



SAND2020-2044C

A hybrid machine learning and reservoir simulation approach to estimating global gas hydrate occurrence using open-source software



PRESENTED BY

Michael Nole



Acknowledgements

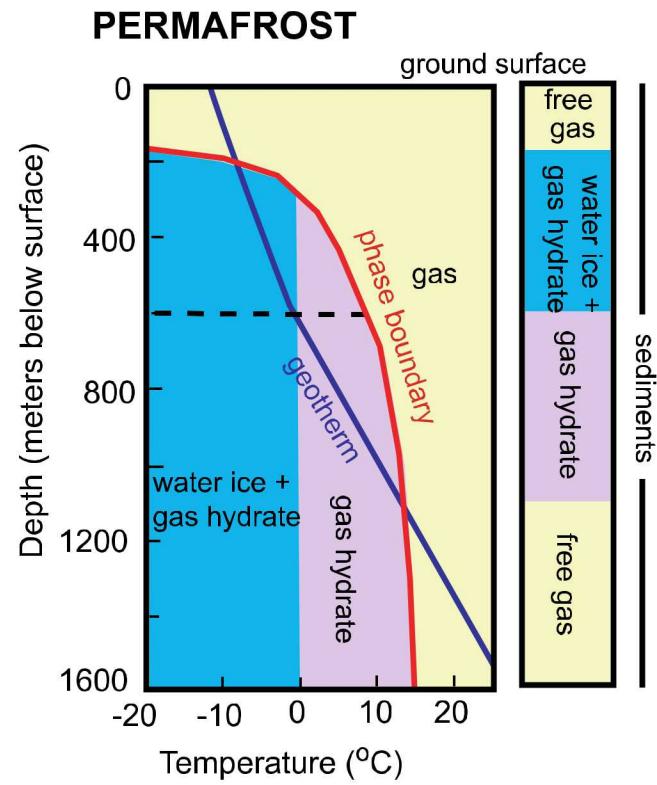
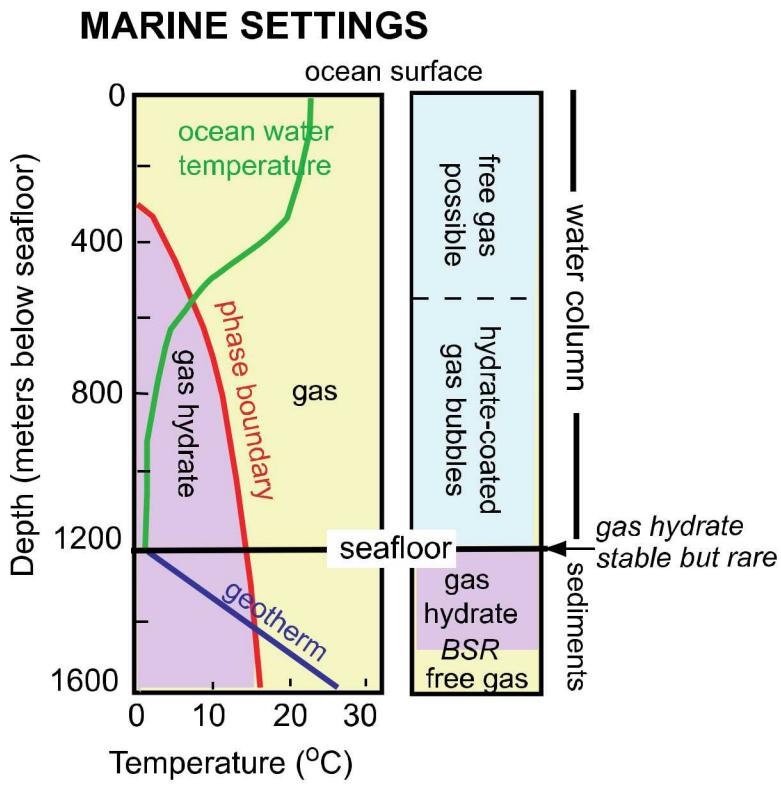
- Billy Eymold, The Ohio State University/Sandia National Labs
- Taylor Lee, Naval Research Laboratory
- Ben Phrampus, Naval Research Laboratory
- Warren Wood, Naval Research Laboratory
- Jenn Frederick, Sandia National Labs



