Real-time reconstruction of 3D ocean temperature fields from reanalysis data and satellite and buoy surface measurements

Bianca Champenois[1], **Themistoklis P. Sapsis**[1]

[1]Massachusetts Institute of Technology,
Department of Mechanical Engineering,
77 Massachusetts Av., Cambridge, MA 02139

Key Points:

- We present a framework to model 3D ocean temperature fields and their uncertainty from real-time surface temperature sensor measurements.
- Our approach uses a convolutional neural network to capture structure from physics-based numerical models.
- The framework is validated with in-situ measurements of ocean temperature at various depths.

Corresponding author: Bianca Champenois, bchamp@mit.edu

Abstract

Despite advancements in computational science, nonlinear geophysical processes still present important modeling challenges. Physical sensors (e.g. satellites, AUVs, and buoys) can collect data at specific points, but are often sparse or inaccurate. We present a framework to build improved spatiotemporal models that combine dynamics inferred from high-fidelity models and sensor measurements. Specifically, we consider ocean temperature because it is an indicator for ocean acidification, and we are motivated by a data set of sensor measurements only available at the ocean's surface. We first apply standard principal component analysis (PCA) at every ocean surface coordinate to a reanalysis data set of the time-evolving 3D temperature field. Next, a conditionally Gaussian model implemented through a temporal convolutional neural network (TCN) is built to predict the time coefficients of the PCA modes, and their variance, as a function of surface temperature. The 2D surface temperature field is estimated by a multi-fidelity Gaussian process regression scheme, for which buoy data have highest fidelity and satellite data have lower fidelity. The surface temperature is then used as input to the neural network to probabilistically predict the PCA coefficients and reconstruct the full 3D temperature field. The results are compared to in-situ measurements at all depths, and the model can be leveraged for optimal sampling and path planning. Overall, the proposed framework can build less expensive and more accurate conditionally Gaussian models in real time, and it can be applied to other geophysical systems for which data from sensors and numerical models are available.

Plain Language Summary

Nonlinear geophysical systems are challenging to model. Sensor measurements are sparse and physics-based numerical simulations are expensive. Here we present a method to build models of geophysical systems that combine physics-based numerical simulations with real-time sensor measurements. We apply this method to a model of ocean temperature for which sensor measurements are only available at the surface of the ocean. We use data science techniques to extract the vertical structure of the temperature field from the physics-based simulation. Then, we train a neural network on the data from the numerical simulation to predict the temperature over depth as a function of surface temperature. Finally, we reconstruct the full 3D temperature field given real-time satellite and buoy measurements, and we compare the predictions with measurements at all depths. Our model also provides an estimate for the variance associated with the physical system. We discuss how our model can be used for active sampling and path planning.

1 Introduction

Environmental and geophysical systems can be modeled with nonlinear equations that typically require complex and computationally expensive numerical solvers. Even with highly accurate numerical methods, model errors still exist, and there can be losses of predictability due to intrinsic instabilities in the system. Such challenges can be mitigated with physical sensors which can be used to collect additional information on quantities of interest. However, sensors only provide information about the system locally in space (e.g. buoys or drifters) or with a high degree of sparsity (e.g. satellite data). There is a need for improved data assimilation techniques that can estimate the state and uncertainty of a system in real time.

To provide a specific example, we consider the temperature of the Massachusetts and Cape Cod Bays, an area with great biodiversity (fish, shellfish, whales, etc.) and significant fishing and tourism industries. The ability to predict temperature is helpful in tracking ocean health and ocean acidification properties (Juranek et al., 2009). Coastal waters are more susceptible to temperature rise and acidification because the influx of freshwater changes the composition of the ocean and reduces its buffering capacity. Naturally, ocean acidification has greater implications in coastal waters where most fisheries are located (Gledhill et al.,

2015; Wallace et al., 2014; Ekstrom et al., 2015; Friedland et al., 2020). Ocean temperature is governed by a set of high dimensional nonlinear equations. These nonlinear equations can, in principle, be solved to evaluate the temperature field using a numerical scheme, such as finite volumes, but such approaches are typically computationally demanding and need to be repeated for each new set of boundary, initial, and excitation conditions (Lermusiaux et al., 2006). The temperature can also be determined using in-situ buoys and satellites. The buoys provide reliable measurements, but they are scarce. On the other hand, a satellite can cover the whole domain, but there are gaps in the data due to cloud coverage, and the measurements are less accurate. Most importantly, sensor measurements are only available at or near the surface of the ocean, leaving the bottom depths of the ocean unaccounted for (Klemas & Yan, 2014; B. Li et al., 2017).

The goal of this work is to leverage data science techniques to characterize the vertical structure of the ocean temperature field from physics-based numerical simulations, and subsequently combine machine-learning methods with real-time sensor measurements of surface temperature to reconstruct and hindcast the full 3D temperature field and its uncertainties. We consider temperature data because it is readily available, and it is a useful indicator for ocean acidification, but the techniques discussed can be applied to other quantities of interest such as salinity, dissolved inorganic carbon (DIC), aragonite, and pH. We develop a framework that can estimate the full 3D field of ocean temperature from real-time satellite and buoy measurements of surface temperature. In addition, the model can estimate the uncertainty associated with both the system and the model. Specifically, we use a combination of data science techniques including principal component analysis (PCA), temporal convolutional neural networks (TCN) and Gaussian process regression (GPR). As a result, we develop a computationally inexpensive model for the Massachusetts and Cape Cod Bays that leverages data from physics-based numerical models, buoys, and satellites to predict the temperature and uncertainty in real time at all points in the domain of interest. The model is also useful to make decisions about where and how to sample future data (Sapsis, 2020; X. Yang et al., 2010) and to evaluate the quality of new sensors.

The framework is organized into multiple steps as outlined in Figure 1. The first two steps are independent. First, in step 1, we use multi-fidelity Gaussian process regression (GPR) to estimate the ocean surface temperature by merging information from satellites and in-situ buoys (Babaee et al., 2020). Next, in step 2, we use the reanalysis data to build a data-driven reduced order model and derive a functional relationship between 3D temperature and surface temperature; this connection is possible given the reduced-order vertical structure of the problem that we obtain from principal component analysis (PCA). Finally, in step 3, we input the real-time 2D surface temperature measurements into the reduced-order model to obtain a real-time estimate for the 3D temperature field and its uncertainty. The framework can be modified or rearranged based on the type and location of new data that become available.

2 Reanalysis Data and Measurements

Our starting point is reanalysis data consisting of a time-evolving 3D temperature field of the Northeast Coastal Ocean from the FVCOM (Finite Volume Community Ocean Model) simulation from Chen et al. (2003). Reanalysis data refer to data from a numerical simulation that integrates real world measurements into the computation. The model uses a fractional step method to solve the spatially and temporally evolving fields for velocity, density, temperature, and salinity, among other variables. A study by B. Li et al. (2017) found that the model agreed well with in-situ measurements with a root mean squared error of 2.28 °C. We consider a truncated portion of the domain in the Massachusetts and Cape Cod Bays that consists of 45 sigma levels from January 2005 to December 2013 (9 years total). Here, a sigma level refers to a layer of the sigma coordinate system. In the sigma coordinate system, horizontal layers follow the model terrain, so for a given (x, y) point, each horizontal layer has the same thickness (Mellor et al., 2002). This coordinate system

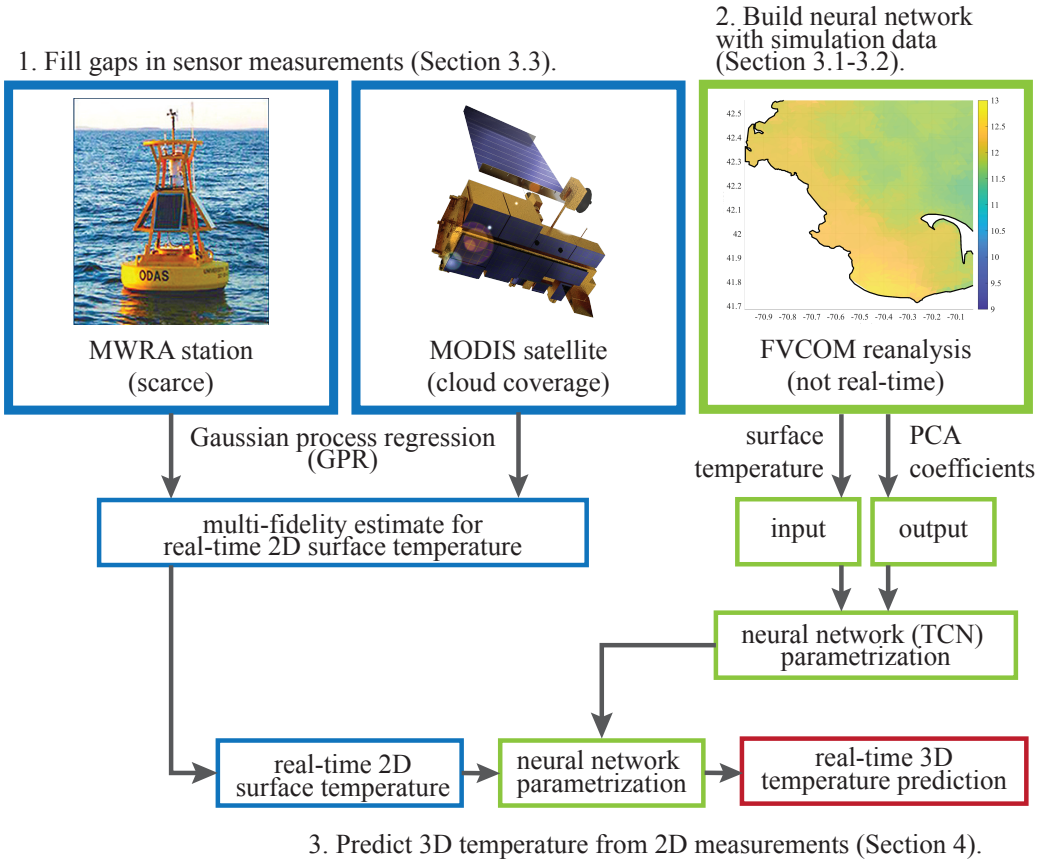

Figure 1. Framework. Flow chart describing the developed framework for real-time estimation of the 3D ocean temperature field. Reanalysis data are employed to estimate a reduced-order model. Ocean surface information, obtained from satellite and buoy measurements, are used as input.

is a convenient way to discretize the domain because it results in a continuous temperature field. As an example, a snapshot of the data from September 13th, 2012 at sigma level -0.5 is plotted in Figure 2.

