

# A Multi-Fidelity Framework for Ocean Temperature Reconstruction Based on Model-Inferred Dynamics and Real Time Satellite and Buoy Measurements

Bianca Champenois<sup>1</sup> and Themistoklis Sapsis<sup>1</sup>

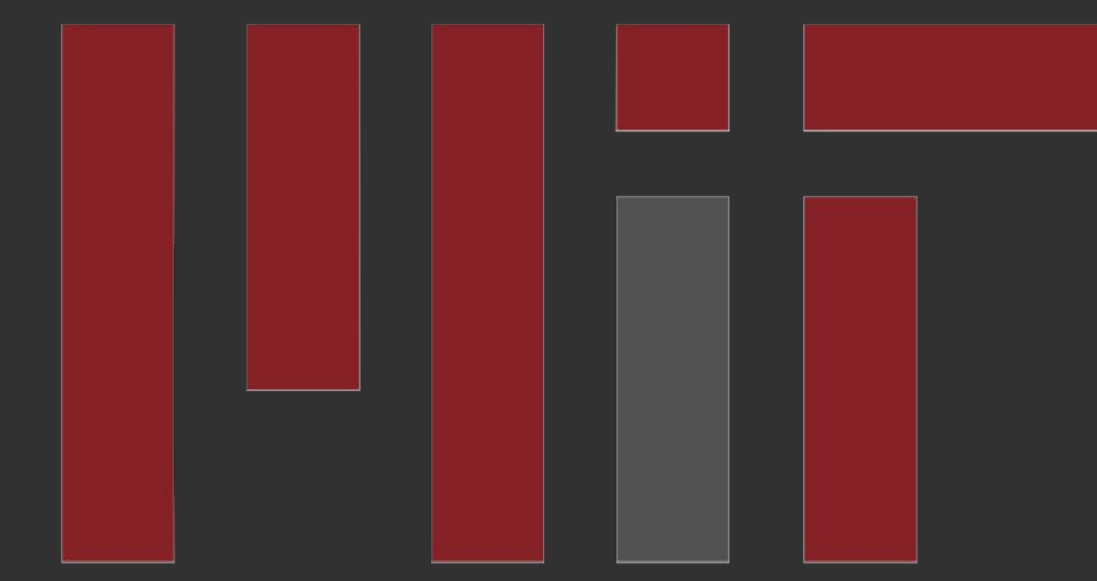
<sup>1</sup>Massachusetts Institute of Technology

November 21, 2022

## Abstract

Despite advancements in computational science, nonlinear geophysical processes still present important modeling challenges. Physical sensors (such as satellites, AUVs, or buoys) can collect data at specific points or regions, but these are often scarce or inaccurate. Here, we present a method to build improved spatio-temporal models that combine dynamics, inferred from high-fidelity numerical models (reanalysis data), and data from sensors. We are motivated by a data set of ocean temperature where sensor measurements are only available at the surface of the ocean. We first employ reanalysis data in the form of a 3D temperature field, and apply standard principal component analysis (PCA) at every ocean surface coordinate. For each coordinate, the vertical structure of the field can be represented with just two PCA modes and their corresponding time coefficients, significantly reducing the dimensionality of the data. Next, a conditionally Gaussian model, implemented through a temporal convolutional neural network, is built to predict the time coefficients of the PCA modes (i.e. vertical structure), as well as their variance, as a function of the surface temperature. These probabilistic predictions are made with the satellite data as input, and they are used with the PCA modes to stochastically reconstruct the full temperature field. The estimated temperature field is then combined with data from buoys through a multi-fidelity Gaussian process regression scheme, where the buoys have the highest fidelity and the satellite-based predictions have lower fidelity. The techniques described provide a framework for building less expensive and more accurate models of conditionally Gaussian estimates for full 3D fields, and they can be applied to geophysical systems where data from both sensors and numerical simulations are available. We implement these techniques to estimate the full 3D temperature field of the Massachusetts and Cape Cod Bay where temperature can serve as a useful indicator for ocean acidification. Finally, we discuss how the developed ideas can be leveraged to make more informed decisions about optimal in-situ sampling and path planning.





# A multi-fidelity framework for ocean temperature reconstruction based on model-inferred dynamics and real time satellite and buoy measurements