Key Points

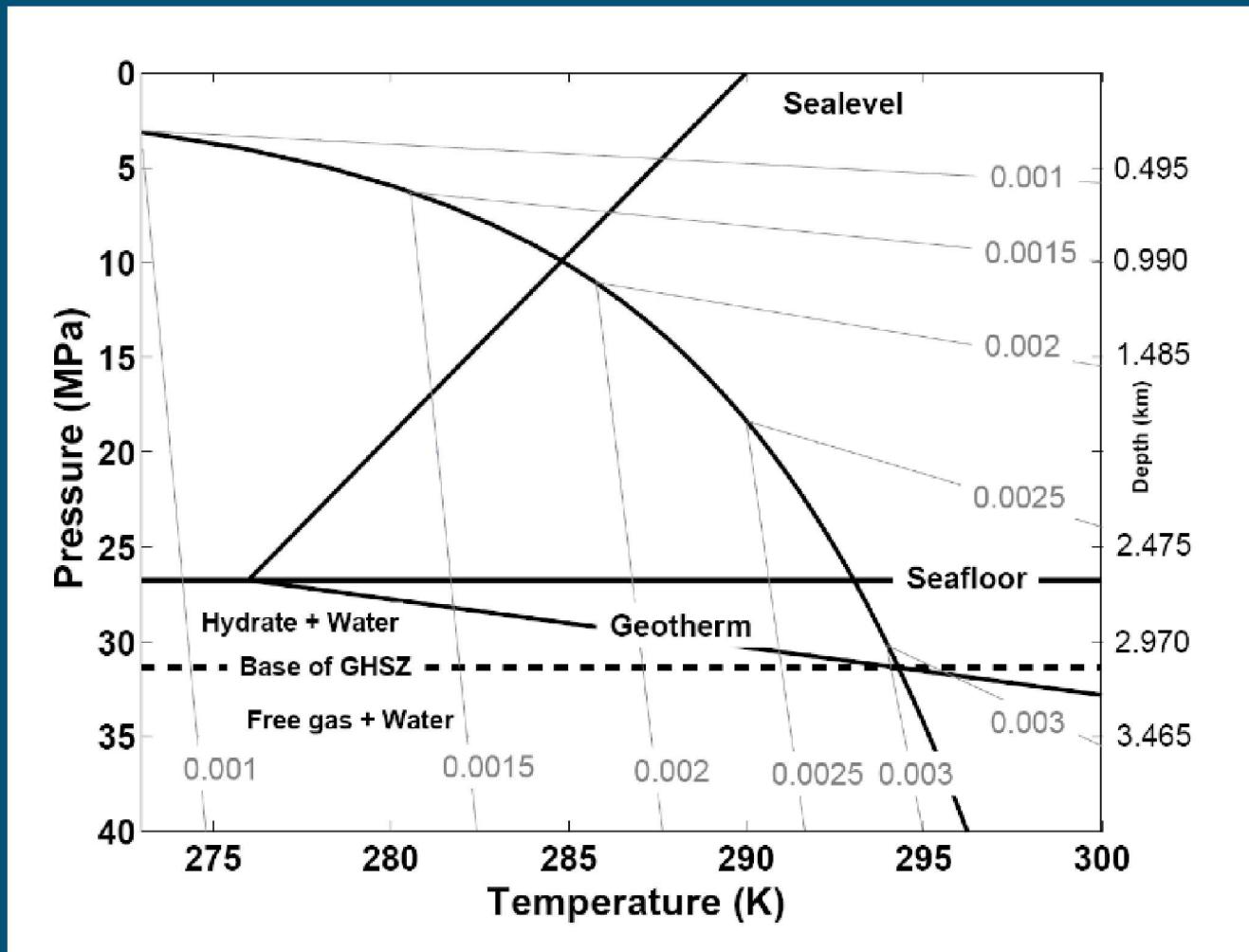
- The expression of gas hydrate in nature is controlled by complex interactions between gas sources, migration mechanisms, hydrate reservoirs, and (thermodynamic) seals
- Forward modeling these systems often requires:
 - Multi-phase, multi-component simulation
 - 3D capability to honor lateral heterogeneity and quantify competing influences of different migration mechanisms
- Predicting continuous distributions of inputs to a simulator is challenging when operating data-limited
 - How do we overcome this challenge, and can we quantify associated uncertainties?
- We can use machine learning to estimate the inputs to a simulator based off of proximity in geologic predictor space instead of simply spatial proximity

Thermodynamics



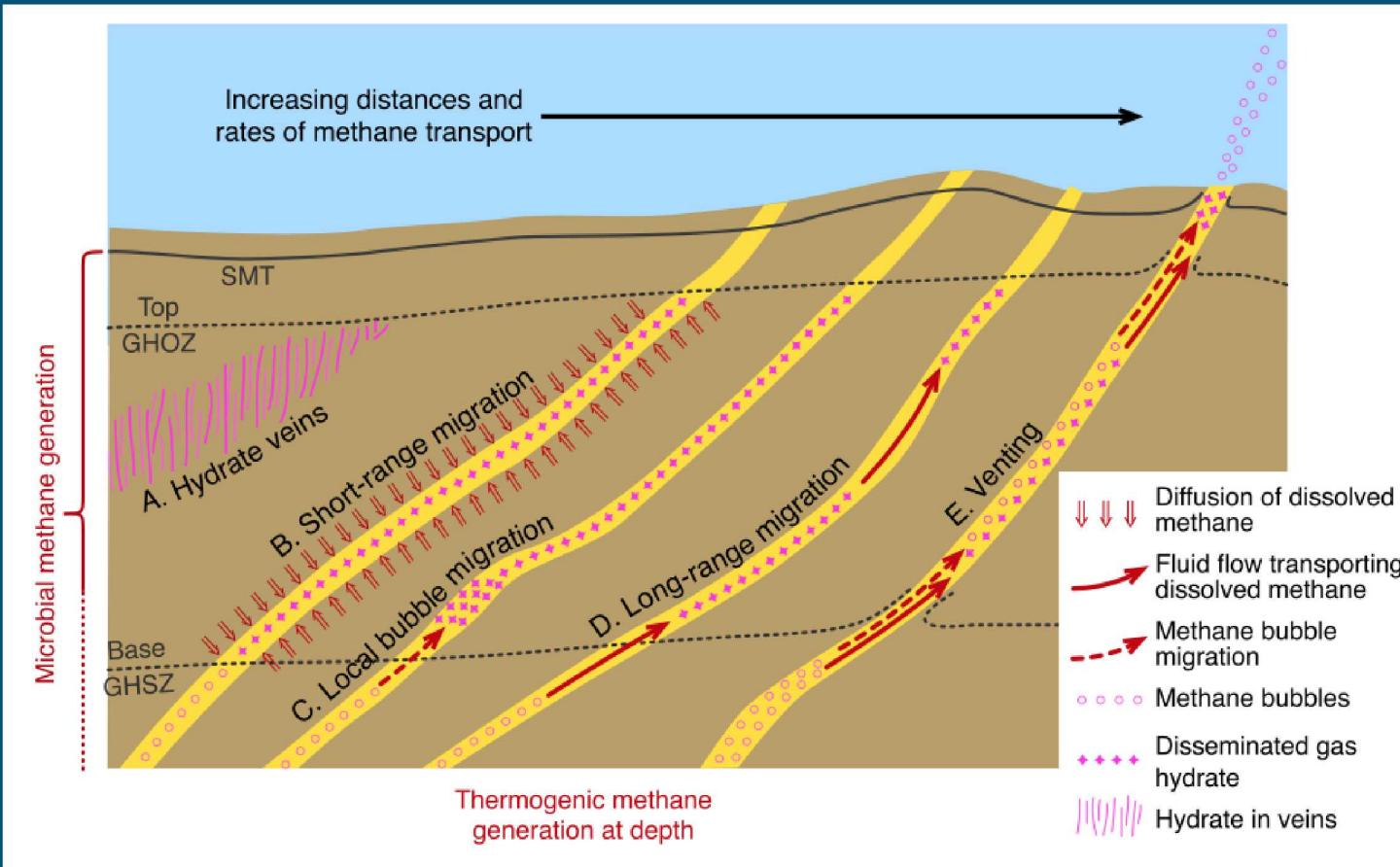
Ruppel, 2007

Methane Availability



Bhatnagar et al., 2007

Transport



Motivation

Sandia National Laboratories

U.S. NAVAL RESEARCH LABORATORY

TEXAS
The University of Texas at Austin

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

Forecasting Marine Sediment Properties On and Near the Arctic Shelf with Geospatial Machine Learning

Jennifer M. Frederick¹, Warren Wood², Michael Nole¹, Ben Phrampus², Hugh Daigle³, Hongku Yoon¹, Brian Young¹, and Ken Sale¹

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!

