

A Probabilistic Modeling Framework for Forecasting Gas Hydrate Using Geospatial Machine Learning

Jennifer M. Frederick¹, William Eymold^{1,2}, Michael Nole¹, Warren Wood³, Ben Phrampus³, Hugh Daigle⁴

1 Sandia National Laboratories
2 The Ohio State University
3 U.S. Naval Research Laboratory
4 University of Texas at Austin

Contact Information:

Jennifer M. Frederick
Sandia National Laboratories
email: jmfrede@sandia.gov
phone: 1.505.313.3233

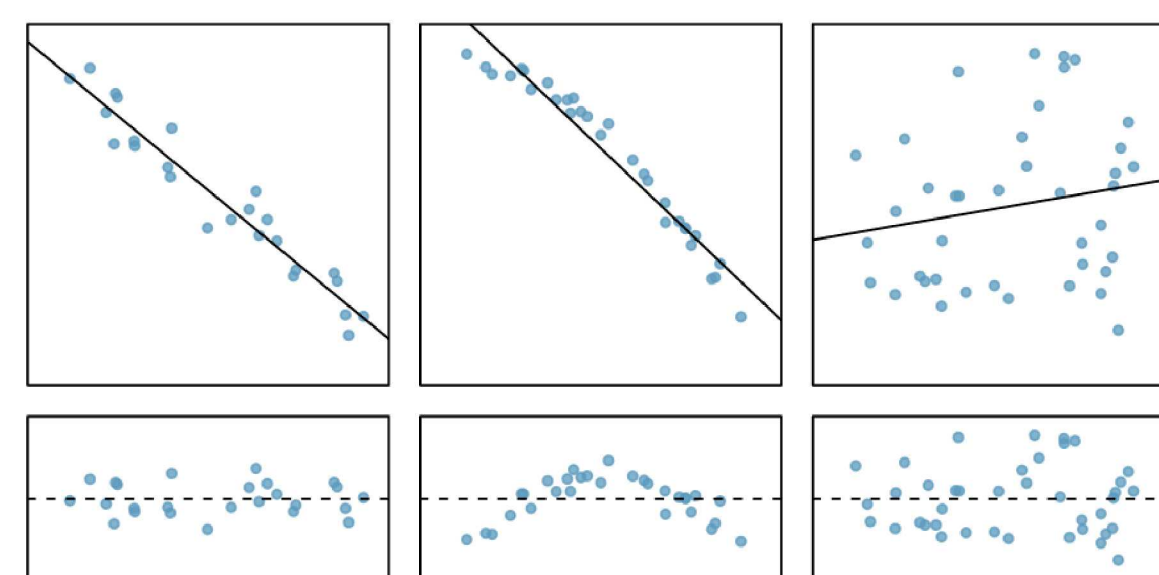
Geospatial machine learning (GML) can be used to build a system that predicts or forecasts seafloor properties like we forecast the weather. GML can produce maps of continuous seafloor properties with estimates of uncertainty, while also integrating physically consistent models. It is superior to traditional interpolation methods.

Global Observations (data)



