A. (2012). Hotspot of accelerated sea-level rise on the Atlantic coast of North America. *Nature Climate Change*, *2*(12), 884–888. https://doi.org/10.1038/nclimate1597

Sasse, T. P., McNeil, B. I., & Abramowitz, G. (2013). A novel method for diagnosing seasonal to inter-annual surface ocean carbon dynamics from bottle data using neural networks. *Biogeosciences*, *10*(6), 4319–4340. https://doi.org/10.5194/bg-10-4319-2013

Sauzède, R., Bittig, H. C., Claustre, H., Pasqueron de Fommervault, O., Gattuso, J.-P., Legendre, L., & Johnson, K. S. (2017). Estimates of water-column nutrient concentrations and carbonate system parameters in the global ocean: A novel approach based on neural networks. *Frontiers in Marine Science*, *4*. https://doi.org/10.3389/fmars.2017.00128

Schlegel, R. W., Oliver, E. C. J., & Chen, K. (2021). Drivers of marine heatwaves in the Northwest Atlantic: The role of air–sea interaction during onset and decline. *Frontiers in Marine Science*, *8*. https://doi.org/10.3389/fmars.2021.627970

Seegers, B. N., Stumpf, R. P., Schaeffer, B. A., Loftin, K. A., & Werdell, P. J. (2018). Performance metrics for the assessment of satellite data products: An ocean color case study. *Optics Express*, *26*(6), 7404–7422. https://doi.org/10.1364/OE.26.007404

Siedlecki, S., Salisbury, J., Gledhill, D., Bastidas, C., Meseck, S., McGarry, K., et al. (2021). Projecting ocean acidification impacts for the Gulf of Maine to 2050: New tools and expectations. *Elementa: Science of the Anthropocene*, *9*(1), 00062. https://doi.org/10.1525/elementa.2020.00062

Signorini, S. R., Mannino, A., Najjar, R. G., Friedrichs, M. A. M., Cai, W.-J., Salisbury, J., et al. (2013). Surface ocean pCO_2 seasonality and sea-air CO_2 flux estimates for the North American East Coast. *Journal of Geophysical Research: Oceans*, *118*(10), 5439–5460. https://doi.org/10.1002/jgrc.20369

Tans, & Keeling. (2021). Mauna Loa CO_2 Annual Mean Data [Dataset]. Retrieved from https://gml.noaa.gov/ccgg/trends/data.html

Townsend, D. W., Rebuck, N. D., Thomas, M. A., Karp-Boss, L., & Gettings, R. M. (2010). A changing nutrient regime in the Gulf of Maine. *Continental Shelf Research*, *30*(7), 820–832. https://doi.org/10.1016/j.csr.2010.01.019

Turk, D., Dowd, M., Lauvset, S. K., Koelling, J., Alonso-Pérez, F., & Pérez, F. F. (2017). Can empirical algorithms successfully estimate aragonite saturation state in the Subpolar North Atlantic? *Frontiers in Marine Science*, *4*. https://doi.org/10.3389/fmars.2017.00385

Vance, J. M., Currie, K., Zeldis, J., Dillingham, P. W., & Law, C. S. (2022). An empirical MLR for estimating surface layer DIC and a comparative assessment to other gap-filling techniques for ocean carbon time series. *Biogeosciences*, *19*(1), 241–269. https://doi.org/10.5194/bg-19-241-2022

Velo, A., Pérez, F. F., Tanhua, T., Gilcoto, M., Ríos, A. F., & Key, R. M. (2013). Total alkalinity estimation using MLR and neural network techniques. *Journal of Marine Systems*, *111–112*, 11–18. https://doi.org/10.1016/j.jmarsys.2012.09.002

Vlahos, P., Chen, R. F., & Repeta, D. J. (2002). Dissolved organic carbon in the Mid-Atlantic Bight. *Deep Sea Research Part II: Topical Studies in Oceanography*, *49*(20), 4369–4385. https://doi.org/10.1016/S0967-0645(02)00167-4

Wallace, R. B., Baumann, H., Grear, J. S., Aller, R. C., & Gobler, C. J. (2014). Coastal ocean acidification: The other eutrophication problem. *Estuarine, Coastal and Shelf Science*, *148*, 1–13. https://doi.org/10.1016/j.ecss.2014.05.027

Wang, Z. A., Lawson, G. L., Pilskaln, C. H., & Maas, A. E. (2017). Seasonal controls of aragonite saturation states in the Gulf of Maine. *Journal of Geophysical Research: Oceans*, *122*(1), 372–389. https://doi.org/10.1002/2016JC012373

Wang, Z. A., Ni, X., Rheuban, J. E., & Morkeski, K. (2023). Carbonate chemistry data from select cruises at the Northeast US Coast Shelf long term ecological research site from 2018-07-20 to 2019-02-04 (NCEI accession 0278969). [Dataset]. NOAA National Centers for Environmental Information. Unpublished Dataset. Retrieved from https://www.ncei.noaa.gov/archive/accession/0278969

Wang, Z. A., Wanninkhof, R., Cai, W.-J., Byrne, R. H., Hu, X., Peng, T.-H., & Huang, W.-J. (2013). The marine inorganic carbon system along the Gulf of Mexico and Atlantic Coasts of the United States: Insights from a transregional coastal carbon study. *Limnology & Oceanography*, *58*(1), 325–342. https://doi.org/10.4319/lo.2013.58.1.0325

Wanninkhof, R., Barbero, L., Byrne, R., Cai, W.-J., Huang, W.-J., Zhang, J.-Z., et al. (2015). Ocean acidification along the Gulf Coast and East Coast of the USA. *Continental Shelf Research*, *98*, 54–71. https://doi.org/10.1016/j.csr.2015.02.008

Williams, N. L., Juranek, L. W., Johnson, K. S., Feely, R. A., Riser, S. C., Talley, L. D., et al. (2016). Empirical algorithms to estimate water column pH in the Southern Ocean. *Geophysical Research Letters*, *43*(7), 3415–3422. https://doi.org/10.1002/2016GL068539

Wright-Fairbanks, E. K., Miles, T. N., Cai, W., Chen, B., & Saba, G. K. (2020). Autonomous observation of seasonal carbonate chemistry dynamics in the Mid-Atlantic Bight. *Journal of Geophysical Research: Oceans*, *125*(11), e2020JC016505. https://doi.org/10.1029/2020JC016505

Xu, Y.-Y., Cai, W.-J., Wanninkhof, R., Salisbury, J., Reimer, J., & Chen, B. (2020). Long-term changes of carbonate chemistry variables along the North American East Coast. *Journal of Geophysical Research: Oceans*, *125*(7), e2019JC015982. https://doi.org/10.1029/2019JC015982

Zeng, J., Nojiri, Y., Landschützer, P., Telszewski, M., & Nakaoka, S. (2014). A global surface ocean fCO_2 climatology based on a feed-forward neural network. *Journal of Atmospheric and Oceanic Technology*, *31*(8), 1838–1849. https://doi.org/10.1175/JTECH-D-13-00137.1

Zhang, L., Wu, L., & Zhang, J. (2011). Simulated response to recent freshwater flux change over the Gulf Stream and its extension: Coupled ocean–atmosphere adjustment and Atlantic–Pacific teleconnection. *Journal of Climate*, *24*(15), 3971–3988. https://doi.org/10.1175/2011JCLI4020.1

Zhao, C. L., & Tans, P. P. (2006). Estimating uncertainty of the WMO mole fraction scale for carbon dioxide in air. *Journal of Geophysical Research*, *111*(D8), D08S09. https://doi.org/10.1029/2005JD006003