Machine Learning Data Flow

Only use the best predictors, based on individual predictions via 10-fold validation. Predictors must conform to the machine learning algorithm.

Geospatial Machine Learning Algorithm

Find Correlations

vector of observed values vector of predictor values

Geospatial Predictions

Based on various observed data and hundreds of different machine learning predictors, GML produces continuous maps of various seafloor properties, such as porosity, sediment types, total organic carbon content, etc.

Geospatial Validation

GML produces validation of seafloor properties and their uncertainty, which is then used in prediction maps. A well-sampled parameter space will reduce uncertainty.

Integrate

Integrate physical models to produce predictions of seafloor geo-acoustic and geo-mechanical properties.

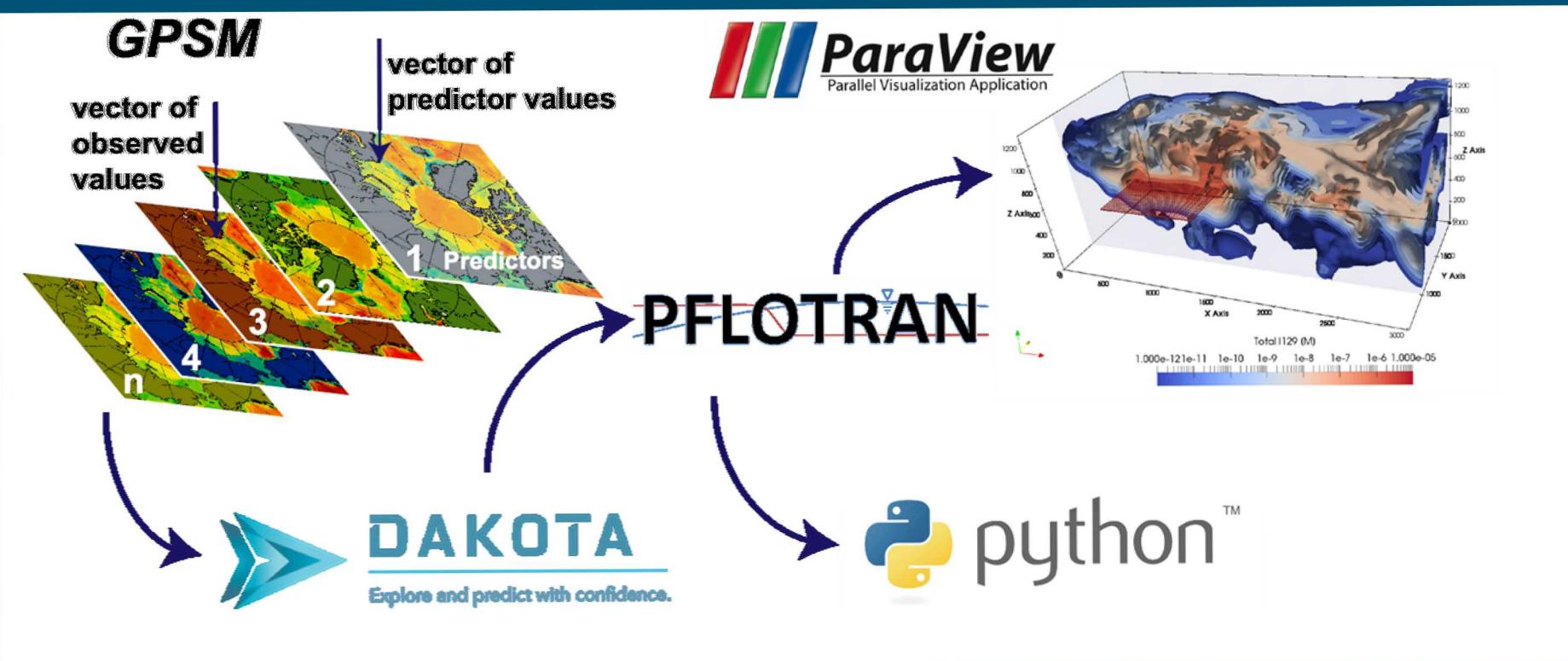
Integrate physical models using GML generated seafloor parameters to map geo-acoustic and geo-mechanical sediment properties.