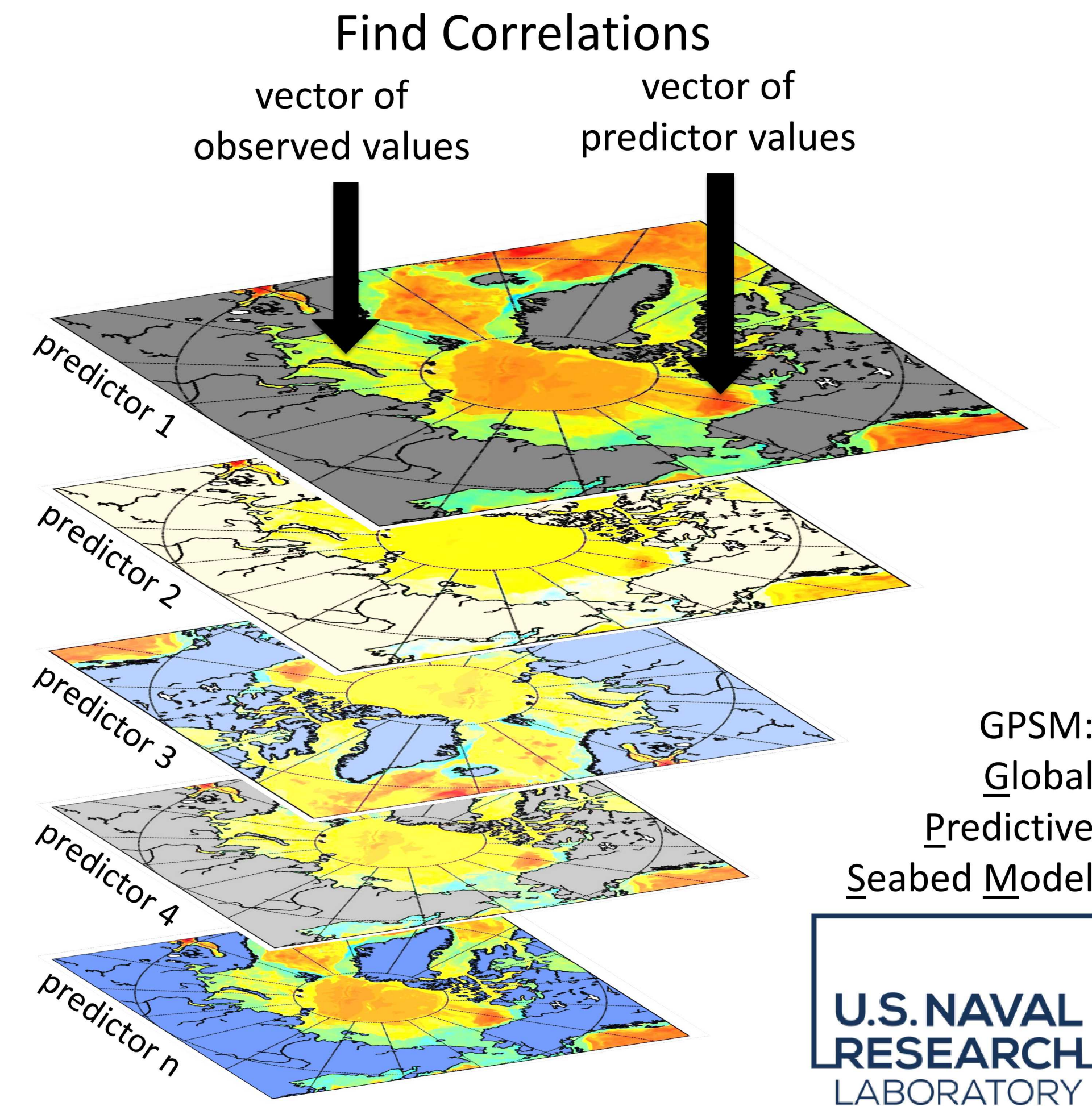
Collect and use all known data on seafloor, organized as a gridded dataset. Data outside of the Arctic can and should be used!

Feature Selection & Validation

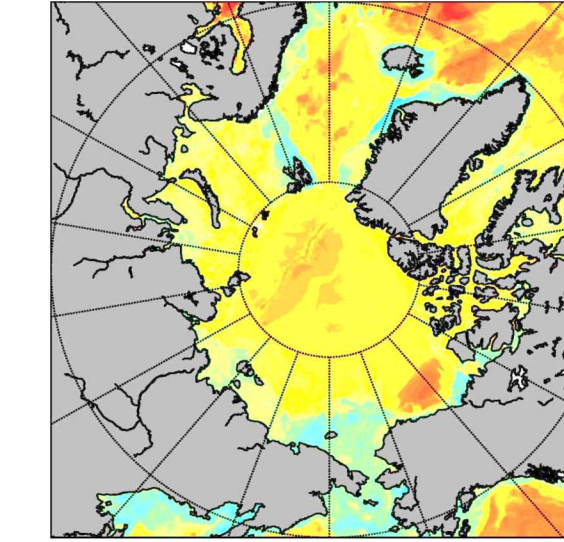


Only use the best predictors, based on individual predictive skill via 10-fold validation. Predictors must perform better than random noise.