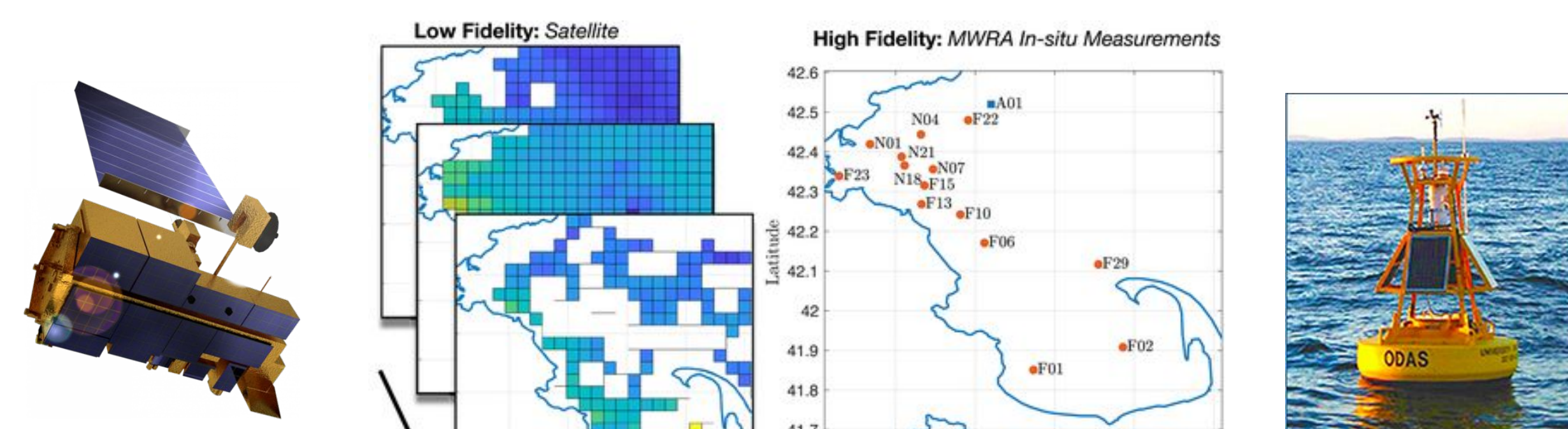


Bianca Champenois, Themistoklis Sapsis

## Objective

Estimate a full 3D temperature field of the ocean using a combination of numerical simulations and physical sensors.

## Background

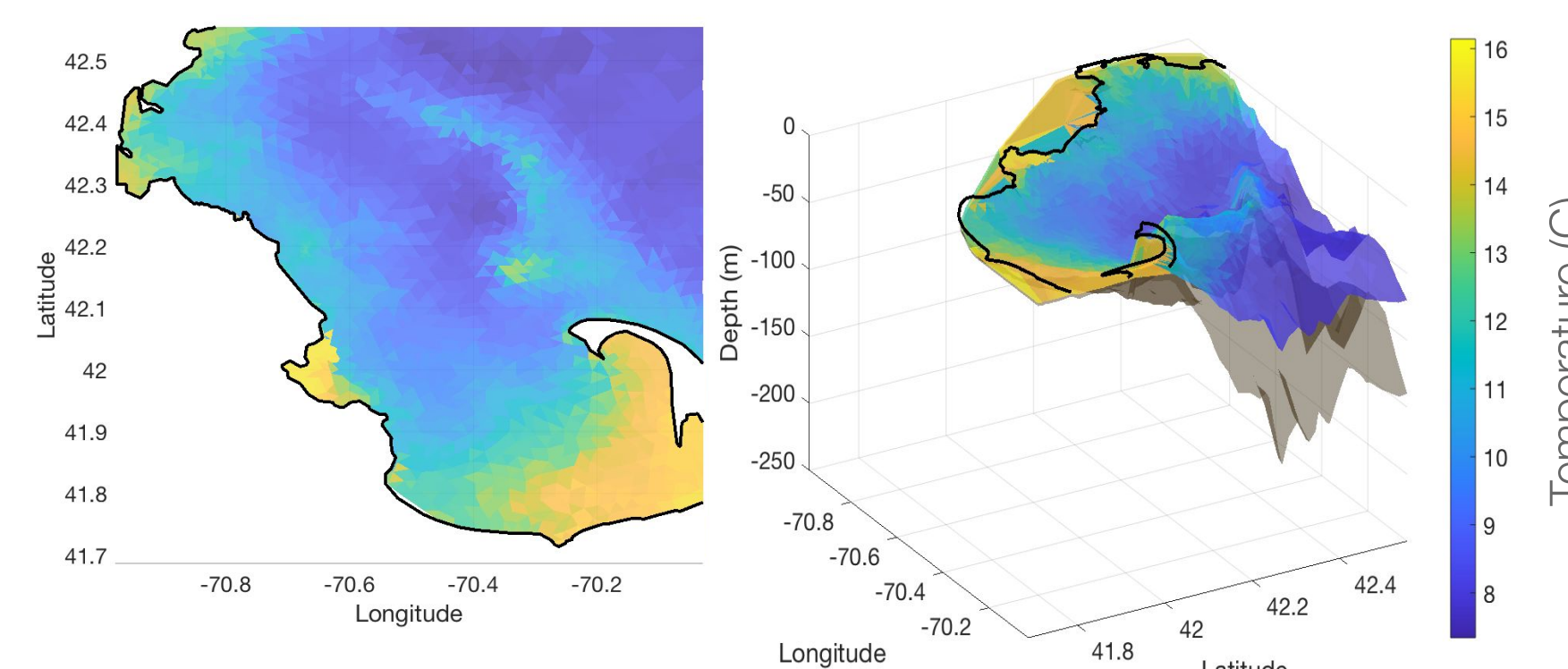


NASA Terra MODIS

MWRA NERACOOS Buoy

Babaei et al. 2020

$$\frac{\partial T}{\partial t} + u \frac{\partial T}{\partial x} + v \frac{\partial T}{\partial y} + w \frac{\partial T}{\partial z} = \frac{\partial}{\partial z} \left( K_h \frac{\partial T}{\partial z} \right) + F_T$$