In addition to the data from the finite volume scheme, we have surface temperature data from physical sensors: satellites, in-situ stations, and buoys. Satellites measure sea surface temperature by quantifying the energy of wavelengths coming from the ocean. Different satellites operate at varying resolutions and levels of accuracy (Chao et al., 2009), but the main challenge associated with using satellite data is that there are gaps due to cloud coverage. In this project, we have access to daily satellite imagery from the MODerate-resolution Imaging Spectroradiometer (MODIS) Terra. In Figure 3 we observe that each day has a different amount of cloud coverage. Most importantly, many days during winter months have no available satellite measurements. In contrast to satellites, in-situ stations and buoys are not affected by cloud coverage. Measurements are available from the Massachusetts Water Resources Authority (MWRA) (Figure 3), but they are only collected on a monthly basis, and there are only 14 locations. The MWRA stations gather data by collecting samples of water at multiple depths and directly measuring the temperature. While this method is more accurate, it is also very costly. Finally, we also have National Oceanic and Atmospheric Administration (NOAA) measurements from environmental monitors on lobster traps and

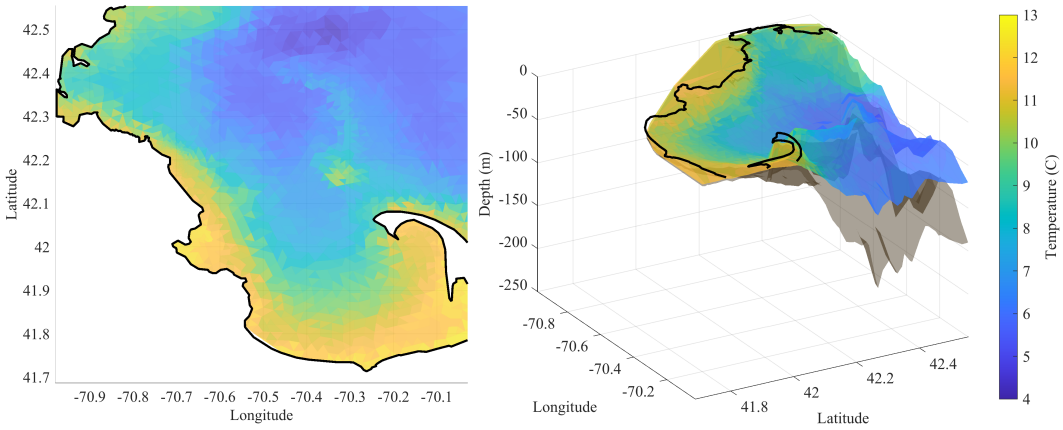

Figure 2. Reanalysis Data of the Northeast Coastal Ocean from FVCOM. The temperature field is plotted for September 13th, 2012 at sigma level -0.5.

large trawlers (eMOLT) and the Northeast Fisheries Science Center (NEFSC). Unlike for the MWRA measurements, there is less information about the accuracy and maintenance of the NOAA sensors. Consequently, we only use surface data from the MWRA stations to build our model, but the model can be augmented with any number of available sources of data. At the validation stage we employ in-depth measurements from the aforementioned sensors to assess the quality of our model.

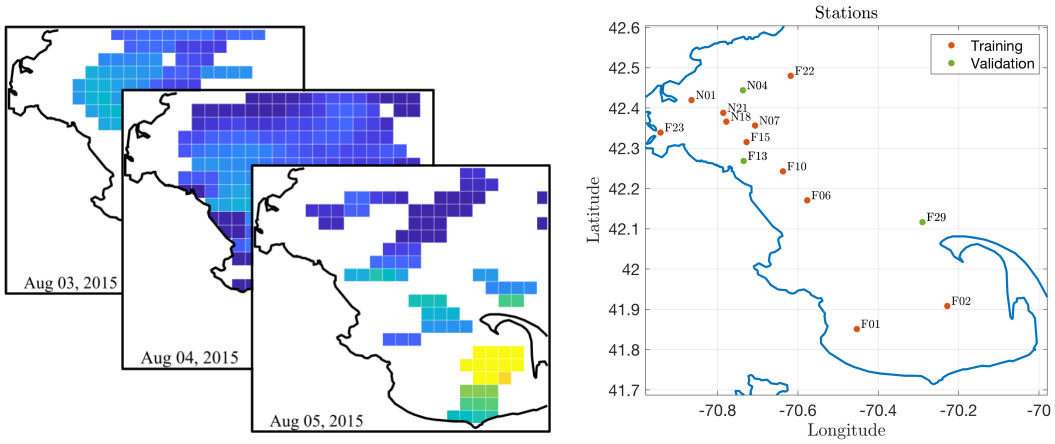

Figure 3. Sensor Data. The low fidelity data (satellite) is only available on days with low cloud coverage. The high fidelity data (buoys) is local in space and sparse.

3 Framework Description

3.1 Temperature Field Order-Reduction Using Vertical PCA

We first apply standard principal component analysis to the reanalysis data set to reduce the dimensionality while retaining patterns and information. Principal component analysis (PCA), also known as proper orthogonal decomposition (POD) or Karhunen–Loève decomposition, among other names, has long been used in many fields. In the context of

fluid mechanics, weather prediction (Lorenz, n.d.; Hannachi et al., 2007), and oceanography, PCA extracts features or trends from large empirical data sets to accurately reconstruct the dynamics of the system using a small number of empirical orthogonal functions (EOF) and corresponding coefficients. Significant work has been done on the use of empirical orthogonal functions to reconstruct spatio-temporal sea surface temperature (SST) for which empirical measurements from sensors are available (Everson et al., 1996; Berliner et al., 2000; Ganzedo et al., 2011; Smith et al., 1996). In some cases, the basis is used to fill gappy data (Everson & Sirovich, 1995). Here, we use PCA to extract the vertical structure, or 3D features, of existing reanalysis data. This process allows us to represent the vertical structure of the temperature field with just a few modes at each location of the ocean surface, and it can be proven that PCA results in an optimal orthogonal transformation that captures maximum variance. We are interested in the vertical structure of the temperature field because most of the energy of the system is coming from solar radiative flux which is normal to the surface of the ocean, and the vertical modes capture vertical mixing and diffusion. At each horizontal location i, (x_i, y_i), the temperature field is discretized into n depths and m time steps.

$$\mathbf{T}_i = \begin{bmatrix} T(z_1,t_1) & ... & T(z_1,t_m) \\ T(z_2,t_1) & ... & T(z_2,t_m) \\ & ... & \\ T(z_n,t_1) & ... & T(z_n,t_m) \end{bmatrix} \tag{1}$$

Using this data matrix, we evaluate the eigenvectors

$$\mathbf{T}_i \mathbf{T}_i^T \boldsymbol{\phi}_{ij} = \lambda \boldsymbol{\phi}_{ij}, j = 1, ..., n \tag{2}$$

Finally, for each location i, the in-depth structure of the temperature is represented using 2 vertical modes and a mean temperature mode:

$$\mathbf{T}_{i,proj}(t) = \sum_{j=1}^{2} q_{ij}(t)\boldsymbol{\phi}_{ij} + \bar{\mathbf{T}}_i(t) \tag{3}$$

The eigenvalues obtained from the decomposition confirm that we have a low rank problem as the first two modes capture more than 85% of the data's energy and are sufficient for reconstructing the temperature field (Figure 4b). The spatial modes $\boldsymbol{\phi}_{ij}$ represent the vertical structure of the field and vary with respect to the horizontal location. The first mode roughly corresponds to the thermocline (Figure 4d). The coefficients $q_{ij}(t)$ and mean temperature $\bar{\mathbf{T}}_i(t)$ are functions of time and are extracted from the reanalysis data set via projection. The error between the PCA projection and the original reanalysis field is shown in Figure 5 for different sigma levels. Of course, for the case where there is no full 3D information, a functional relationship between surface information and these coefficients needs to be determined. This is the scope of the next section.

3.2 Machine Learning Functional Relationships Between PCA Coefficients and Surface Temperature

Next, we machine learn a functional relationship between the surface temperature and the temperature over depth at each horizontal location i, (x_i, y_i). We choose surface temperature as the input of the neural network because it is readily accessible from sensor measurements (Figure 3). We also build a second neural network to predict the associated standard deviation and estimate the uncertainty of our predictions. These uncertainties exclusively model the error made by the neural network in modeling the in-depth PCA coefficients.

Recent developments in machine learning have increased the popularity of using neural networks to model geophysical processes. A previous study from Ali et al. (2004) used a neural network to predict subsurface temperature as a function of surface variables. However, they did not perform data preprocessing to reduce the dimensionality of their outputs,

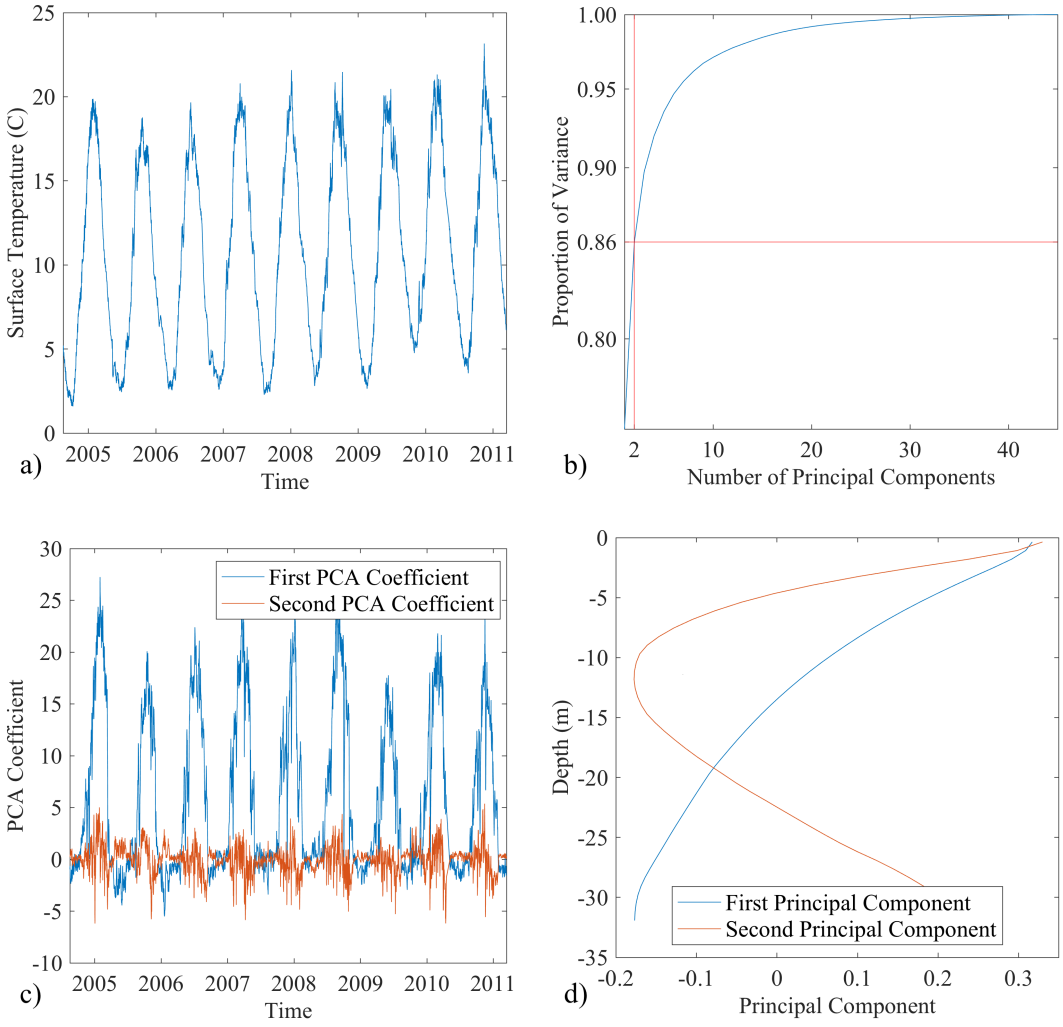

Figure 4. Vertical Order-reduction. a) Time series for the surface temperature at a specific horizontal location, b) energy distribution of the vertical modes, c) time series for the PCA coefficients obtained by projection of the reanalysis temperature field, d) the first two vertical modes.

nor did they predict the variance associated with the processes they were modeling. We specifically build a neural network that predicts the mean and standard deviation of the PCA coefficients $q_{ij}(t)$ and mean temperature $\bar{\mathbf{T}}_i(t)$ obtained in the previous section. Many studies have focused on the use of neural networks to predict such time-varying PCA coefficients (Miao et al., 2019; Maulik et al., 2020; Meng & Karniadakis, 2020; Raissi et al., 2017; W. Yang et al., 2020).