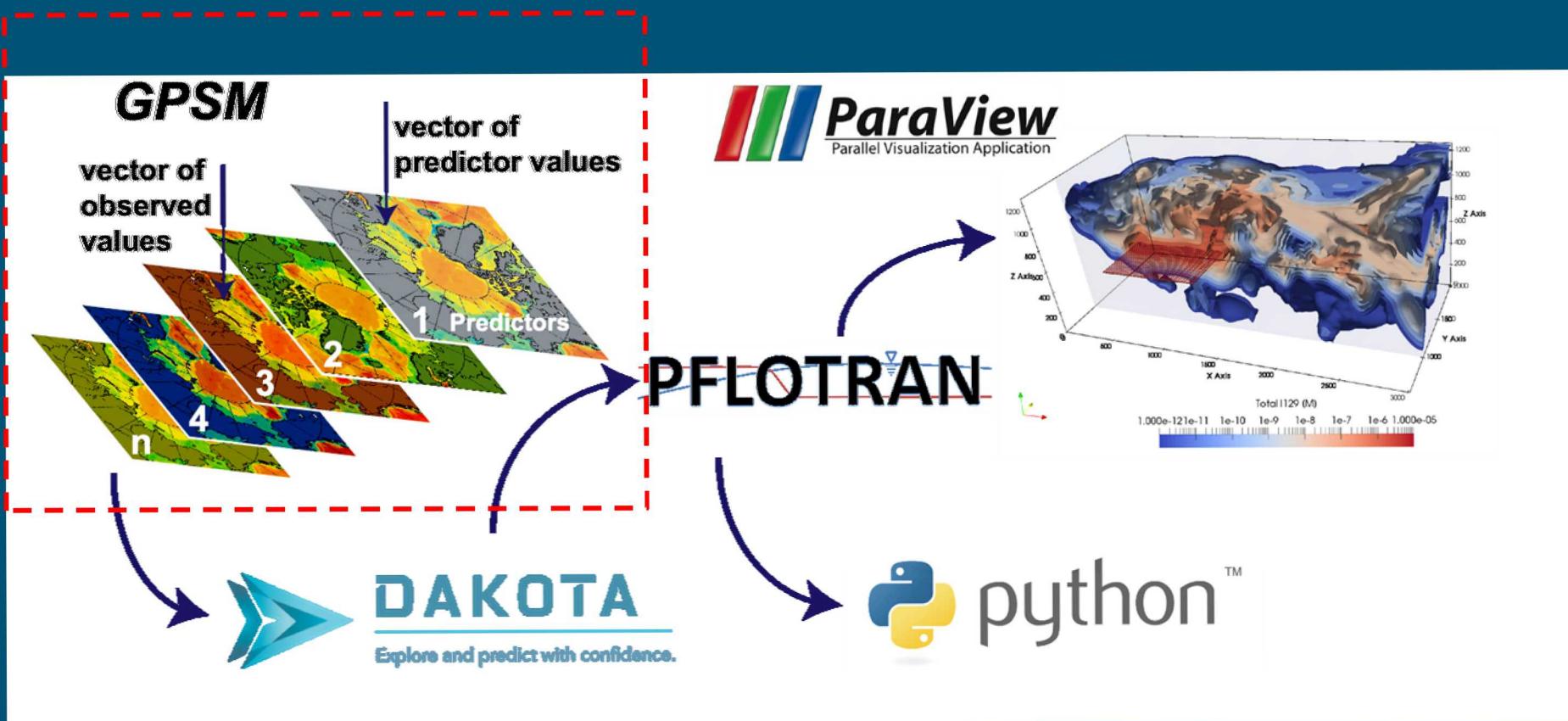
Physical estimation of sediment grain size and thermal properties will help map methane gas phase dynamics.