FVCOM Reanalysis Data

Chen et al. 2003

## Techniques

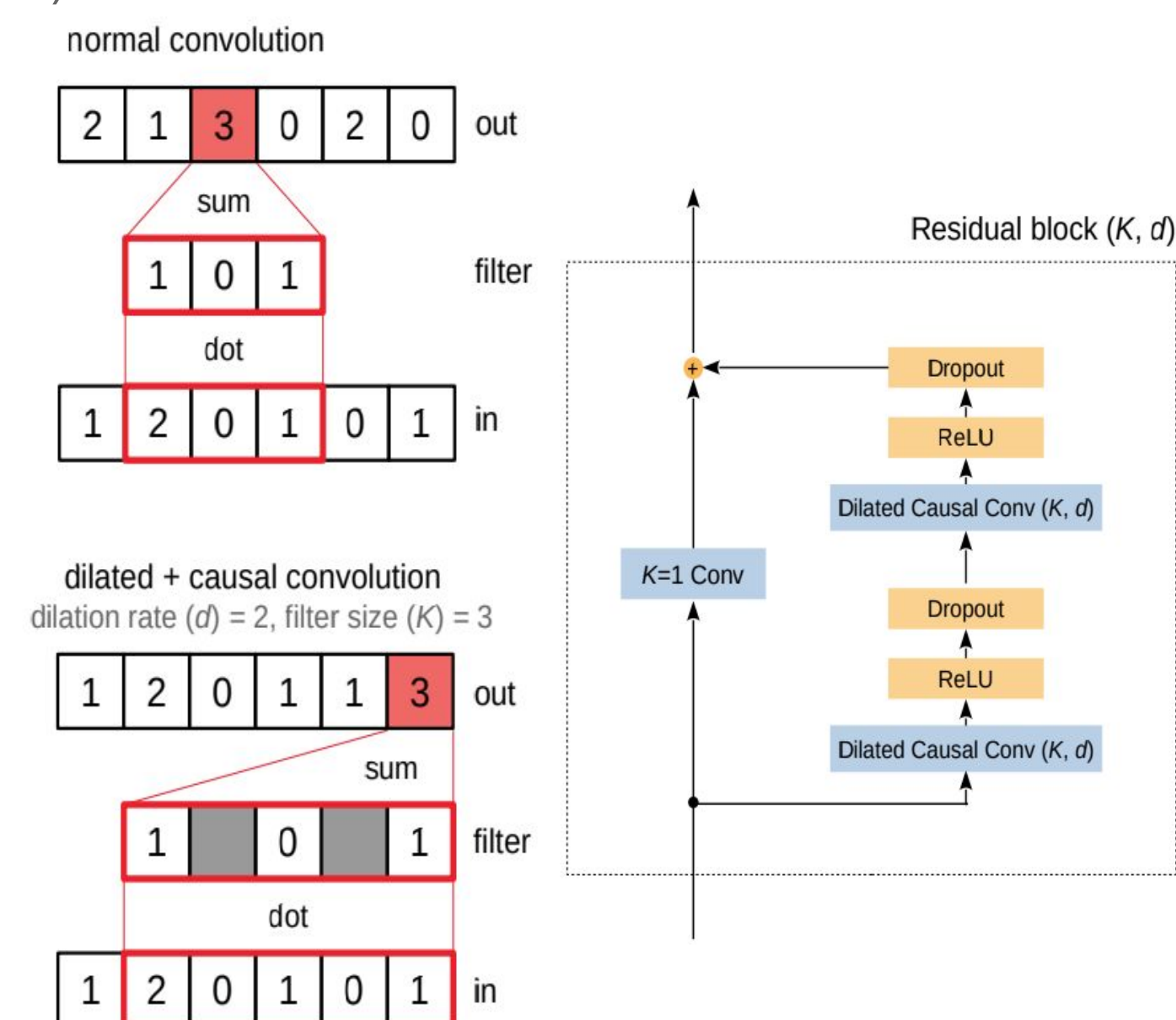
Principal Component Analysis (PCA)

$$\mathbf{T}_{proj}(\cdot, t) = \sum_{i=1}^2 q_i(t) \phi_i + \bar{\mathbf{T}}(t)$$

Gaussian Process Regression (GPR)