In this project, we build a temporal convolutional network (TCN), a type of convolutional neural network (CNN) that performs convolutions on one dimensional time series data. Unlike a traditional CNN, a TCN is causal which is useful for modeling dynamic systems (Wan et al., 2021). TCNs have also been shown to outperform other recurrent neural networks for sequence modeling (Bai et al., 2018; Aksan & Hilliges, 2019; Lara-Benítez et al., 2020). As such, they are increasingly being used in geophysical applications (Yan et al., 2020; Baño Medina et al., 2020; Gan et al., 2021; Bolton & Zanna, 2019; Weyn et al., 2020). We adapt the Stochastic Machine Learning (SMaL) code from Wan et al. (2021) and retain the same residual block architecture (Figure 6). The data are standardized before

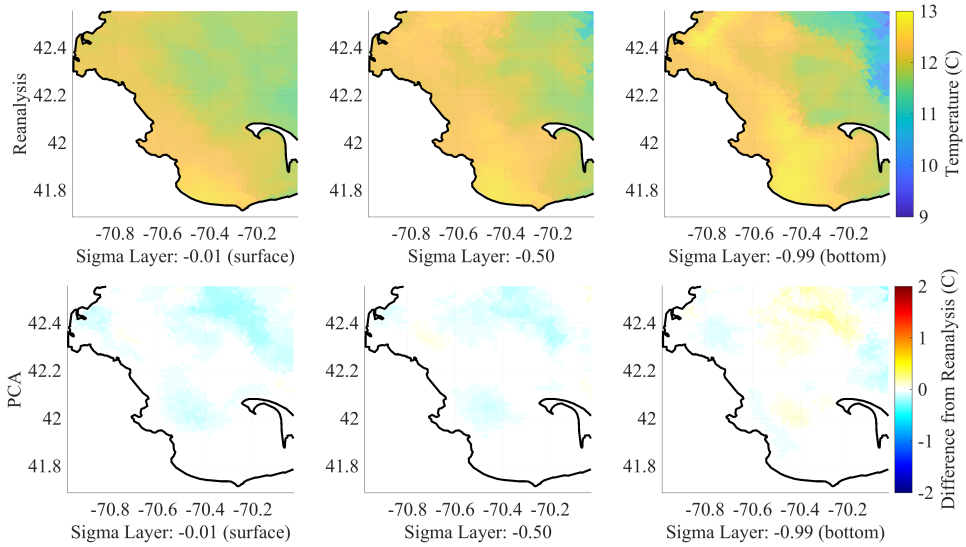

Figure 5. PCA Projection Error. The reanalysis data (top) and the difference between the PCA projection and the reanalysis (bottom) are plotted for March 8th, 2012 at three different sigma layers.

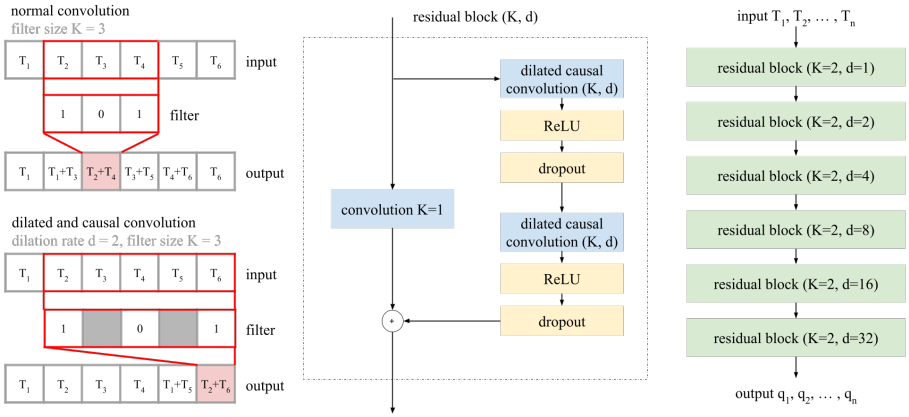

Figure 6. Architecture of TCN. The TCN is built with residual blocks that consist of a sequence of two convolutional layers with ReLU activation and a dropout. The dilation factor of each residual block is doubled at each depth.

training for improved results. The batch size of the neural network is set to 5 for model generalizability. The filter width is set to 2 which results in a small receptive field for reduced computational costs and improved generalizability. The dropout layer has a probability of 0.5 for regularization. The depth of the network is chosen to be 6 layers as this provided an adequate number of degrees of freedom to represent the underlying physical phenomena. The dilation factor is doubled at each depth to cover many different time scales.

187 ### 3.2.1 Loss functions for neural network training

Typically, the weights of a neural network are obtained by minimizing a loss function that quantifies the error between the true data and the model predictions.

$$J(\boldsymbol{\theta}) = \frac{1}{T}\sum L(\hat{\mathbf{y}}(\boldsymbol{\theta}) - \mathbf{y}). \tag{4}$$

Here, we build two neural networks at each location i, (x_i, y_i), one for the mean and one for standard deviation. We emphasize that each horizontal location is treated separately to account for spatial inhomogeneities. We train each network sequentially because we require the mean prediction to train the second neural network for the standard deviation. Furthermore, we optimize different loss functions for each network. To predict the mean of the PCA coefficients, we minimize the mean absolute error (MAE), a standard loss function for neural networks.

$$J_{MAE} = \frac{1}{m}\sum |\hat{y} - y|. \tag{5}$$

To predict the standard deviation of the PCA coefficients, we minimize the mean negative anomaly correlation coefficient (MNACC) (Wan et al., 2021). It is a correlation-based loss function, so it does not scale with magnitude, therefore more effectively penalizing anomalies.

$$J_{MNACC} = \frac{1}{m}\sum \frac{\sum(\hat{z} - [\hat{z}])(z - [z])}{\sqrt{\sum(\hat{z} - [\hat{z}])^2}\sqrt{\sum(z - [z])^2}} \tag{6}$$

$$z = y - y^{ref} \tag{7}$$

188 Here, the reference y^{ref} is the cyclic mean, and for ocean temperature it corresponds to
189 the annual variation due to seasons. Without a reference, this loss reduces to the Pearson
190 correlation coefficient, another standard loss function in many machine learning applications.

191 ### 3.2.2 Choice of number and location of input points

192 While the weights and biases can be found by optimizing a loss function, other param-
193 eters of the neural network need to be fine-tuned through discrete numerical experiments.
194 For example, the choice of input points affects the output of the neural network. We know
195 from the physics of the system that the temperature gradients in the x and y direction also
196 contribute to the vertical temperature profile. As such, we include neighboring points in the
197 input of the neural network to produce a non-local parametrization. We perform numerical
198 experiments to find the number and location of input points that result in the lowest testing
199 error and are best suited for generalizability. We perform these tests on three (x_i, y_i) pairs
200 in the neighborhood domain, denoted A, B, and C in Figure 7, and we adopt the same
parameters for the models of all other (x_i, y_i) pairs. We first test the neural network with

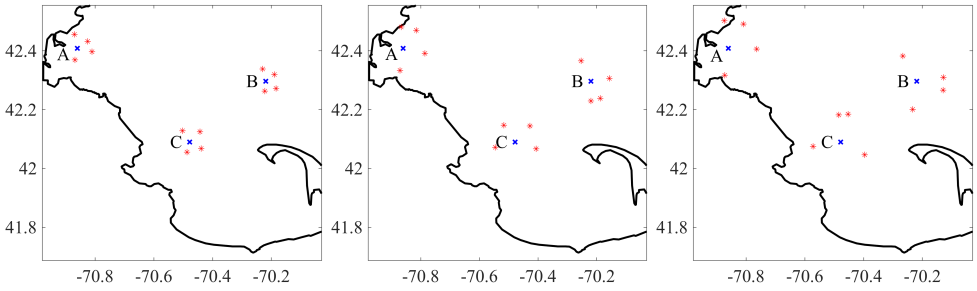

Figure 7. Input Points. The input of the neural network is augmented to include the surface temperature at four nearby points in addition to the surface temperature at the corresponding point of interest. Different neighborhoods, shown here, are tested.

one, two, three, and five neighborhood input points. Then, we experiment with the distance between the input points and the point of interest.

From the results of these numerical experiments, we build the inputs of the TCN with the surface temperature of four additional nearby points for which the distance is between nine and ten kilometers. This distance, obtained from numerical experiments, makes physical sense because it corresponds to submesoscale processes of the ocean (Benway et al., 2019).