Geospatial Machine Learning Algorithm

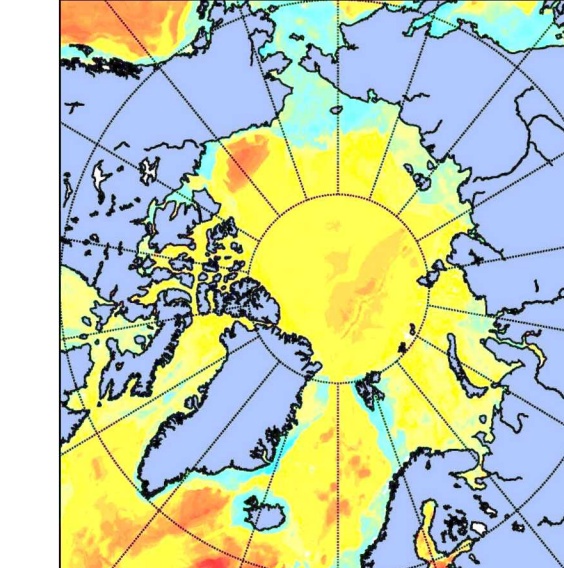


Forecast



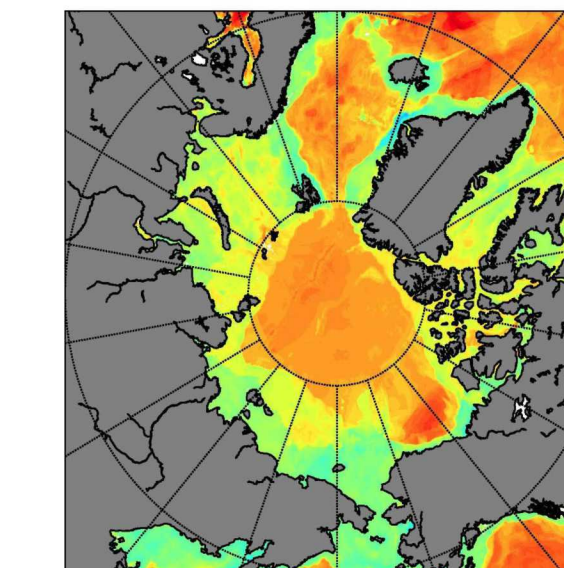
Based on sparse known data, and hundreds of dense calculated predictors, GML produces continuous maps of desired seafloor quantities, such as porosity, sediment type, total organic carbon content, etc.

Uncertainty



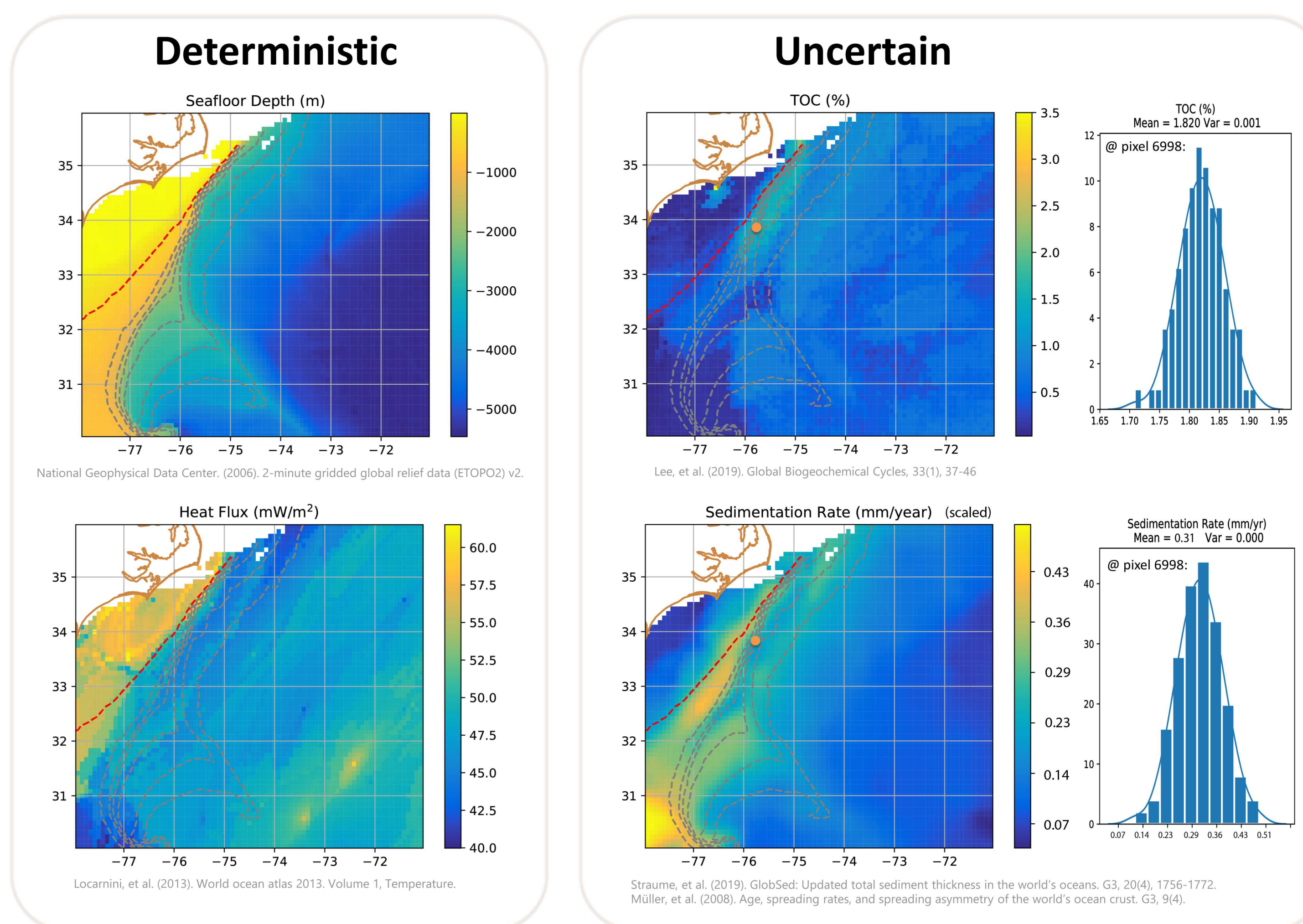
GML produces estimates of seafloor quantities and their uncertainty, which is based on prediction error. A well sampled parameter space will reduce parameter uncertainty.