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

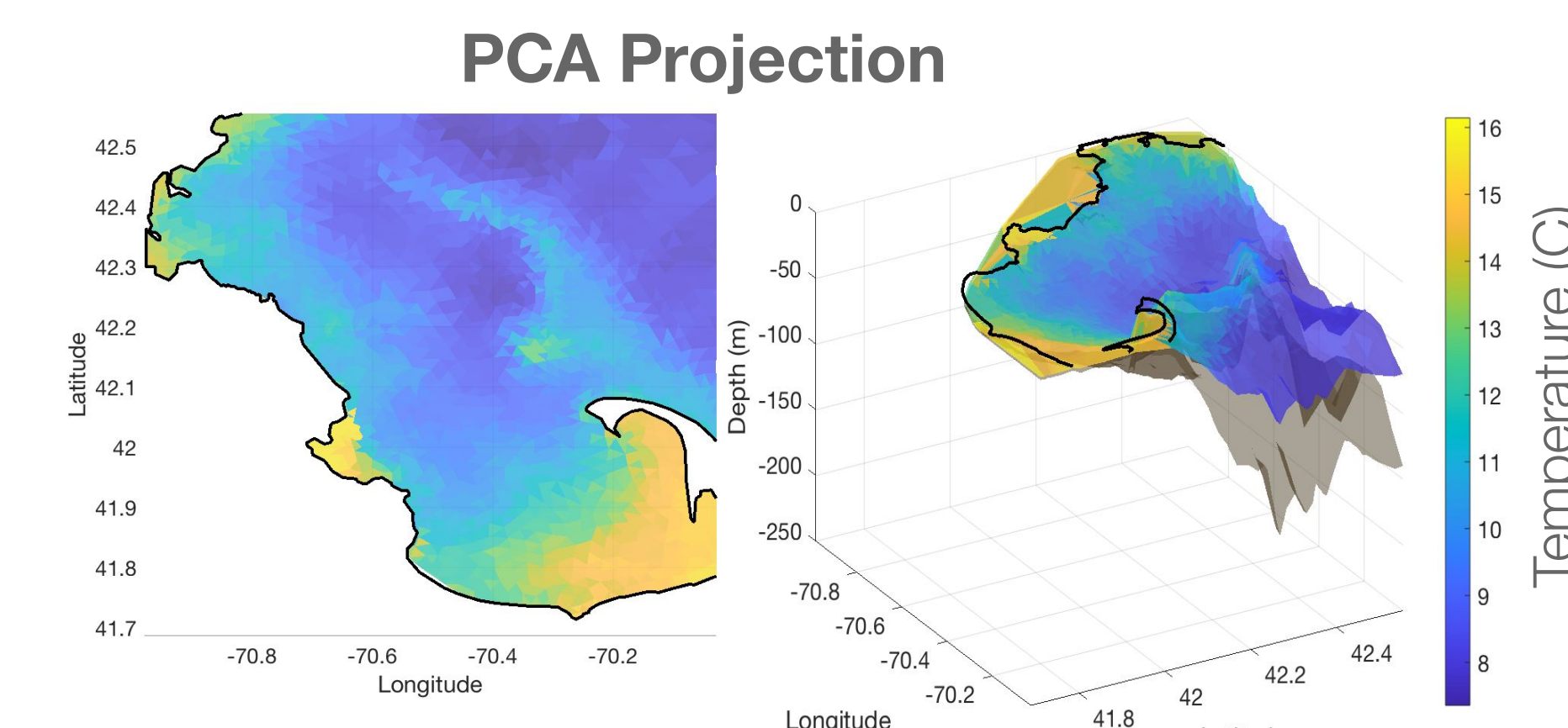
Temporal Convolutional Network (TCN)



Wan et al. 2020

## Framework

**Step 1:** Apply PCA to Each (x,y) Point of the FVCOM Data



$$\mathbf{T} = \begin{bmatrix} T(z_1, t_1) & \dots & T(z_1, t_m) \\ T(z_2, t_1) & \dots & T(z_2, t_m) \\ \vdots & \vdots & \vdots \\ T(z_n, t_1) & \dots & T(z_n, t_m) \end{bmatrix}$$