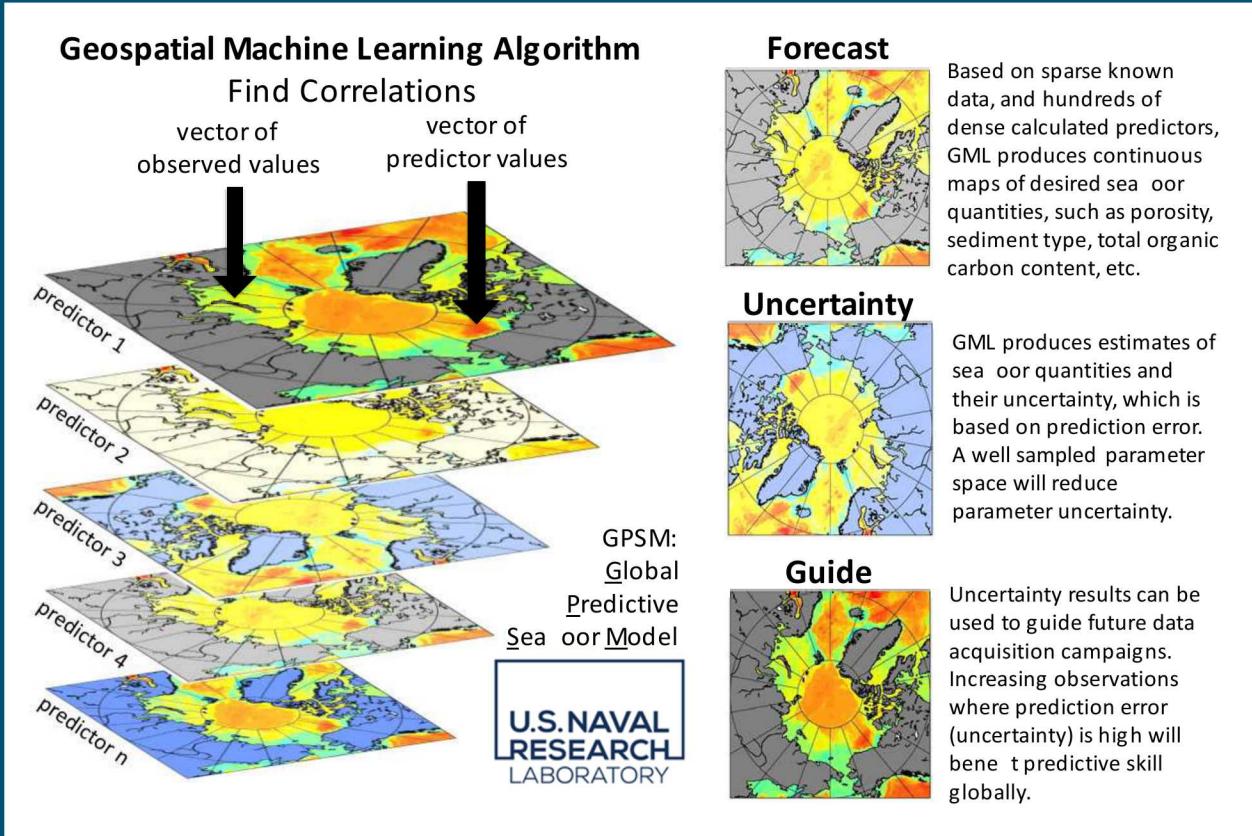
GPSPM: Global Predictive Seafloor Model

U.S. NAVAL RESEARCH LABORATORY

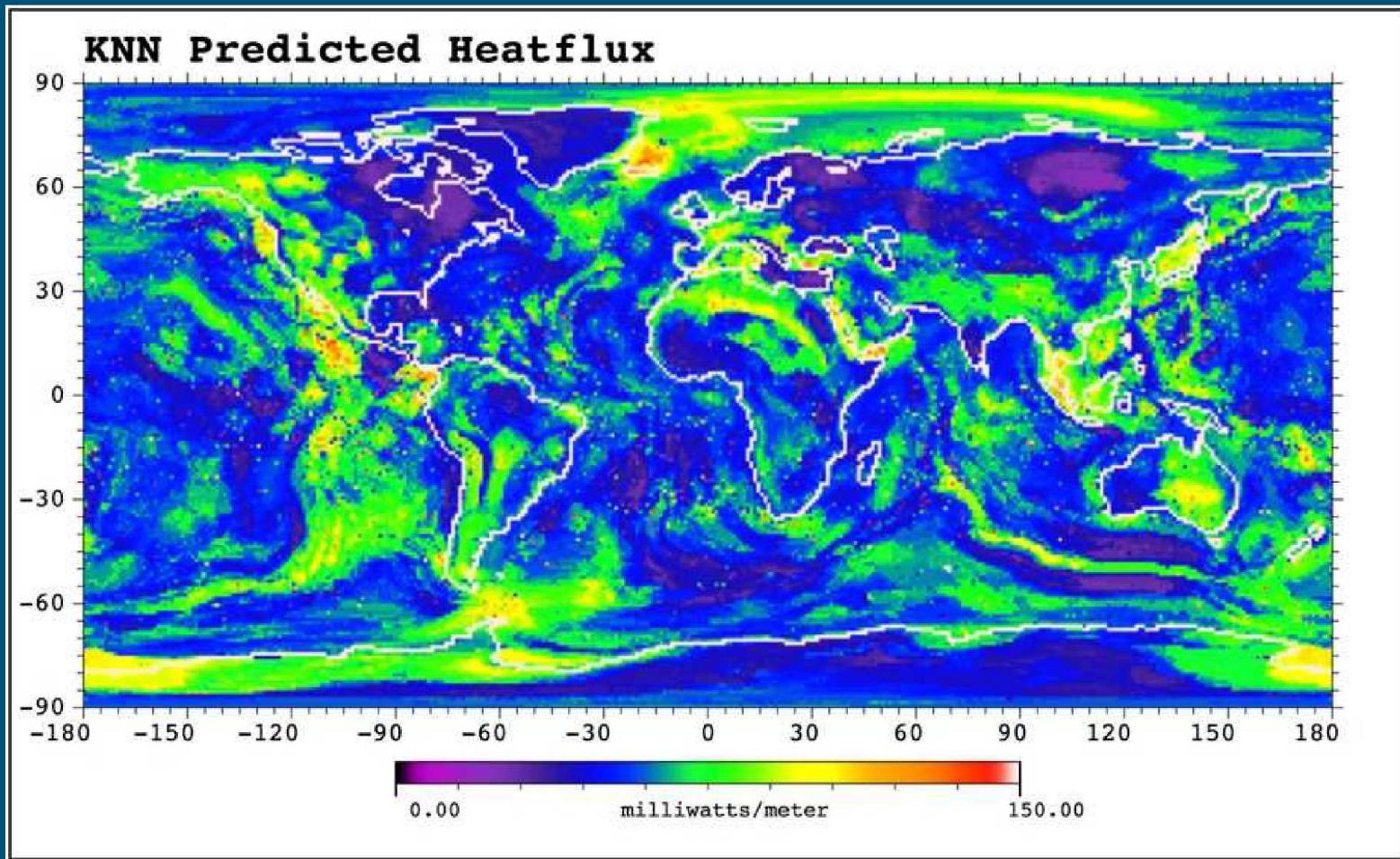
Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525. SAND No. 2019-4278 M