Guide



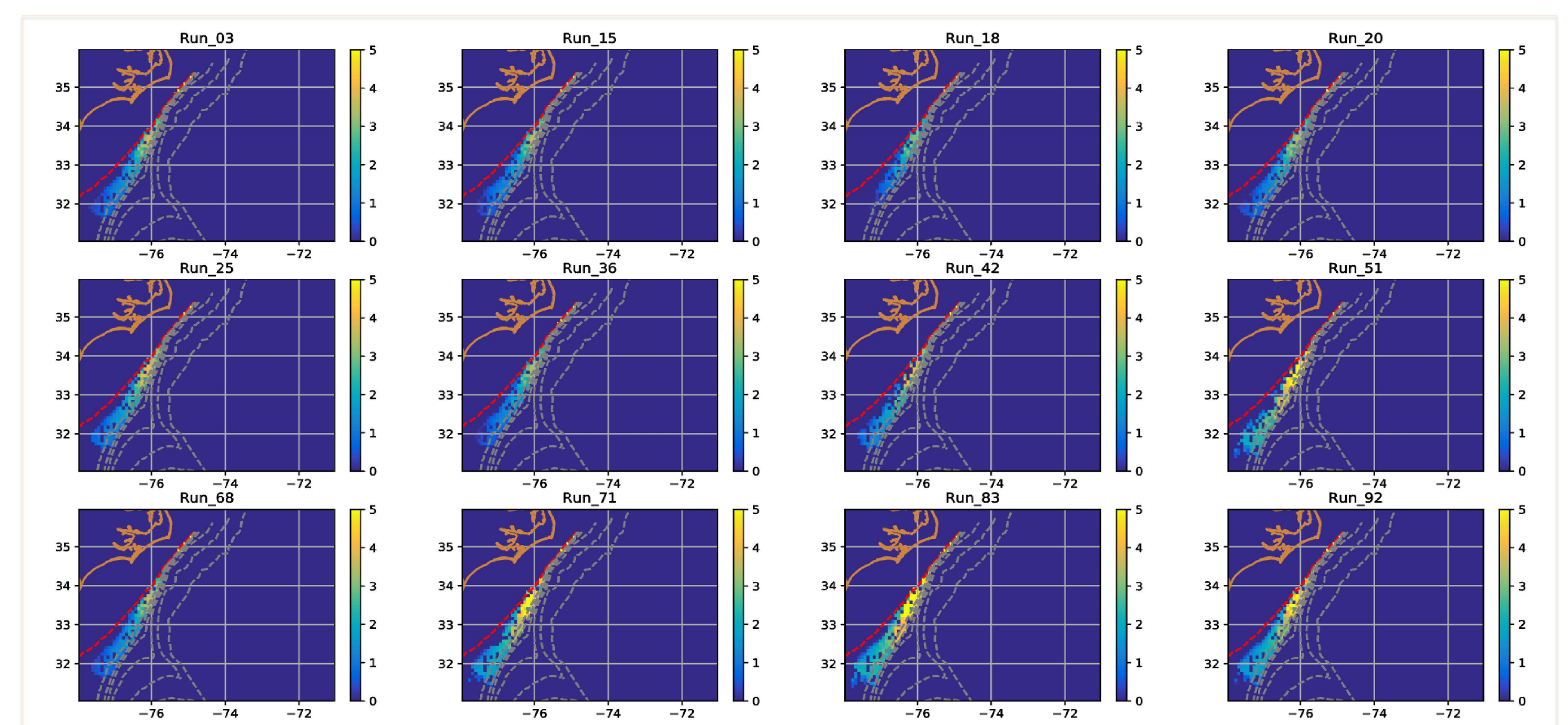
Uncertainty results can be used to guide future data acquisition campaigns. Increasing observations where prediction error (uncertainty) is high will benefit predictive skill globally.