only keep the first two modes

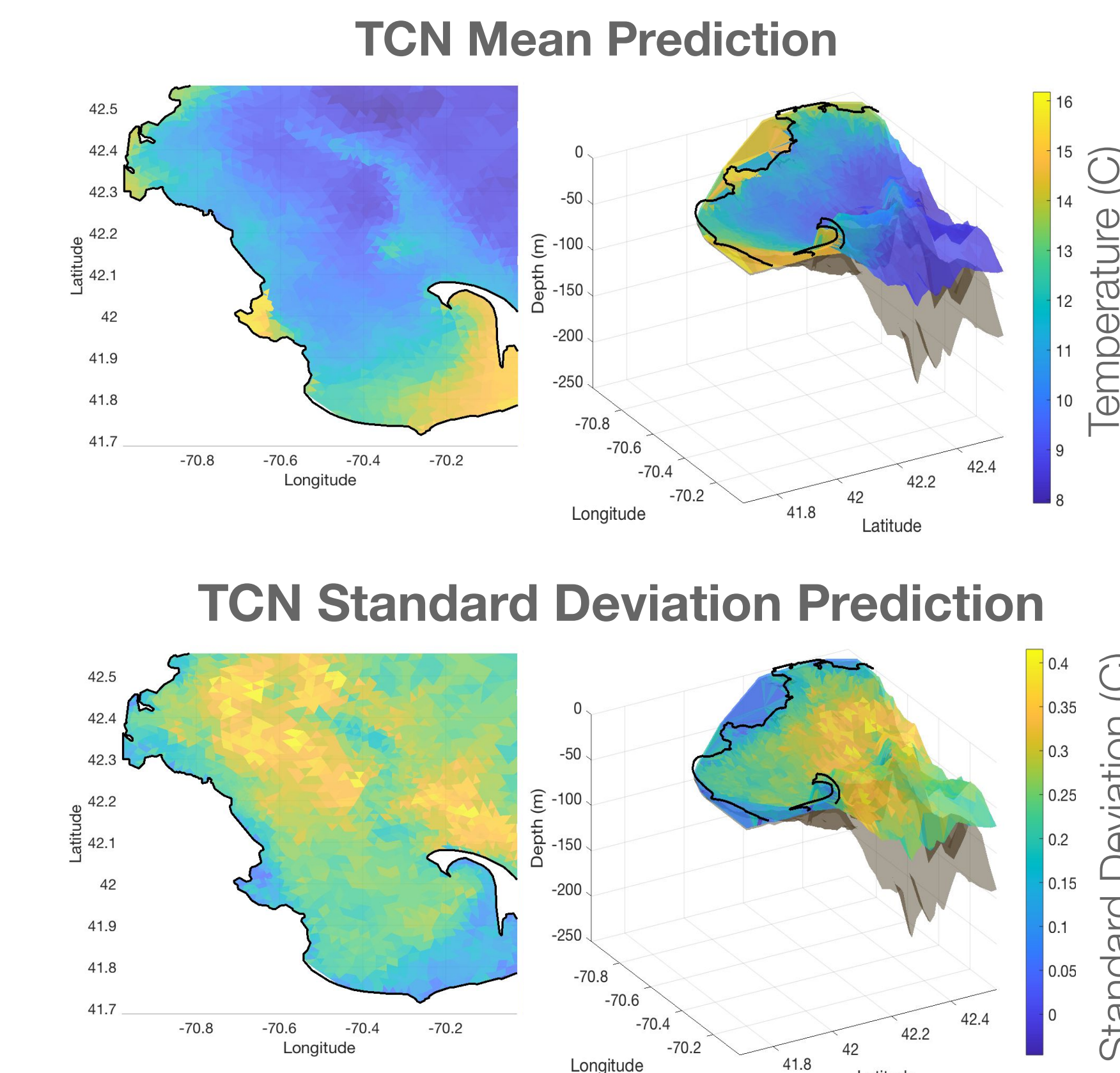
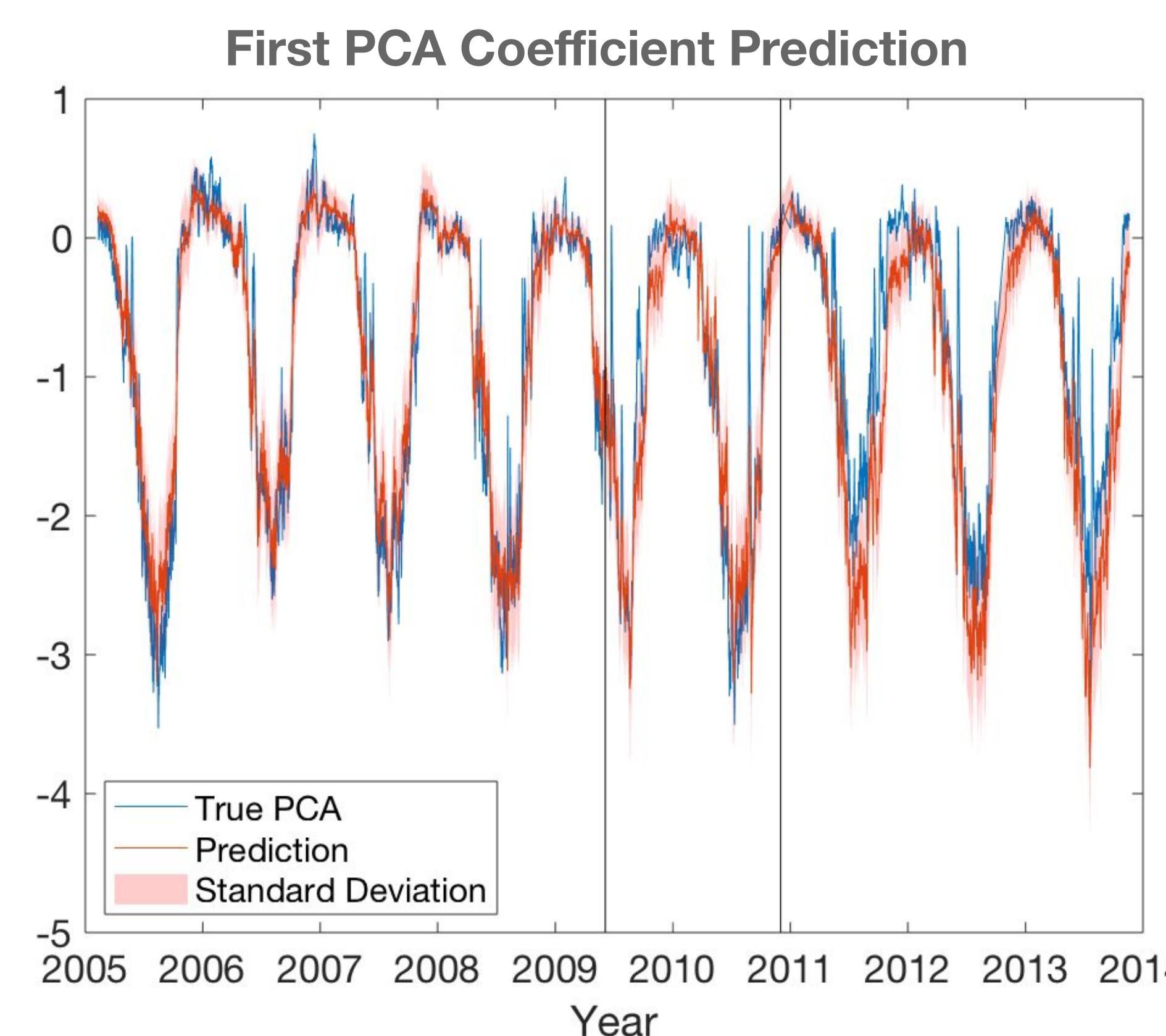
$$\mathbf{T}_{proj}(\cdot, t) = \sum_{i=1}^2 q_i(t) \phi_i + \bar{\mathbf{T}}(t)$$