3.2.3 Choice of memory for the neural network

The temporal convolutional network also has parameters associated with the dynamics in time, i.e. how much memory from the input should be retained in order to achieve the best prediction. Starting with time series arrays of surface temperature, T_S, PCA coefficients, q_1 and q_1, and mean temperature, $\bar{T}$,

$$\mathbf{x} = \begin{bmatrix} T_S(t_0) & T_S(t_1) & T_S(t_2) & ... & T_S(t_n) \end{bmatrix} \tag{8}$$

$$\mathbf{y} = \begin{bmatrix} q_1(t_0) & q_1(t_1) & q_1(t_2) & ... & q_1(t_n) \\ q_2(t_0) & q_2(t_1) & q_2(t_2) & ... & q_2(t_n) \\ \bar{T}(t_0) & \bar{T}(t_1) & q_2\bar{T} & ... & \bar{T}(t_n) \end{bmatrix} \tag{9}$$

we build matrices of smaller sequences on which we apply the convolutional filter.

$$\mathbf{x}_{TCN} = \begin{bmatrix} T_S(t_0) & T_S(t_1) & ... & T_S(t_m) \\ T_S(t_s) & T_S(t_{s+1}) & ... & T_S(t_{m+1}) \\ & ... & & \\ T_S(t_{n-m+1}) & ... & & T_S(t_n) \end{bmatrix} \tag{10}$$

$$\mathbf{y}_{TCN} = \begin{bmatrix} q_1(t_1) & q_1(t_2) & ... & q_1(t_m) \\ q_1(t_s) & q_2(t_{s+1}) & ... & q_1(t_{m+1}) \\ & ... & & \\ q_1(t_{n-m+1}) & ... & & q_1(t_n) \end{bmatrix} \tag{11}$$

When building these smaller sequences, we have the ability to choose how much data to use which affects the performance of the neural network. The sampling rate determines how many time steps to skip within an input time series, the stride, s, determines how many time steps to skip between each successive time series, and the memory length scale, m, determines how many points back in time to consider in one time series. Again, we perform numerical experiments to find the values for these parameters that result in the lowest testing error. The memory length scale is set to be 20 days, and the sampling rate and stride are both set to 1 day. In our final model, each PCA coefficient is predicted using the surface temperature from all of the data from the 20 previous days, a choice that is consistent with ocean time scales for upwelling and eddies (Benway et al., 2019).

3.2.4 Surface temperature constraint

The output of the neural network is used to reconstruct the full 3D temperature field, but we want to ensure that the prediction at the surface of the ocean matches exactly the input surface temperature:

$$q_1\boldsymbol{\phi}_1(z=0) + q_2\boldsymbol{\phi}_2(z=0) + \bar{T} = T(z=0) \tag{12}$$

This requirement can be written as a constraint function

$$f(\hat{\mathbf{y}}(\boldsymbol{\theta})) = q_1\boldsymbol{\phi}_1(z=0) + q_2\boldsymbol{\phi}_2(z=0) + \bar{T} - T(z=0) \tag{13}$$

We embed the soft constraint $\lambda|f(\hat{\mathbf{y}}(\boldsymbol{\theta}))|$ into the loss function

$$J(\boldsymbol{\theta}) = \frac{1}{T}\sum L(\hat{\mathbf{y}}(\boldsymbol{\theta}) - \mathbf{y}) + \lambda|f(\hat{\mathbf{y}}(\boldsymbol{\theta}))| \tag{14}$$

From numerical experiments, we find that the neural network is able to match the surface temperature without the soft constraint. Nevertheless the inclusion of the constraint guarantees that there will be no significant deviations.

3.2.5 Results of the neural network training

By using additional nearby points and previous time steps, we create a non-local parametrization in both space and time. The neural network is built using four years of data for training (mid 2005 - mid 2009), one and a half years for validation (mid 2009 - 2011), and two and a half years for testing (2011 until mid-2013) (Figure 8). The error associated with the neural network predictions are calculated relative to both the original reanalysis data and the PCA reconstruction (Table 1). The predicted time series for a representative horizontal location, as well as the predicted standard deviation, are shown in Figure 8.

Table 1. Neural Network Model Evaluation

			MAE ($°C$)			RMSE ($°C$)		
location	y (target)	$\hat{y}$ (output)	train	val.	test	train	val.	test
full field	FVCOM	TCN	0.2088	0.2999	0.3185	0.3552	0.4961	0.5470
full field	PCA	TCN	0.1942	0.2846	0.3078	0.3359	0.4777	0.5364
surface	FVCOM	TCN	0.4178	0.6818	0.7051	0.6657	1.0275	1.1249
surface	PCA	TCN	0.4406	0.7005	0.7384	0.7005	1.0601	1.1816

The raw outputs of the neural network are simply the PCA coefficients and mean temperature, as well as their standard deviations. However, these raw outputs can be combined with the PCA modes to reconstruct the full 3D temperature field (Figure 9). For each (x, y) pair, it takes one minute to train a neural network on a standard CPU. Once the neural network is fully optimized, it only takes a few seconds to make a prediction.

3.3 Filling Gaps in the Surface Sensor Data

Satellites provide useful information about surface temperature, but they are significantly affected by cloud coverage. Work has been done to improve measurements from satellites and to blend data from multiple satellites (Chin et al., 2017; Zhu et al., 2019). In many projects, in-situ buoy measurements are used to either validate or improve the accuracy of models (Zhu et al., 2018; Donlon et al., 2007; A. Li et al., 2013; Reynolds, 1988; Reynolds & Smith, 1994; Zhu et al., 2015; Chao et al., 2009). One recent approach that has been shown to obtain quick, accurate, and useful results is Gaussian process regression (GPR) (Babaee et al., 2020; Raissi et al., 2019). GPR is not ideal because the matrix inversion becomes slow for large numbers of input points. However, GPR is convenient for problems with a low number of input points. Furthermore, unlike with other machine learning techniques, the hyperparameters of the model, specifically those of the kernel, have an intuitive physical meaning and can be set according to properties of the system. Here, we use GPR to essentially *extrapolate* the available surface data. Note that we use the term extrapolation (as opposed to interpolation) since in many cases the available surface data are so sparse that interpolation is not meaningful. The features (inputs) of the model are the longitude, latitude, and time, and the value that is being predicted is the surface temperature. For points at which sensor data are available, we keep the original data, but for points at which there are no measurements, we predict the temperature using nearby points both in time and space.

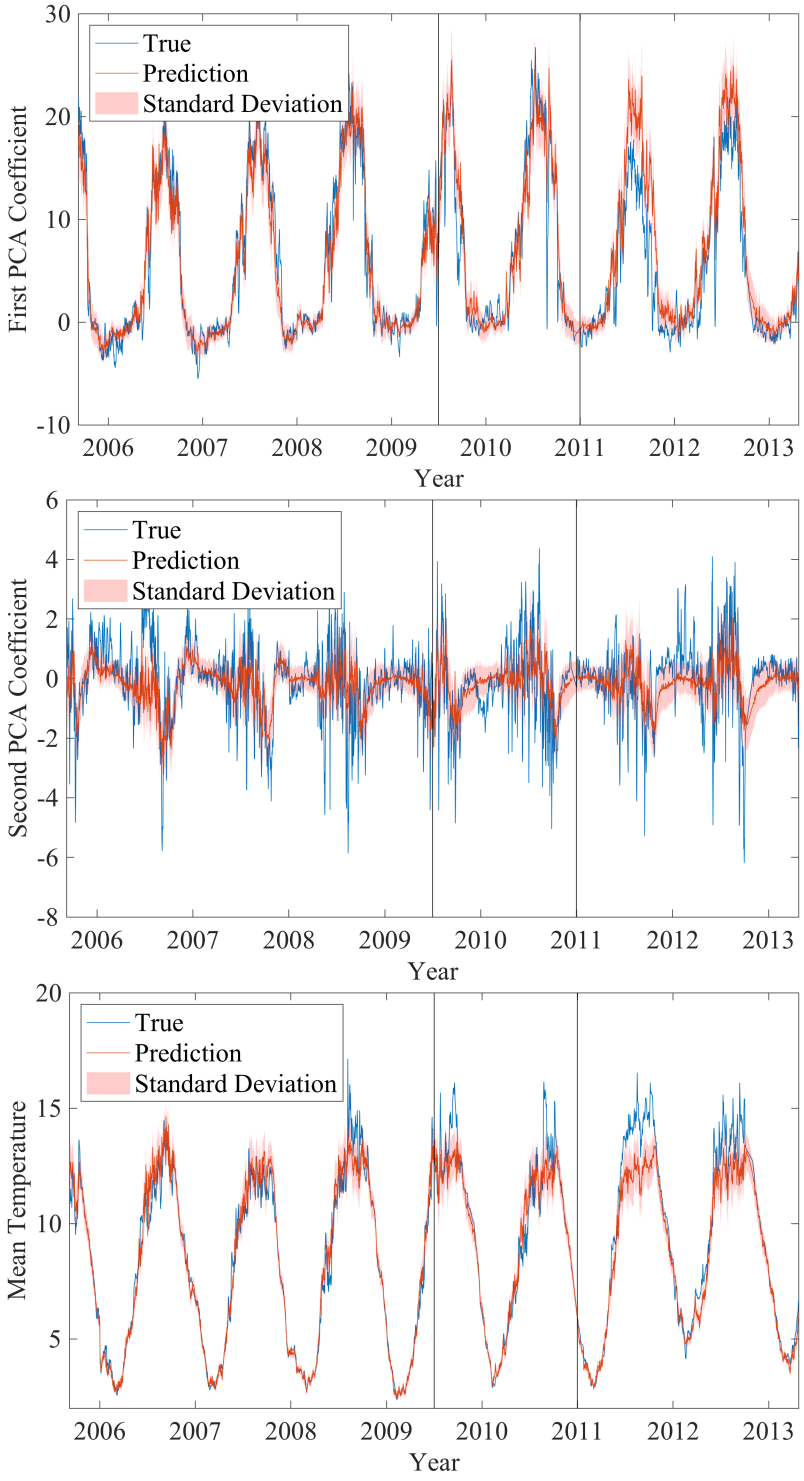

Figure 8. TCN Predictions. The first and second PCA coefficients and the mean temperature, as well as their standard deviation, are predicted for the reanalysis data. The black lines delineate the training, validation, and test sets, respectively.

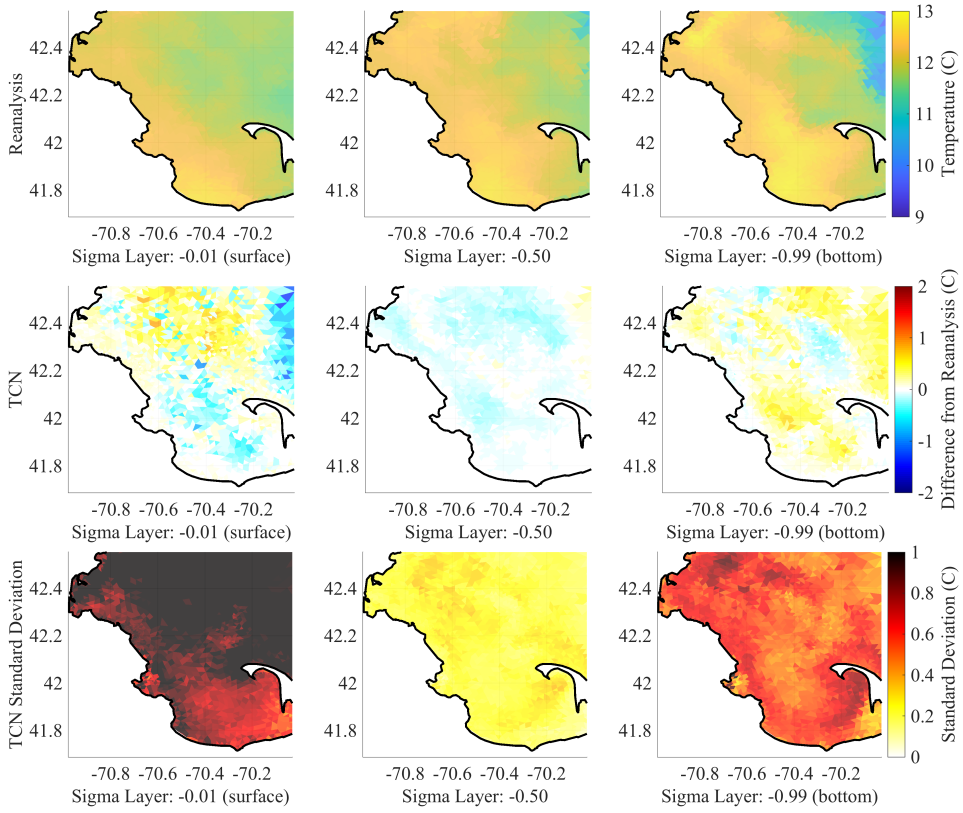

Figure 9. TCN 3D Reconstruction of the Temperature Field. From top to bottom, the reanalysis data, TCN prediction, and TCN standard deviation prediction are plotted for March 8th, 2012 at three different sigma levels.