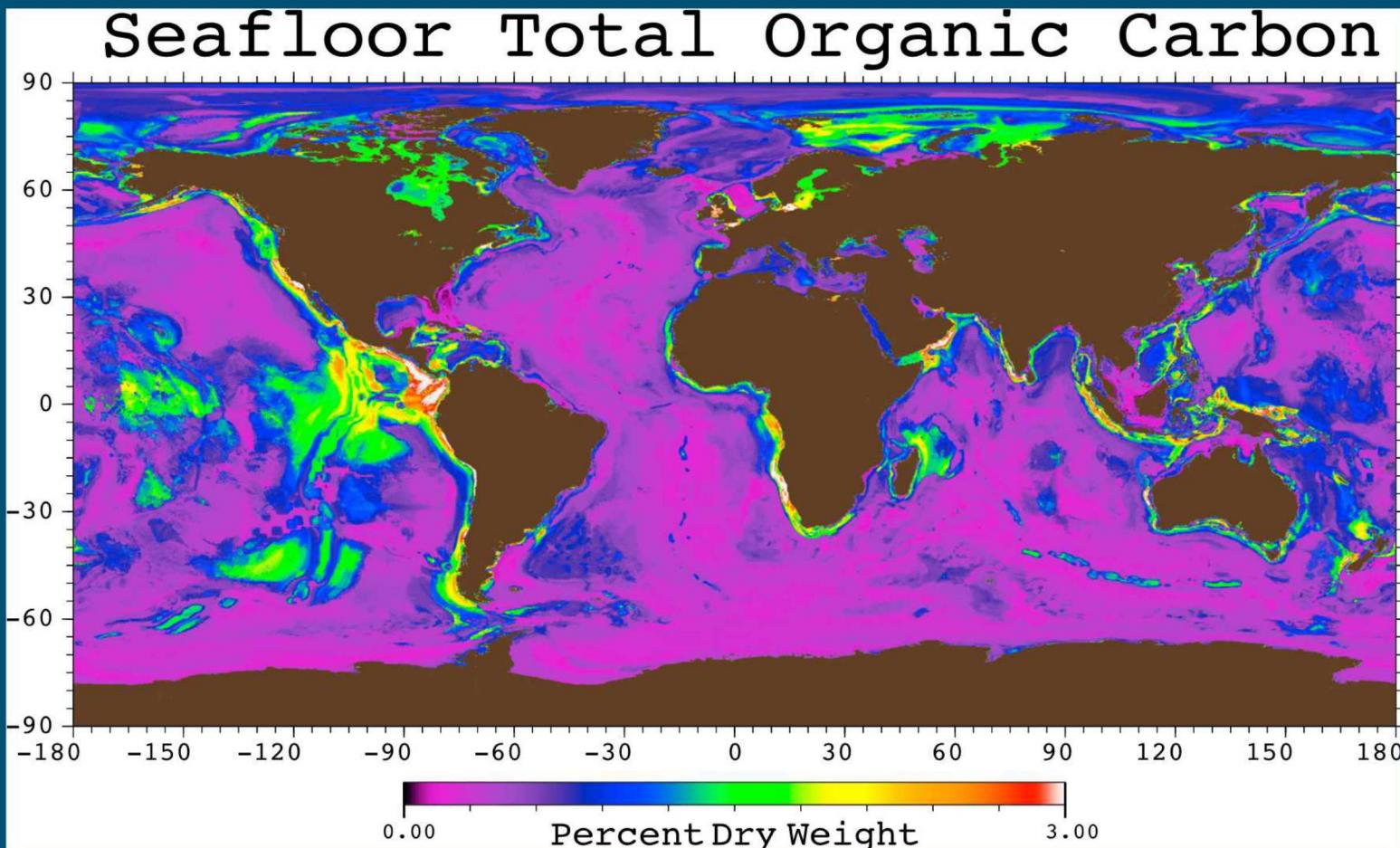




GPSM machine-learned maps



GPSM machine-learned maps



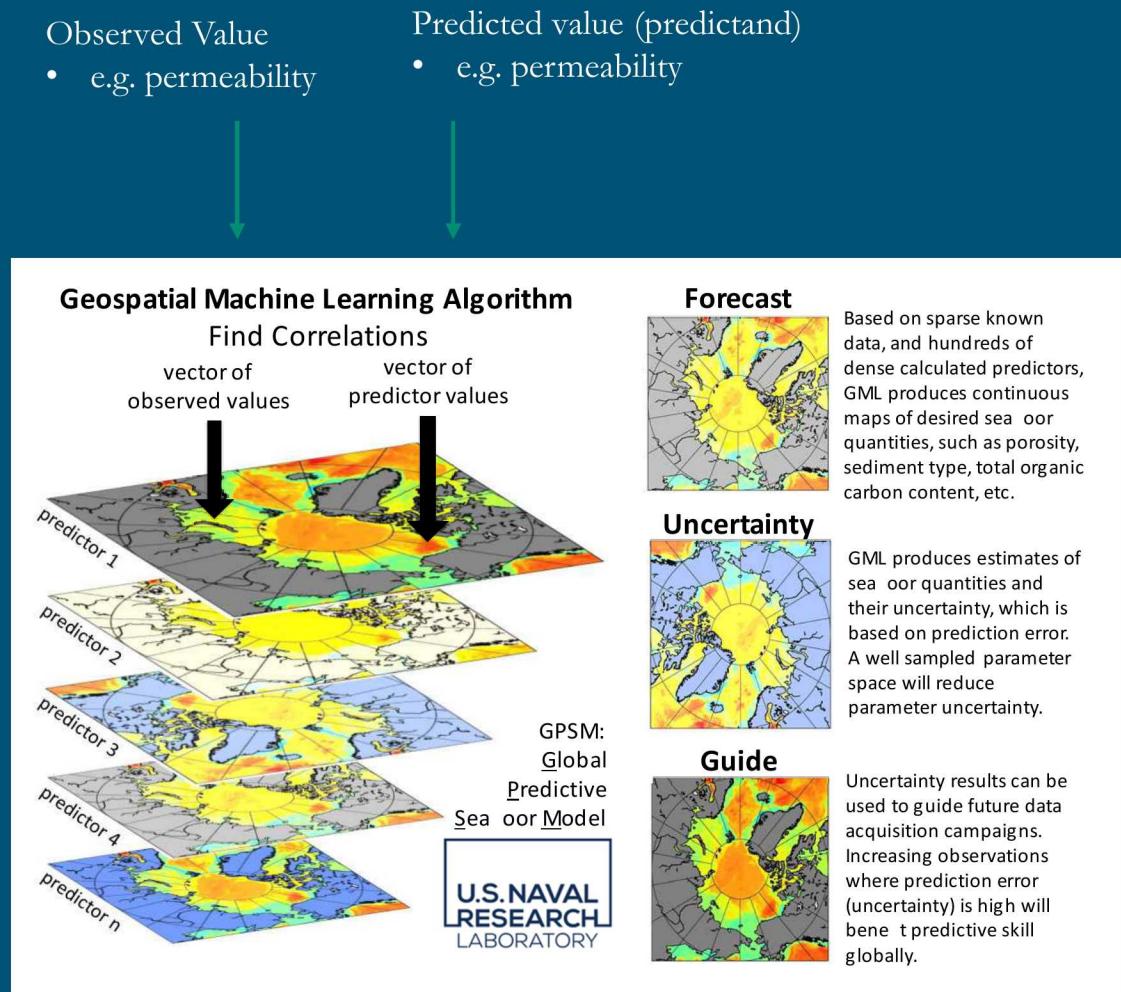
Lee et al., 2019

Machine learning \leftrightarrow Mechanistic Simulation

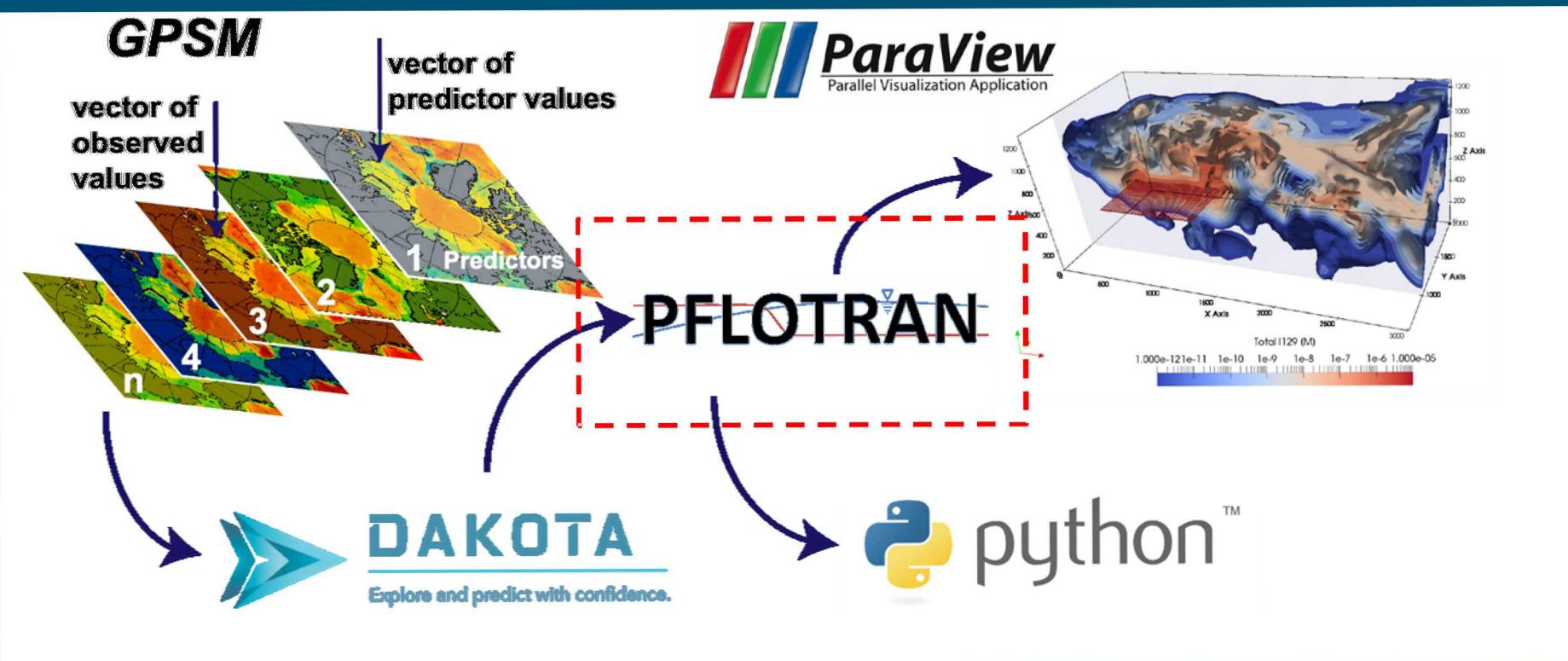
- Probabilistic maps quantify uncertainties in the seafloor parameters we are interested in
- How do these **uncertainties propagate** downward?
- What can confidence intervals on e.g. seafloor temperature, seafloor depth, seafloor organic carbon content, or heat flux tell us about the **likelihood of shallow gas or gas hydrate**?