PCA coefficients

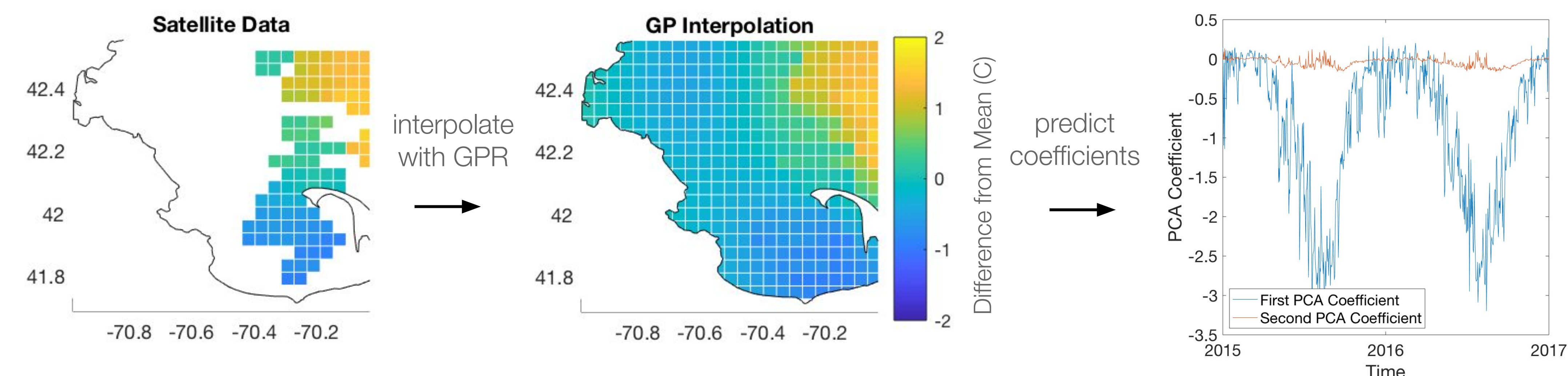
PCA modes

Holmes et al. 1996

**Step 2:** Build a TCN to Predict the PCA Coefficients as a Function of Surface Temperature



**Step 3:** Predict the PCA Coefficients from the Satellite Data to Reconstruct the 3D Field



**Step 4:** Build a Multi-Fidelity GPR Model Using the TCN Predictions and Buoy