3.3.1 Gaussian process regression

The mean and variance are predicted using the kernel, K, which relates all of the available data points (Rasmussen & Williams, 2006). Specifically, the mean prediction is

$$\bar{\mathbf{f}}_* = \mathbf{m}(X_*) + K(X_*, X)[K(X, X) + \sigma_n^2 I]^{-1}(\mathbf{y} - \mathbf{m}(X)) \tag{15}$$

and the variance is

$$cov(\mathbf{f}_*) = K(X_*, X_*) - K(X_*, X)[K(X, X) + \sigma_n^2 I]^{-1} K(X, X_*) \tag{16}$$

For our application, the mean function $\mathbf{m}(X_*)$ is explicitly set to be the spatial mean (Figure 10) of the available satellite data. To avoid inverting large matrices, we keep the size of the kernel small by building a new GPR model for each time step. The input data consists of the available data on the day of interest, data from one day before and one day after. In other words, we only use data from three days to predict the surface temperature for one day, and we repeat this process for all time steps. The features for each time step k are

$$\begin{bmatrix} x_i & y_i & t_{k-1} \\ x_i & y_i & t_k \\ x_i & y_i & t_{k+1} \end{bmatrix} = \begin{bmatrix} x_i & y_i & -1 \\ x_i & y_i & 0 \\ x_i & y_i & 1 \end{bmatrix} \tag{17}$$

where (x_i, y_i) are all of the available spatial points at each time step k.

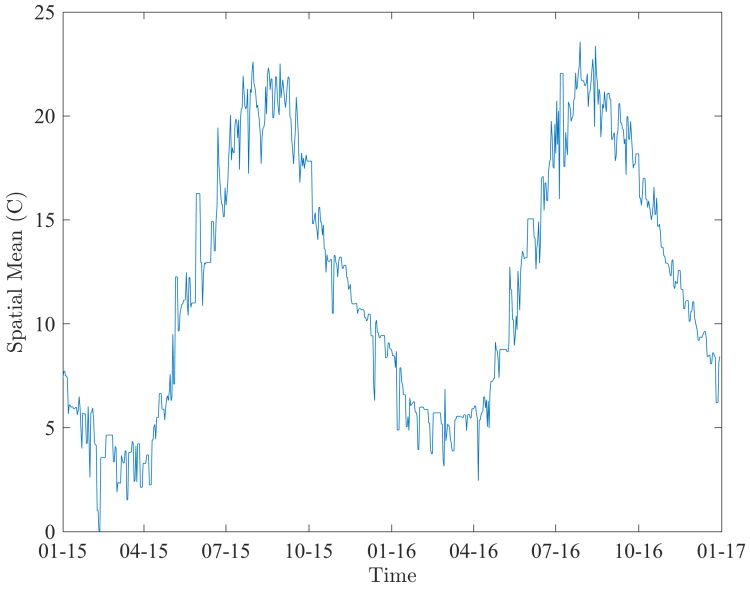

Figure 10. Spatial Mean of the Satellite Data. The mean over the whole spatial domain is plotted for each day.

3.3.2 Hyperparameter selection

For the kernel, we use the radial basis function (RBF) with automatic relevance determination as the covariance function.

$$cov(f(\mathbf{x_p}), f(\mathbf{x_q})) = k(\mathbf{x_p}, \mathbf{x_q}) = \sigma_f^2 * exp(-\frac{1}{2}(\mathbf{x_q} - \mathbf{x_p})^T \theta (\mathbf{x_q} - \mathbf{x_p})) \tag{18}$$

The signal variance σ_f and characteristic lengthscales $\mathbf{\Theta}$ are hyperparameters of the model. The characteristic lengthscale represents how far apart two points need to be for their function values to become uncorrelated. The inverse of the lengthscale represents how relevant a given feature is. The automatic relevance determination chooses different characteristic lengthscales for each input to determine the relevant inputs. As such, there are three characteristic lengthscales: one for the input longitude, one for the input latitude, and one for the input time. The noise variance, σ_n, is not a parameter of the kernel, but it can also be considered one of the hyperparameters of the whole system. This parameter assumes that we know the uncertainty of the sensors.

Typically, the hyperparameters are found by optimizing the following loss function.

$$\log p(\mathbf{y}|X) = -\frac{1}{2}\mathbf{y}^T(K + \sigma_n^2 I)^{-1}\mathbf{y} - \frac{1}{2}\log|K + \sigma_n^2 I| - \frac{n}{2}\log 2\pi \tag{19}$$

However, to avoid overfitting and to generalize the models, we manually set the same hyperparameters for all days, changing only the training data for each day. For days with no available training data, we take the average over 10 days (5 previous and 5 following days). For the spatial lengthscales, we choose a value of 0.25 degrees or 25 kilometers, which is equivalent to six "gridpoints" or "pixels," where one gridpoint is the spatial granularity. This choice assigns more weight to spatial points that are within 25 kilometers of the point of interest; it corresponds to the mesoscales of the ocean (Benway et al., 2019). For the time lengthscale, we set the hyperparameter to one day. Finally, we choose to set the noise variance to $\sigma_n = 0.1$, and we set the signal variance to $\sigma_f = 0.3$ by taking the average of minimizing the objective function over all models.

3.3.3 Multi-fidelity Gaussian process regression

We improve the model by incorporating the buoy data, which has lower uncertainty than the satellite data, through a recursive multi-fidelity Gaussian process regression scheme. The satellite and buoy measurements are shown to have good agreement (Babaee et al., 2020). Given s levels of fidelity, the model with the lowest fidelity is denoted with $\mathbf{x_1}$, y_1, $\bar{\mathbf{f}}_{*1}$, and the model with the highest fidelity is denoted with $\mathbf{x_s}$, y_s, $\bar{\mathbf{f}}_{*s}$ (Perdikaris et al., 2015). The prediction for the model with the lowest fidelity follows the Gaussian process regression steps from equations (15) and (16)

$$\bar{\mathbf{f}}_1(\mathbf{x}_*) = K(X_*, X_1)[K(X_1, X_1) + \sigma_{n1} I]^{-1}\mathbf{y}_1,$$

with covariance

$$cov(\bar{\mathbf{f}}_1) = K(X_*, X_*) - K(X_*, X_1)[K(X_1, X_1) + \sigma_{n1} I]^{-1}K(X_1, X_*).$$

Each following model has the form

$$\bar{\mathbf{f}}_t(\mathbf{x}_*) = \rho_{t-1}\bar{\mathbf{f}}_{t-1} + \delta_t \qquad t = 2, ..., s \tag{20}$$

In this project, there are only two levels of fidelity, so the prediction for the highest level of fidelity, $s = 2$, can be computed with the following equation

$$\bar{\mathbf{f}}_2(\mathbf{x}_*) = \rho\bar{\mathbf{f}}_1(\mathbf{x}_*) + \mu_d + K(X_*, X_2)[K(X_2, X_2) + \sigma_{n2} I]^{-1}(\mathbf{y} - \rho\bar{\mathbf{f}}_1(\mathbf{x}_2) - \mu_d).$$

Its corresponding covariance is

$$cov(\bar{\mathbf{f}}_2) = \rho^2 cov(\bar{\mathbf{f}}_1) + K(X_*, X_*) - K(X_*, X_2)[K(X_2, X_2) + \sigma_{n2} I]^{-1}K(X_2, X_*),$$

where, ρ and μ_d are hyperparameters that are different for each level of fidelity. Like σ_f and θ of the covariance function, ρ and μ_d can be chosen through maximum likelihood estimation or other optimization techniques. We use the Emukit (Paleyes et al., 2019) Python package, which builds on the GPy Python package, to build the multi-fidelity model. Such techniques have already been used to predict surface temperature ((Babaee et al., 2020)), but our model differs with respect to the choice of input points. Babaee et al. (2020) used all of the available data to build a model while we only use spatial points from three time steps. Because we use less data at each time step, our model is faster at making predictions, and therefore more practical for real-time modeling. For consistency, we set ρ and μ_d to be the same as those from the optimized model in Babaee et al. (2020).

3.3.4 Results of surface temperature extrapolation

The results of the extrapolation are shown in Figure 11 both for a day with high cloud coverage (March 8th, 2016) and for a day with minimal cloud coverage (September 13th, 2016). The uncertainty of the extrapolation is higher in regions with significant cloud coverage. Overall, the results from using just three days compare favorably with those from Babaee et al. (2020), while the new model is significantly faster.