GPSM Maps of Continuous Input Parameters



A Demonstration at Blake Ridge Gas Hydrate Province

Produce Ensemble Simulation Results By Propagating Uncertain Parameter Inputs
Sampling on the GPSM-derived input parameters generates numerous realizations (100 in this demonstration) for gas hydrate saturation at the Blake Ridge (after 500,000 years). Uncertain parameters are assumed to be normally distributed with a mean value and standard deviation.

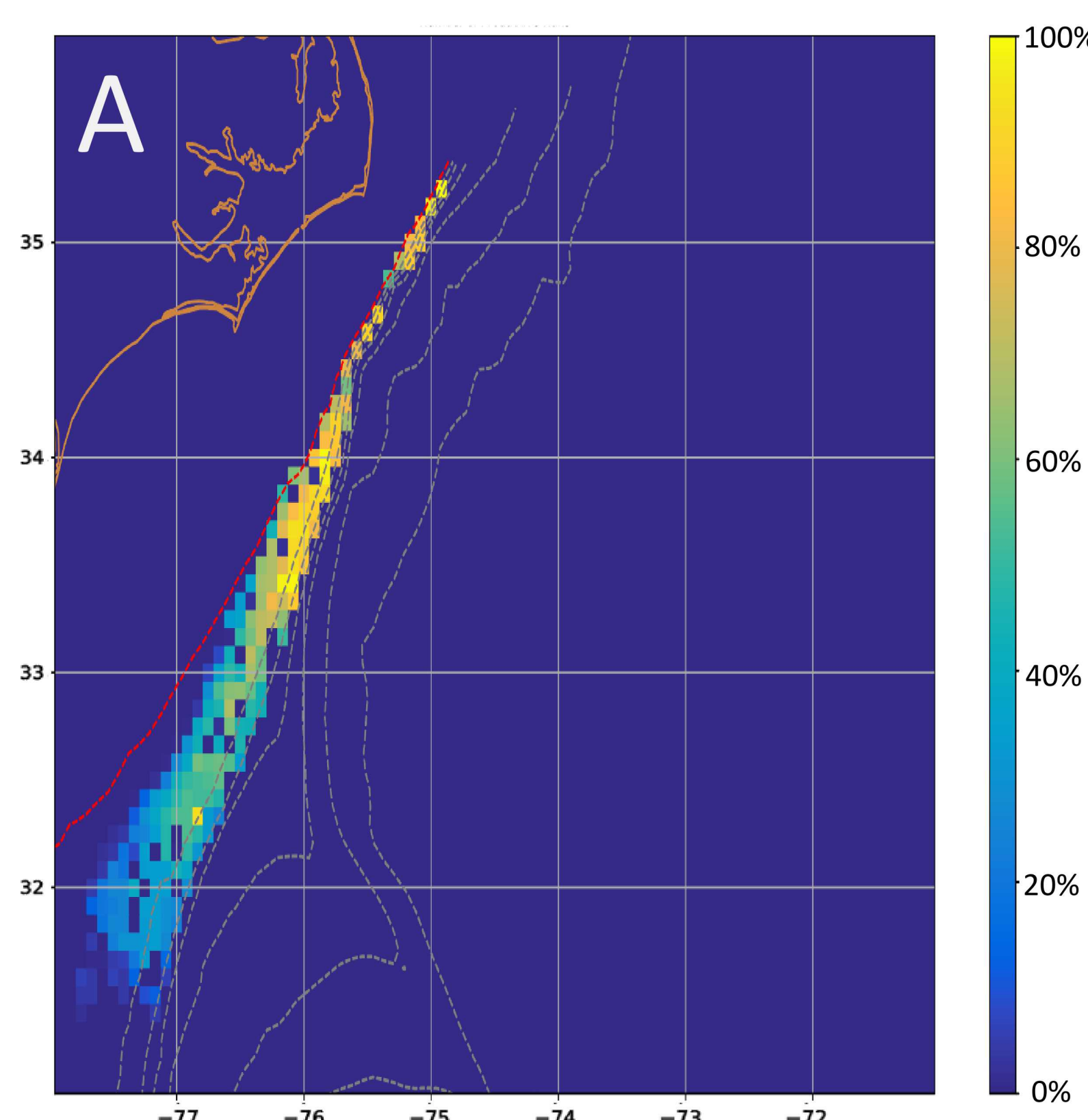


PFLOTRAN

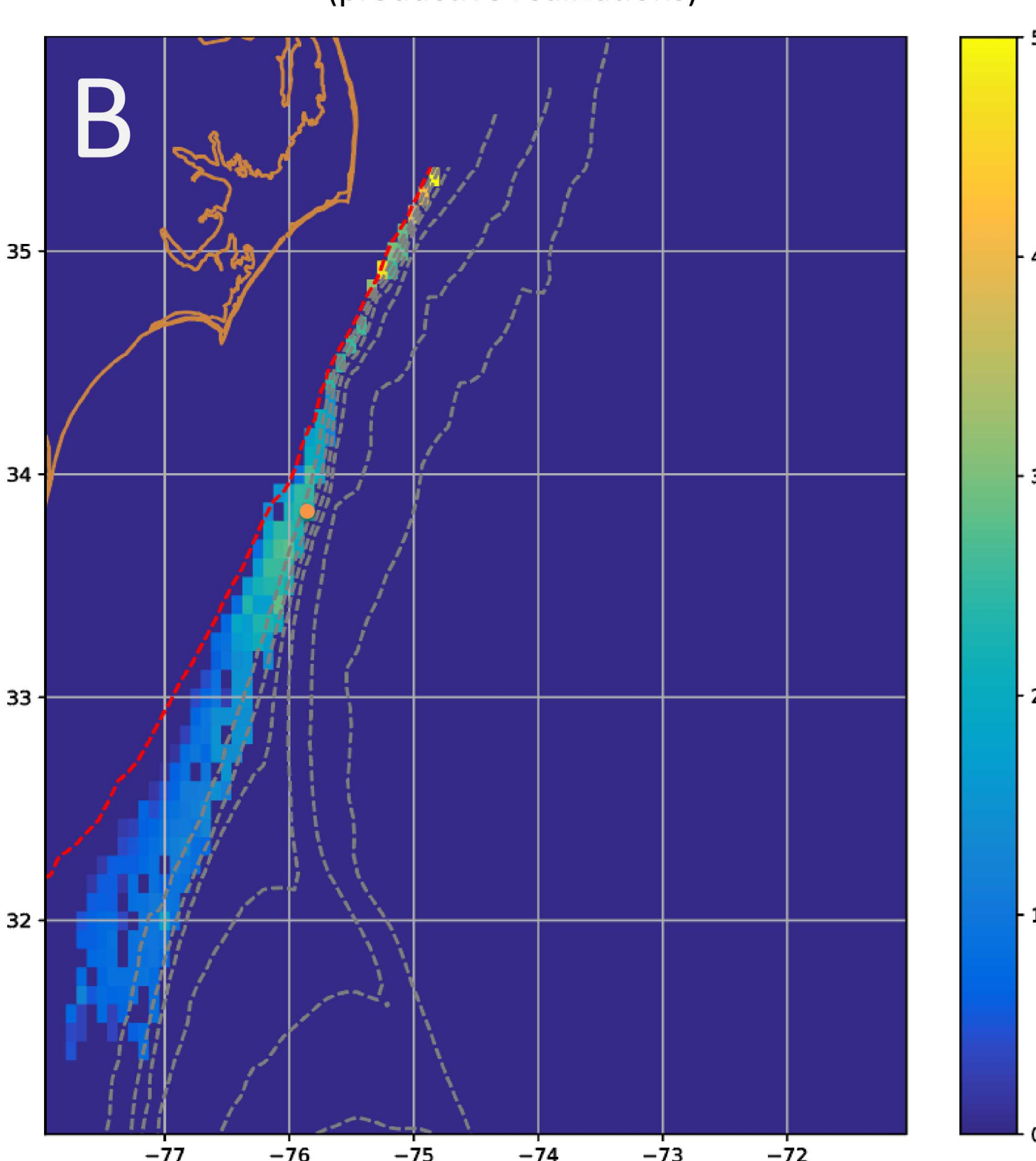
Ensemble simulation results are processed into a single probabilistic forecast

The ensemble results of interest are compiled into a histogram. Assuming a normal distribution, the mean and variance are calculated from the histogram to create a probabilistic map.