Data

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

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

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

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

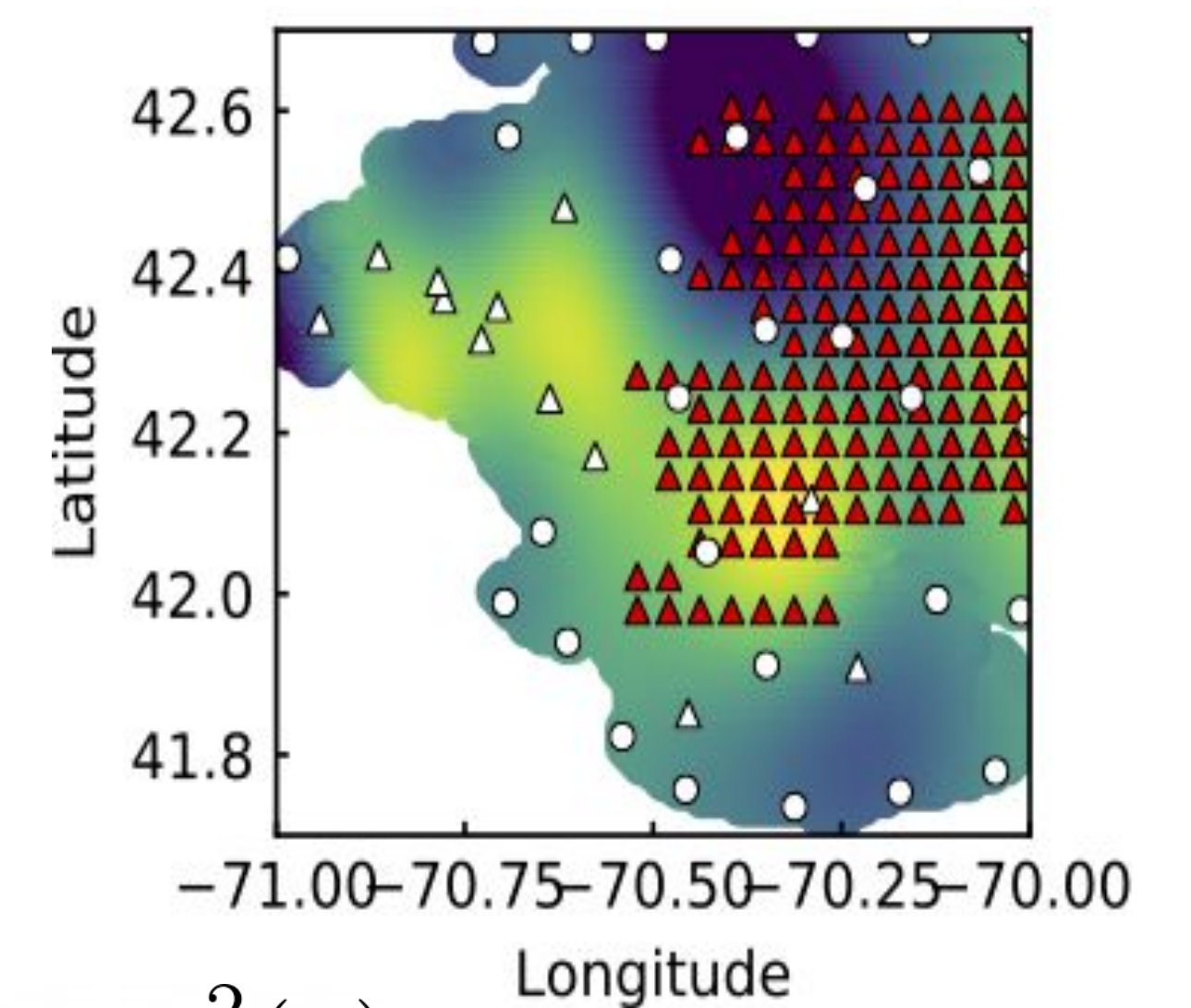
low fidelity

high fidelity

Rasmussen and Williams 2004

## Future Work

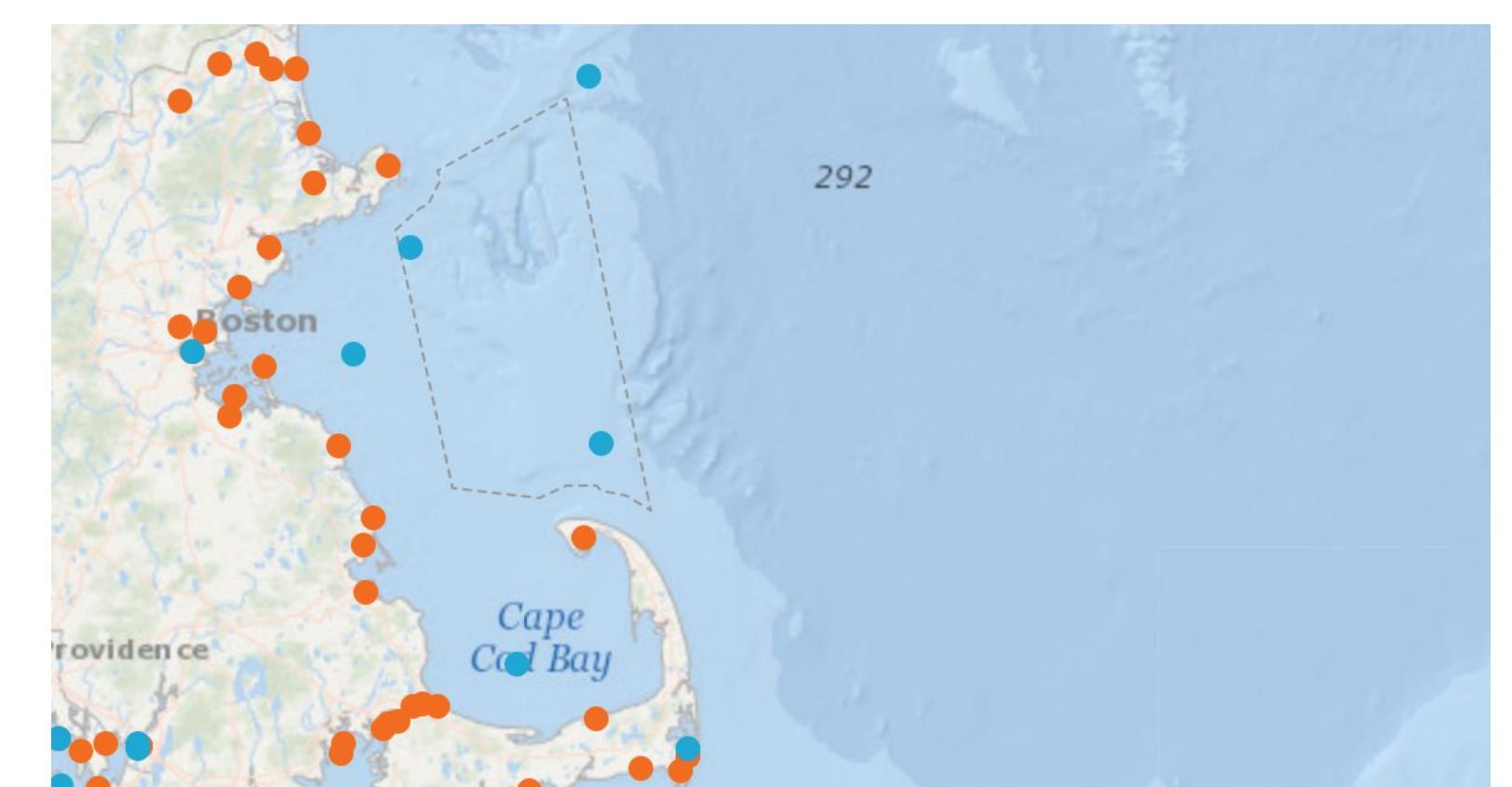