Ex.

1. ϕ
2. α
3. v_{sed}
4. dist.
5. etc.



We need thermodynamic models

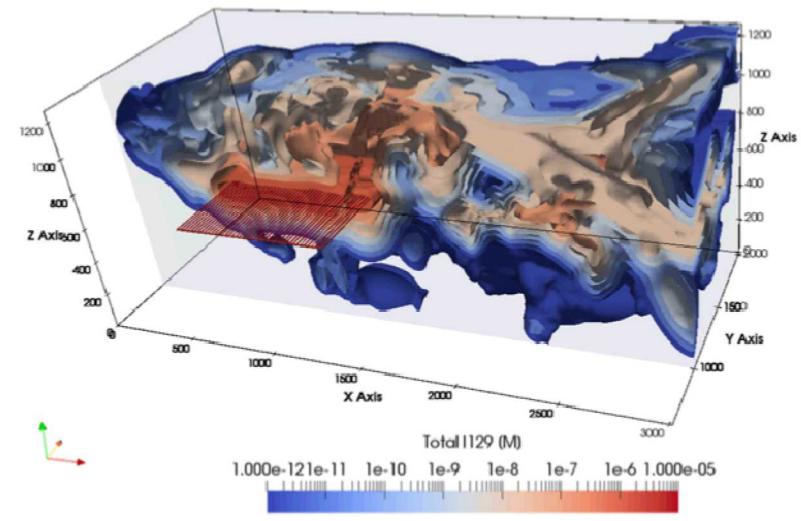




$$\frac{\partial m_a}{\partial t} = -\nabla \cdot (\rho_l X_a^l \mathbf{q}_l + \rho_g X_a^g \mathbf{q}_g + \mathbf{J}_a^l + \mathbf{J}_a^g) + q_a^G,$$

$$\frac{\partial m_w}{\partial t} = -\nabla \cdot (\rho_l X_w^l \mathbf{q}_l + \rho_g X_w^g \mathbf{q}_g + \mathbf{J}_w^l + \mathbf{J}_w^g) + q_w^G,$$

$$\frac{\partial e}{\partial t} = -\nabla \cdot (\rho_l H_l \mathbf{q}_l + \rho_g H_g \mathbf{q}_g - \kappa_{\text{eff}} \nabla T) + q_e^G,$$