4 Results and Validation of the Full 3D Temperature Field

Finally, we utilize the real-time estimate for surface temperature obtained from GPR as input to the TCN to obtain the PCA coefficients and the mean temperature, as well as their uncertainty, at each horizontal location for the day of interest. When estimating surface temperature, we left out measurements from three stations (N04, F13, F29 from Figure 3) which we saved for validation. These stations also collect measurements for temperature over multiple depths, which we divide into shallow (0-25m), medium (25-45m), and deep (>45m). We also have NOAA measurements from environmental monitors on lobster traps and large

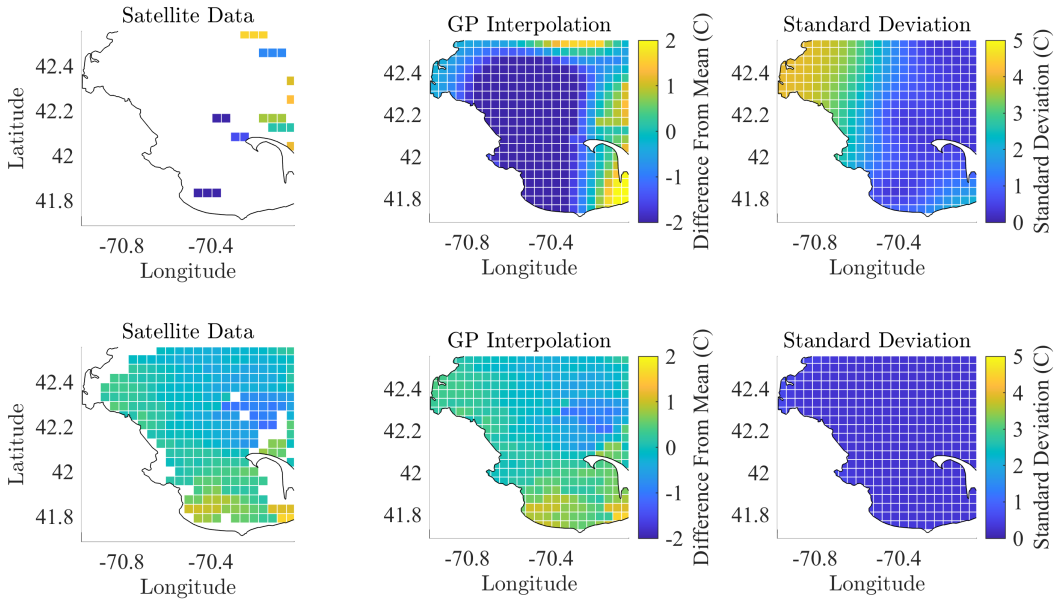

Figure 11. Results of Extrapolation for Two Different Days. The available satellite and buoy data are extrapolated to obtain a surface temperature field over the full domain. Each row represents a different day with high cloud coverage (upper row) and low cloud coverage (lower row)

trawlers (eMOLT in Figure 12) and the Northeast Fisheries Science Center (NEFSC in Figure 12).

The neural network predictions from the real-time sensor measurements are plotted in Figure 13 for 2015 and 2016. The neural network provides an estimate for the mean and standard deviation (red shading) of the quantities of interest. The predicted PCA coefficients are then projected onto the deterministic PCA modes and summed with the predicted PCA mean to reconstruct the full 3D temperature and uncertainty fields. The results of the full 3D reconstruction are plotted in Figure 14 for March 8th, 2016 and September 16th, 2016 at three sigma layers.

We validate the results of our full model by comparing the predictions from the neural network to the in-situ measurements that are not used in the training phase. For the MWRA measurements, which are the most reliable in-situ measurements, we find that our model performs similarly to other comprehensive, but very expensive, ocean models. The mean absolute error of our predictions is 1.37°C, and 79% of predictions fall withing two degrees of the truth. We also observe that the model performs best for days with the most amount of available satellite data (80-100%). We found more significant errors when we applied our analysis to the eMOLT and NEFSC data sets (Figure 15); some of the errors seem to be associated with system trends which require further investigation (e.g. condition of sensors, calibration, etc.). We find that there are no significant improvements from including the buoy measurements when modeling the surface temperature. However, the framework allows us to seamlessly incorporate data from multiple sources which could be useful in applications where data are more sparse. Furthermore, the framework provides an estimate for uncertainty given the level of accuracy of each sensor.

5 Conclusions

We introduced a fast and accurate framework, based on recently developed machine learning techniques and reanalysis data obtained from comprehensive ocean models, to

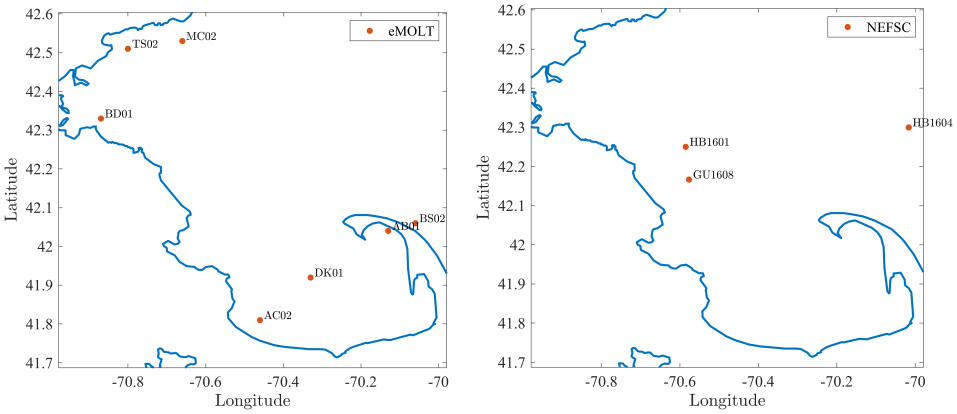

Figure 12. Location of In-situ NOAA Sensors for Validation. Left: environmental monitors on lobster traps and large trawlers (eMOLT). Right: Northeast Fisheries Science Center (NEFSC).

reconstruct 3D ocean temperature fields from real-time sensor measurements of surface temperature. We compared the results from our framework to in-situ measurements, and we found that the error associated with our predictions is comparable to that of other state of the art models that are significantly more expensive. We also demonstrated how our model can evaluate the quality of sensor measurements from different sources. In the future, we plan to use our model's estimates of uncertainty to make decisions about the system, a process often referred to as active sampling or optimal sampling. For example, we can define and optimize an acquisition function to decide where to place additional sensors or plan the trajectory of an ocean drifter. In some cases, properly formulated acquisition functions can be leveraged to identify extreme values (Y. Yang et al., 2022). Overall, the developed model has many applications, ranging from monitoring general ocean health for fisheries to understanding changes in ocean acidification, and the techniques described can be used for other geophysical systems.

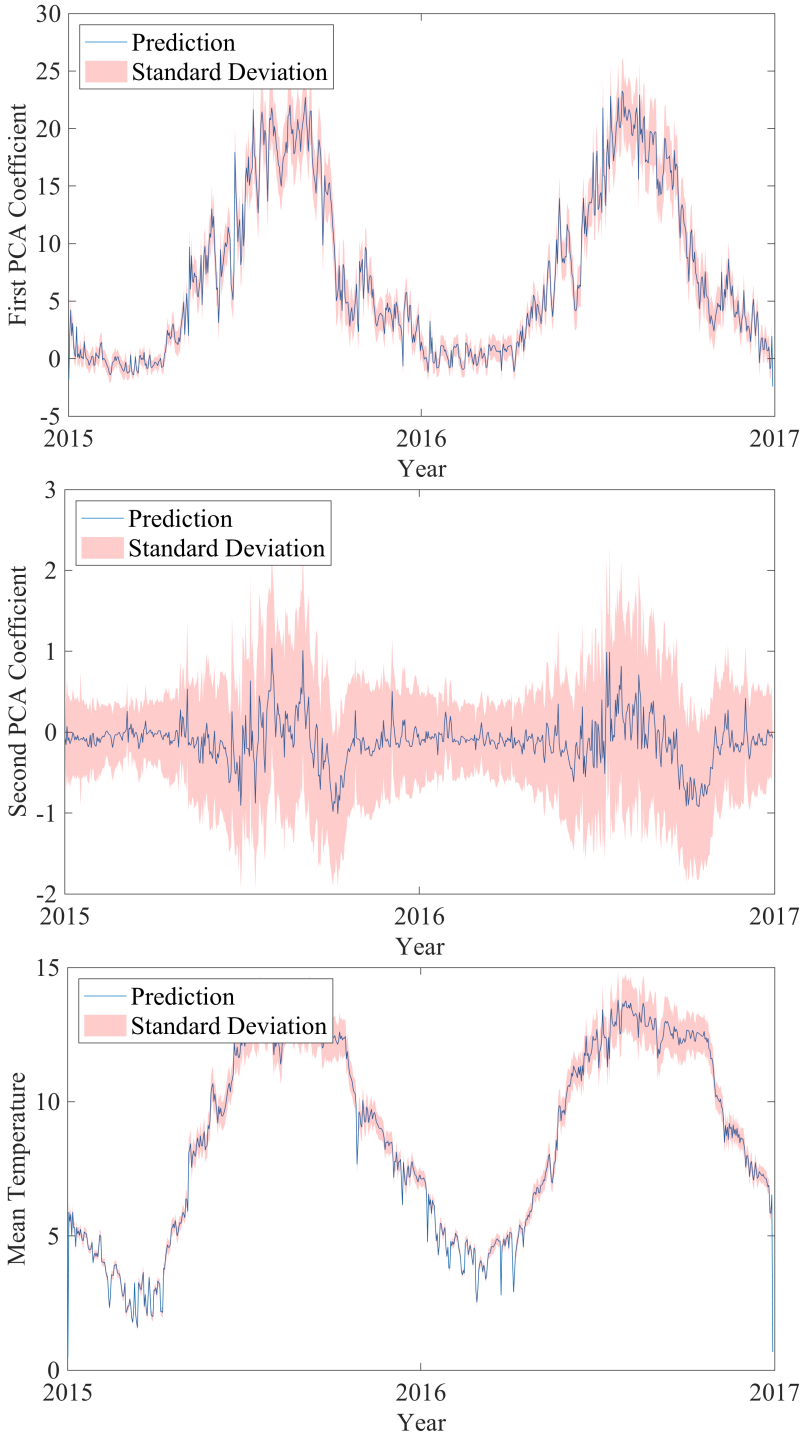

Figure 13. TCN Predictions from Satellite Measurements. The first and second PCA coefficients, the mean temperature, as well as their uncertainties (red shading) are predicted for the available satellite surface temperature.