**Optimal Sampling and Path Planning**



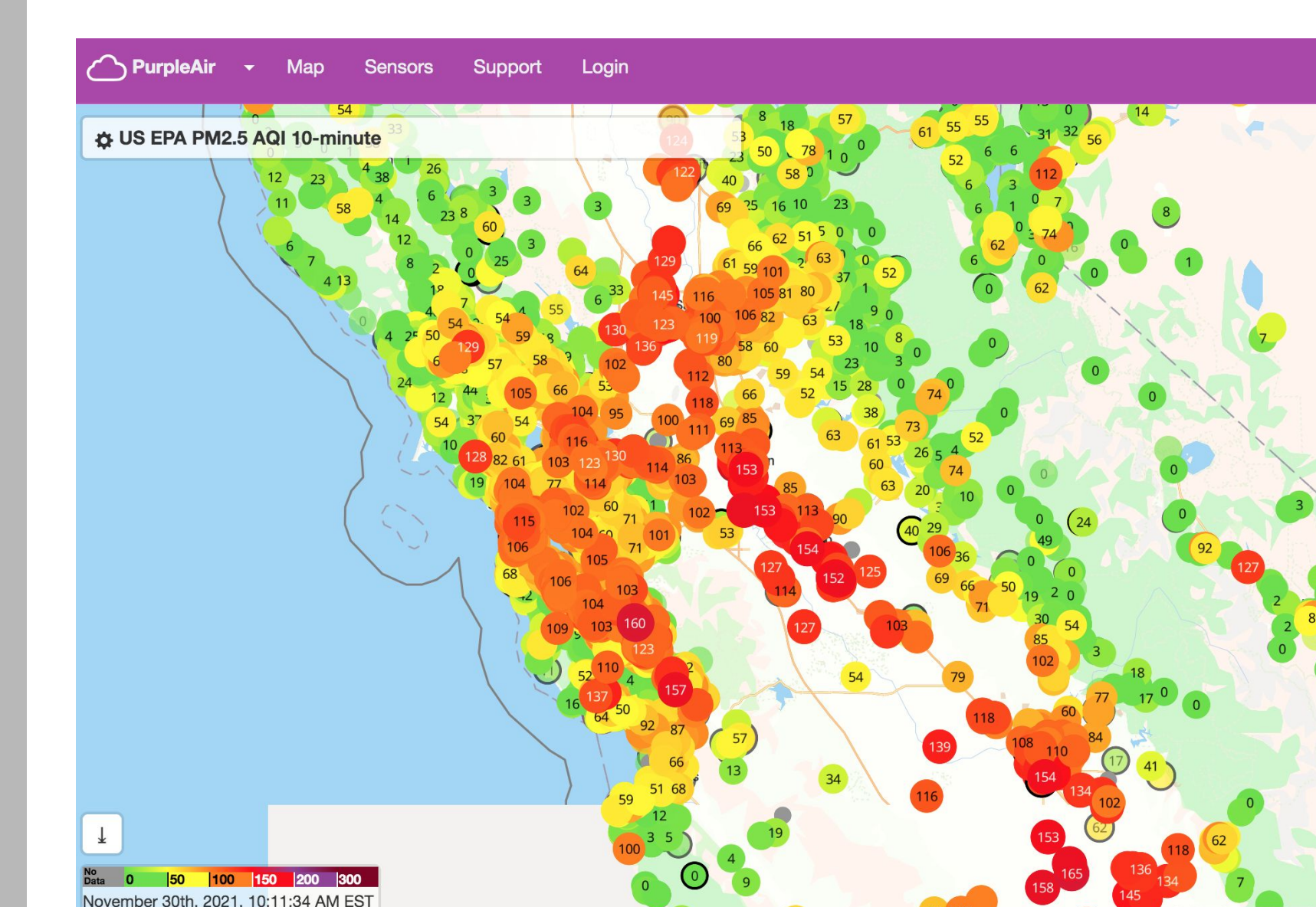
$$a_{US}(x) = \sigma^2(x)$$

Blanchard and Sapsis 2020

**Incorporate Results Into SEAGLASS**



## Other Applications



## Acknowledgements

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. Thank you to Dr. Carolina Bastidas and Michael Defilippo for the data.