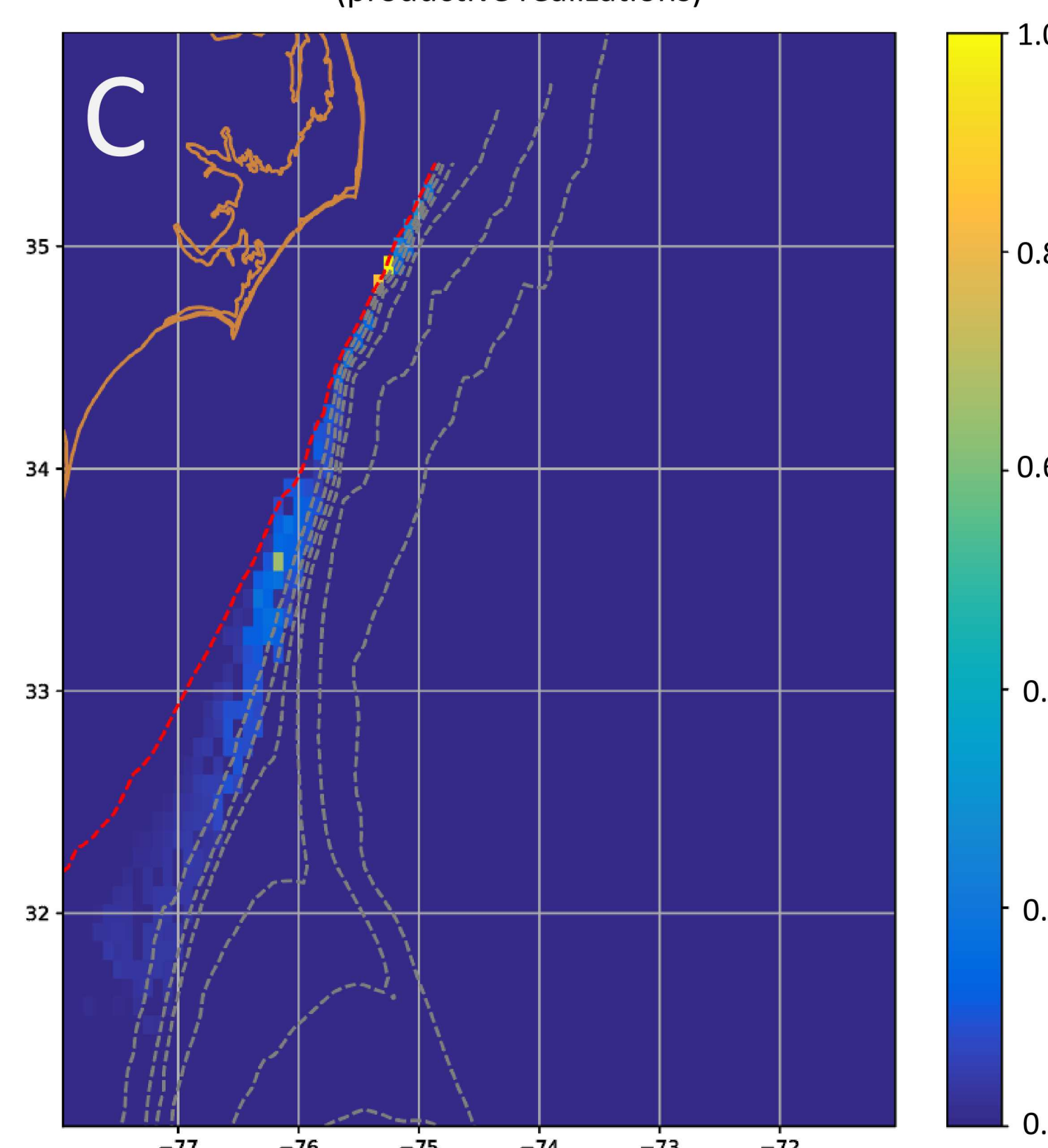
Percent Productive Realizations



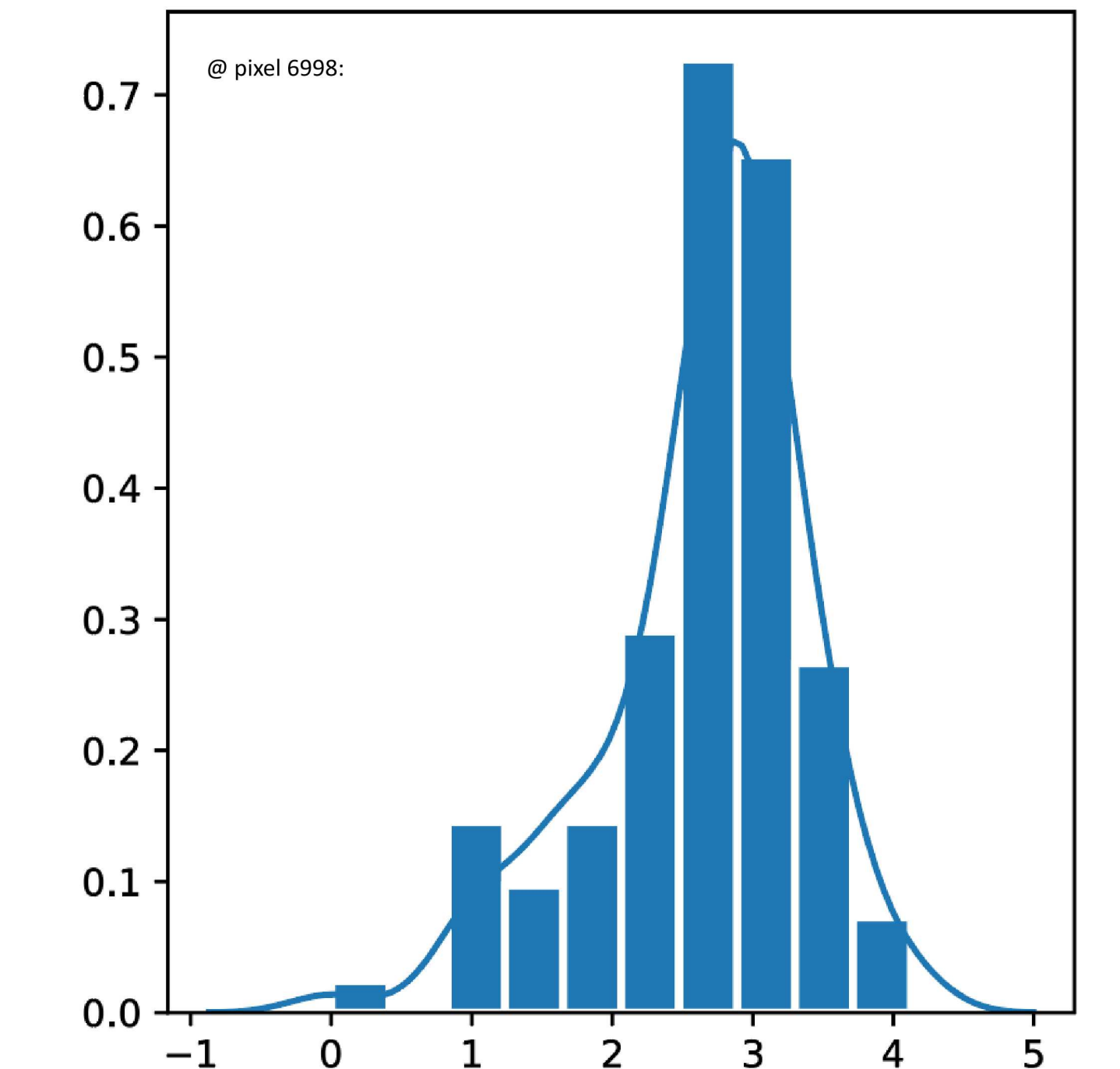
Mean Max. Gas Hydrate Saturation (productive realizations)



Variance of Gas Hydrate Saturation (productive realizations)



Maximum Hydrate Saturation (%)
Mean = 2.645 Var = 0.530



From the number of realizations that produced gas hydrate (Figure A), the mean maximum gas hydrate saturation is shown (Figure B) and the variance of the mean maximum gas hydrate saturation (Figure C). Maps like these can be used to predict seabed sound propagation or quantify gas hydrate resource potential.