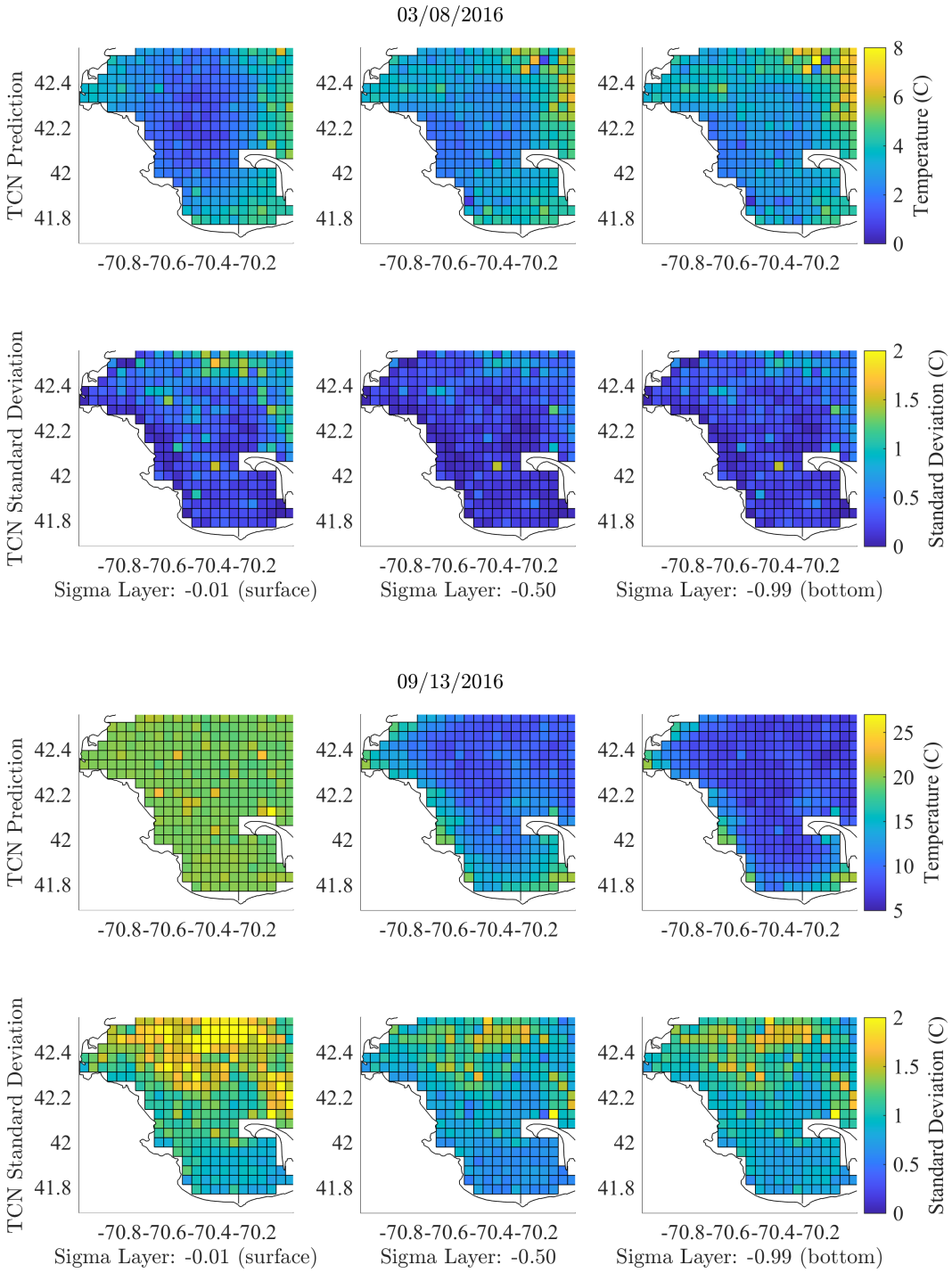

Figure 14. 3D Temperature Field Reconstruction From Real-time Sensor Measurements. The predicted PCA coefficients are projected onto the corresponding modes and summed with the predicted mean temperature to reconstruct the full 3D temperature. The results and associated uncertainty are plotted for March 8th, 2016 and September 13th, 2016 at three sigma layers.

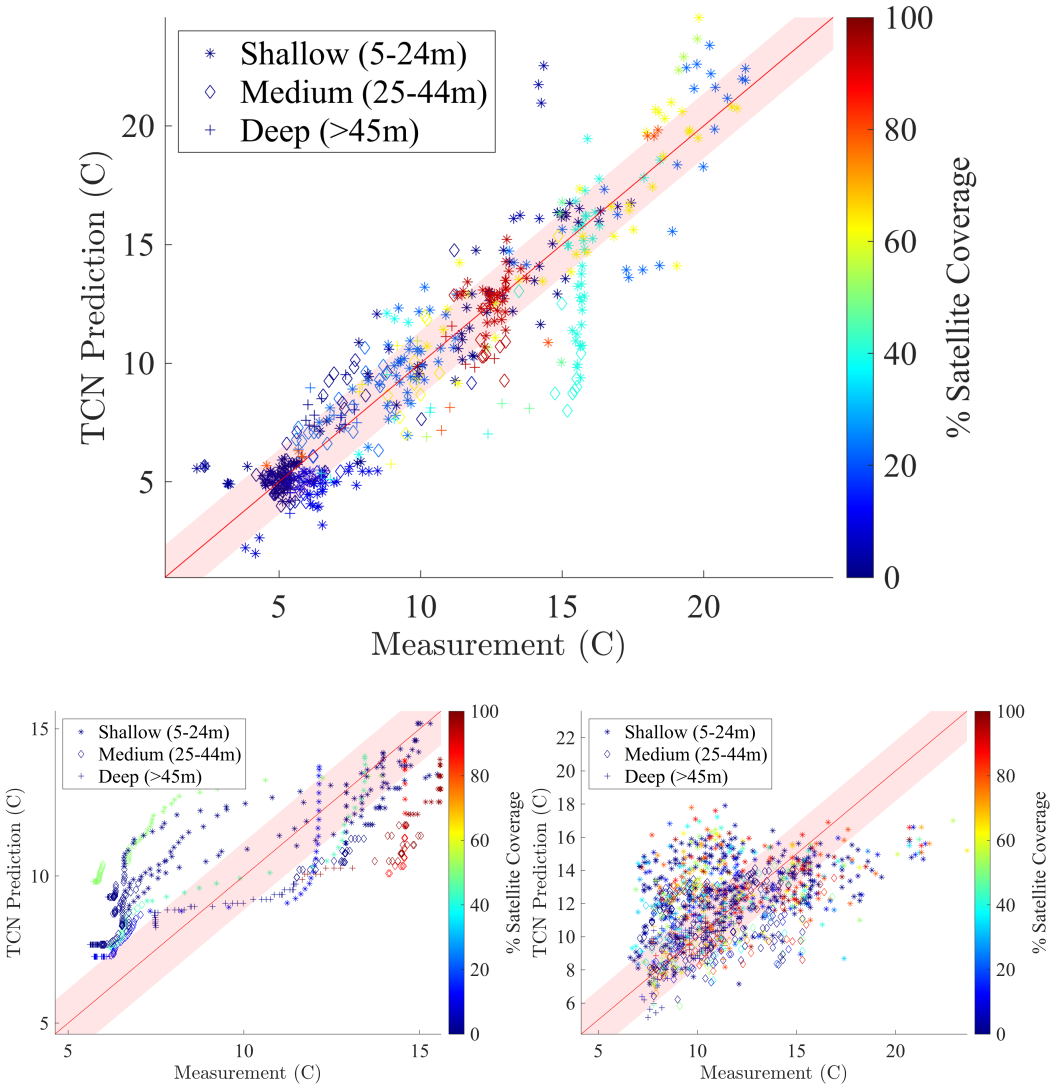

Figure 15. Validation Comparison between buoy measurements and predictions from the neural network at different depths for days with different amounts of satellite coverage. We split the data based on its source: MWRA (top plot), NEFSC (bottom left), EMOLT (bottom right). The red shading corresponds to the standard deviation of the absolute error.

Appendix A Open Research

The Finite Volume Community Ocean Model (FVCOM) data are available from the The Northeast Coastal Ocean Forecast System (NECOFS): http://fvcom.smast.umassd.edu/necofs/. The Moderate-resolution Imaging Spectroradiometer (MODIS) SST data come from the NASA EOSDIS Physical Oceanography Distributed Active Archive Center (PO.DAAC) at the Jet Propulsion Laboratory, in Pasadena, CA (https://doi.org/10.5067/MODST-1D4N4). The MWRA measurements are accessible at https://www.mwra.com/harbor/html/wq_data.htm. The data from Environmental Monitors on Lobster Traps and Large Trawlers (eMOLT) and the Northeast Fisheries Science Center (NEFSC) oceanographic profile data set can be downloaded from the National Ocean and Atmospheric Administration's Environmental Research Division's Data Access Program (NOAA ERDDAP) website at https://comet.nefsc.noaa.gov/erddap/tabledap/index.html. The temporal convolutional network was built with Tensorflow, and the multi-fidelity Gaussian process regression was implemented with Emukit.

Acknowledgments

This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. 1745302, MIT Sea Grant, and the Harrington Fellowship. We are grateful to Dr. Carolina Bastidas and Michael Defilippo for providing and pre-processing the buoy data. We also thank Professor George Karniadakis and Professor Hessam Babaee for stimulating discussions.

References

Aksan, E., & Hilliges, O. (2019). STCN: Stochastic temporal convolutional networks. In *International conference on learning representations.*

Ali, M. M., Swain, D., & Weller, R. A. (2004). Estimation of ocean subsurface thermal structure from surface parameters: A neural network approach. *Geophysical Research Letters*, *31*(20). doi: https://doi.org/10.1029/2004GL021192

Babaee, H., Bastidas, C., Defilippo, M., Chryssostomidis, C., & Karniadakis, G. (2020, 01). A multi-fidelity framework and uncertainty quantification for sea surface temperature in the Massachusetts and Cape Cod Bays. *Earth and Space Science*, *7*. doi: 10.1029/2019EA000954

Bai, S., Kolter, J. Z., & Koltun, V. (2018). *An empirical evaluation of generic convolutional and recurrent networks for sequence modeling.*

Baño Medina, J., Manzanas, R., & Gutiérrez, J. M. (2020). Configuration and intercomparison of deep learning neural models for statistical downscaling. *Geoscientific Model Development*, *13*(4), 2109–2124. doi: 10.5194/gmd-13-2109-2020

Benway, H. M., Lorenzoni, L., White, A. E., Fiedler, B., Levine, N. M., Nicholson, D. P., ... Letelier, R. M. (2019). Ocean time series observations of changing marine ecosystems: An era of integration, synthesis, and societal applications. *Frontiers in Marine Science*, *6*. Retrieved from `https://www.frontiersin.org/article/10.3389/fmars.2019.00393` doi: 10.3389/fmars.2019.00393

Berliner, L. M., Wikle, C. K., & Cressie, N. (2000). Long-lead prediction of pacific SSTs via Bayesian dynamic modeling. *Journal of Climate*, *13*(22), 3953 - 3968. doi: 10.1175/1520-0442(2001)013⟨3953:LLPOPS⟩2.0.CO;2

Bolton, T., & Zanna, L. (2019). Applications of deep learning to ocean data inference and subgrid parameterization. *Journal of Advances in Modeling Earth Systems*, *11*(1), 376-399. Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018MS001472` doi: https://doi.org/10.1029/2018MS001472

Chao, Y., Li, Z., Farrara, J. D., & Hung, P. (2009). Blending sea surface temperatures from multiple satellites and in situ observations for coastal oceans. *Journal of Atmospheric and Oceanic Technology*, *26*(7), 1415 - 1426. doi: 10.1175/2009JTECHO592.1

Chen, C., Liu, H., & Beardsley, R. C. (2003). An unstructured grid, finite-volume, three-dimensional, primitive equations ocean model: application to coastal ocean and estuaries [dataset]. *Journal of Atmospheric and Oceanic Technology*, *20*(1), 159 - 186. doi: 10.1175/1520-0426(2003)020⟨0159:AUGFVT⟩2.0.CO;2

Chin, T. M., Vazquez-Cuervo, J., & Armstrong, E. M. (2017). A multi-scale high-resolution analysis of global sea surface temperature. *Remote Sensing of Environment*, *200*, 154-169. doi: https://doi.org/10.1016/j.rse.2017.07.029

Donlon, C., Robinson, I., Casey, K. S., Vazquez-Cuervo, J., Armstrong, E., Arino, O., ... Rayner, N. (2007). The global ocean data assimilation experiment high-resolution sea surface temperature pilot project. *Bulletin of the American Meteorological Society*, *88*(8), 1197 - 1214. Retrieved from `https://journals.ametsoc.org/view/journals/bams/88/8/bams-88-8-1197.xml` doi: 10.1175/BAMS-88-8-1197

Ekstrom, J. A., Suatoni, L., Cooley, S. R., Pendleton, L. H., Waldbusser, G. G., Cinner, J. E., ... Portela, R. (2015, Mar 01). Vulnerability and adaptation of us shellfisheries to ocean acidification. *Nature Climate Change*, *5*(3), 207-214. doi: 10.1038/nclimate2508

Everson, R., Cornillon, P., Sirovich, L., & Webber, A. (1996). Empirical eigenfunction analysis of sea surface temperatures in the Western North Atlantic. *AIP Conference Proceedings*, *375*(1), 563-590. doi: 10.1063/1.50998

Everson, R., & Sirovich, L. (1995, Aug). Karhunen–Loève procedure for gappy data. *J. Opt. Soc. Am. A*, *12*(8), 1657–1664. doi: 10.1364/JOSAA.12.001657

Friedland, K. D., Morse, R. E., Manning, J. P., Melrose, D. C., Miles, T., Goode, A. G., ... Powell, E. N. (2020). Trends and change points in surface and bottom thermal environments of the US Northeast Continental Shelf ecosystem. *Fisheries Oceanography*, *29*(5), 396-414. doi: https://doi.org/10.1111/fog.12485

Gan, Z., Li, C., Zhou, J., & Tang, G. (2021). Temporal convolutional networks interval prediction model for wind speed forecasting. *Electric Power Systems Research*, *191*, 106865. doi: https://doi.org/10.1016/j.epsr.2020.106865

Ganzedo, U., Alvera-Azcárate, A., Esnaola, G., Ezcurra, A., & Sáenz, J. (2011). Reconstruction of sea surface temperature by means of DINEOF: a case study during the fishing season in the Bay of Biscay. *International Journal of Remote Sensing*, *32*(4), 933-950. doi: 10.1080/01431160903491420

Gledhill, D. K., White, M. M., Salisbury, J., Thomas, H., Mook, B., Grear, J., ... Doney, S. C. (2015, June). Ocean and coastal acidification off New England and Nova Scotia. *Oceanography*.

Hannachi, A., Jolliffe, I. T., & Stephenson, D. B. (2007). Empirical orthogonal functions and related techniques in atmospheric science: a review. *International Journal of Climatology*, *27*(9), 1119-1152. Retrieved from `https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/joc.1499` doi: https://doi.org/10.1002/joc.1499

Juranek, L. W., Feely, R. A., Peterson, W. T., Alin, S. R., Hales, B., Lee, K., ... Peterson, J. (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). doi: https://doi.org/10.1029/2009GL040778

Klemas, V., & Yan, X.-H. (2014, Mar 01). Subsurface and deeper ocean remote sensing from satellites: an overview and new results. *Progress in Oceanography*, *122*, 1-9. Retrieved from `https://www.sciencedirect.com/science/article/pii/S0079661113002310`

Lara-Benítez, P., Carranza-García, M., Luna-Romera, J. M., & Riquelme, J. C. (2020). Temporal convolutional networks applied to energy-related time series forecasting. *Applied Sciences*, *10*(7). doi: 10.3390/app10072322

Lermusiaux, P. F., Malanotte-Rizzoli, P., Stammer, D., Cummings, J., & Moore, A. M. (2006, March). Progress and prospects of US data assimilation in ocean research. *Oceanography*, *issue volume*.

Li, A., Bo, Y., Zhu, Y., Guo, P., Bi, J., & He, Y. (2013). Blending multi-resolution satellite sea surface temperature (SST) products using Bayesian maximum entropy method. *Remote Sensing of Environment*, *135*, 52-63. doi: https://doi.org/10.1016/j.rse.2013.03.021

Li, B., Tanaka, K. R., Chen, Y., Brady, D. C., & Thomas, A. C. (2017). Assessing the quality of bottom water temperatures from the finite-volume community ocean model (FVCOM) in the Northwest Atlantic Shelf region. *Journal of Marine Systems*, *173*, 21-30. doi: https://doi.org/10.1016/j.jmarsys.2017.04.001

Lorenz, E. N. (n.d.). *Empirical orthogonal functions and statistical weather prediction.* Massachusetts Institute of Technology, Department of Meteorology.

Maulik, R., Egele, R., Lusch, B., & Balaprakash, P. (2020). Recurrent neural network architecture search for geophysical emulation. In *Sc20: International conference for high performance computing, networking, storage and analysis* (p. 1-14). doi: 10.1109/SC41405.2020.00012

Mellor, G. L., Häkkinen, S. M., Ezer, T., & Patchen, R. C. (2002). A generalization of a sigma coordinate ocean model and an intercomparison of model vertical grids. In N. Pinardi & J. Woods (Eds.), *Ocean forecasting: Conceptual basis and applications* (pp. 55–72). Berlin, Heidelberg: Springer Berlin Heidelberg. Retrieved from `https://doi.org/10.1007/978-3-662-22648-3_4` doi: 10.1007/978-3-662-22648-3_4

Meng, X., & Karniadakis, G. E. (2020). A composite neural network that learns from multi-fidelity data: application to function approximation and inverse PDE problems. *Journal of Computational Physics*, *401*, 109020. doi: https://doi.org/10.1016/j.jcp.2019.109020

Miao, Q., Pan, B., Wang, H., Hsu, K., & Sorooshian, S. (2019). Improving monsoon precipitation prediction using combined convolutional and long short term memory neural network. *Water*, *11*(5). doi: 10.3390/w11050977

Paleyes, A., Pullin, M., Mahsereci, M., Lawrence, N., & González, J. (2019). Emulation of physical processes with Emukit. In *Second workshop on machine learning and the physical sciences, neurips.*

Perdikaris, P., Venturi, D., Royset, J. O., & Karniadakis, G. E. (2015). Multi-fidelity modelling via recursive co-kriging and Gaussian-Markov random fields. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *471*(2179), 20150018. Retrieved from `https://royalsocietypublishing.org/doi/abs/10.1098/rspa.2015.0018` doi: 10.1098/rspa.2015.0018

Raissi, M., Babaee, H., & Karniadakis, G. E. (2019). Parametric Gaussian process regression for big data. *Computational Mechanics*, *64*.

Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2017). Machine learning of linear differential equations using Gaussian processes. *Journal of Computational Physics*, *348*, 683-693. doi: https://doi.org/10.1016/j.jcp.2017.07.050

Rasmussen, C. E., & Williams, C. K. I. (2006). *Gaussian processes for machine learning*. Cambridge, MA, USA: MIT Press.

Reynolds, R. W. (1988). A real-time global sea surface temperature analysis. *Journal of Climate*, *1*(1), 75 - 87. doi: 10.1175/1520-0442(1988)001⟨0075:ARTGSS⟩2.0.CO;2

Reynolds, R. W., & Smith, T. M. (1994). Improved global sea surface temperature analyses using optimum interpolation. *Journal of Climate*, *7*(6), 929 - 948. doi: 10.1175/1520-0442(1994)007⟨0929:IGSSTA⟩2.0.CO;2

Sapsis, T. P. (2020). Output-weighted optimal sampling for Bayesian regression and rare event statistics using few samples. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *476*(2234), 20190834. doi: 10.1098/rspa.2019.0834

Smith, T. M., Reynolds, R. W., Livezey, R. E., & Stokes, D. C. (1996). Reconstruction of historical sea surface temperatures using empirical orthogonal functions. *Journal of Climate*, *9*(6), 1403 - 1420. doi: 10.1175/1520-0442(1996)009⟨1403:ROHSST⟩2.0.CO;2

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. doi: https://doi.org/10.1016/j.ecss.2014.05.027

Wan, Z. Y., Dodov, B., Lessig, C., Dijkstra, H., & Sapsis, T. P. (2021, October). A data-driven framework for the stochastic reconstruction of small-scale features with application to climate data sets. *Journal of Computational Physics*, *442*, 110484. doi: 10.1016/j.jcp.2021.110484

Weyn, J. A., Durran, D. R., & Caruana, R. (2020). Improving data-driven global weather prediction using deep convolutional neural networks on a cubed sphere. *Journal of Advances in Modeling Earth Systems*, *12*(9), e2020MS002109. Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2020MS002109` (e2020MS002109 10.1029/2020MS002109) doi: https://doi.org/10.1029/2020MS002109

Yan, J., Mu, L., Wang, L., Ranjan, R., & Zomaya, A. Y. (2020). Temporal convolutional networks for the advance prediction of ENSO. *Scientific Reports*. doi: https://doi.org/10.1038/s41598-020-65070-5

Yang, W., Yao, Q., Ye, K., & Xu, C.-Z. (2020, Feb 01). Empirical mode decomposition and temporal convolutional networks for remaining useful life estimation. *International Journal of Parallel Programming*, *48*(1), 61-79. doi: 10.1007/s10766-019-00650-1

Yang, X., Venturi, D., Chen, C., Chryssostomidis, C., & Karniadakis, G. E. (2010). EOF-based constrained sensor placement and field reconstruction from noisy ocean measurements: application to Nantucket Sound. *Journal of Geophysical Research: Oceans*, *115*(C12). doi: https://doi.org/10.1029/2010JC006148

Yang, Y., Blanchard, A., Sapsis, T., & Perdikaris, P. (2022). Output-weighted sampling for multi-armed bandits with extreme payoffs. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *478*(2260), 20210781. Retrieved from `https://royalsocietypublishing.org/doi/abs/10.1098/rspa.2021.0781` doi:

10.1098/rspa.2021.0781

Zhu, Y., Bo, Y., Zhang, J., & Wang, Y. (2018). Fusion of multisensor SSTs based on the spatiotemporal hierarchical Bayesian model. *Journal of Atmospheric and Oceanic Technology*, *35*(1), 91 - 109. doi: 10.1175/JTECH-D-17-0116.1

Zhu, Y., Kang, E., Bo, Y., Zhang, J., Wang, Y., & Tang, Q. (2019, 02). Hierarchical Bayesian model based on robust fixed rank filter for fusing MODIS SST and AMSR-E SST. *Photogrammetric Engineering Remote Sensing*, *85*, 119-131. doi: 10.14358/PERS.85.2.119

Zhu, Y., Kang, E. L., Bo, Y., Tang, Q., Cheng, J., & He, Y. (2015). A robust fixed rank kriging method for improving the spatial completeness and accuracy of satellite SST products. *IEEE Transactions on Geoscience and Remote Sensing*, *53*(9), 5021-5035. doi: 10.1109/TGRS.2015.2416351