Open-source software development



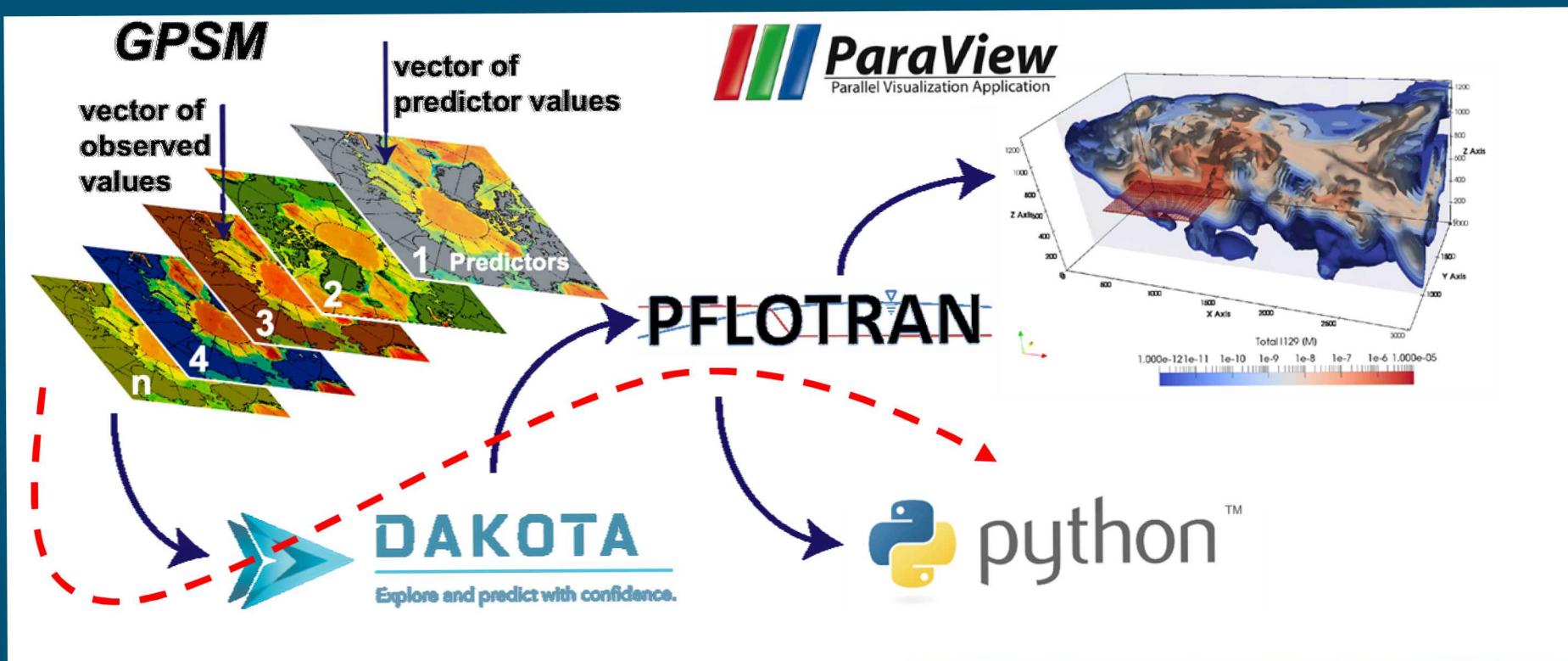
- Version control (git)
- Regression testing
- Online documentation/user's manual
- Continuous integration

 A screenshot of a GitHub repository page for "pflotran". The left sidebar shows navigation links: Source, Commits (which is selected and highlighted in blue), Branches, Pull requests, Pipelines, Deployments, Issues, Wiki, and Downloads. The main content area is titled "Commits" and shows a list of recent commits. Each commit is represented by a colored dot (representing the author) followed by the commit hash, author, date, and a brief message. The commits are ordered from most recent at the top to oldest at the bottom. The commit list includes:

Author	Commit	Message	Date
Tom Seidl	c47ee89	Merged in dtseidl/pfotran/tom/fmdm-surrogate (pull request #230) changed DECAY_TIME to use time units Approved-by: Gle...	2019-11-26
Tom Seidl	9937fd8	changed MECHANISM FMDMSURROGATE to FMDM_SURROGATE side note: time units for the DECAY_TIME sub-block assume...	2019-11-26
Tom Seidl	a227fec	changed DECAY_TIME to require time units	2019-11-25
Tom Seidl	6264747	Merged pflotran/pfotran into master	2019-11-25
Glenn Hammond	bbb42e3	Initial pass at fixing indexing of source/sink terms in general; still needs fix for numerical Jacobian	2019-11-22
Glenn Hammond	af4db21	fixed bug in regression script where timeout_error flag is uninitialized.	2019-11-22
Heeho Park	3815a2f	Merged in heeho/cpr-preconditioner-hypre-have-fix (pull request #228) PETSC_HAVE_LIBHYPRE changed to PETSC_HAVE_LI...	2019-11-21
Michael Nole	5ef26c5	Minor changes following comments on pull request.	2019-11-21
Michael Nole	5d630b4	Tweak.	2019-11-21
Michael Nole	d1b8426	Refactor to have the PM hold all hydrate_parameters, to which others will point.	2019-11-21
Michael Nole	c282cf7	Move methanogenesis parameters into the PM and out of patch. Requires changes outside of HMODE.	2019-11-21
Glenn Hammond	c6e8b83	Merge branch 'master' of https://bitbucket.org/pfotran/pfotran	2019-11-20
Glenn Hammond	aa086d1	Changed -malloc 0 -> -malloc_debug no	2019-11-20
Glenn Hammond	c43ba1b	Merged in glenn/add-major-fail-to-regression (pull request #226) Addition of MAJOR FAIL category in regression tests; time o...	2019-11-20
Glenn Hammond	b32ed5c	Added T for time out error to regression testing legend	2019-11-20
Tom Seidl	8a6edb97	Merged in dtseidl/pfotran/tom/fmdm-surrogate (pull request #223) Tom/fmdm surrogate Approved-by: Glenn Hammond geham...	2019-11-20
Tom Seidl	114a2e3	Revisions for pull request #223: Changed tolerance for fmdm_ann_surrogate regression test. Replaced real, integer, and logical...	2019-11-20
Glenn Hammond	3df5b83	Merged in glenn/add-h5py-to-travis (pull request #225) Added python-h5py to apt-get Approved-by: Glenn Hammond geham...	2019-11-19
Glenn Hammond	47c3690	Added python-h5py to apt-get	2019-11-19
Michael Nole	8c92b6d	Merged in michael/hmode (pull request #224) Michael/hmode Approved-by: Glenn Hammond gehammo@sandia.gov	2019-11-19
Michael Nole	d4244af	hydrate regression test config file update.	2019-11-19
Michael Nole	70e45a6	Move Srg into rel_perm base	2019-11-19
Michael Nole	cdf8dbf	Working radiolysis model.	2019-11-19
Glenn Hammond	b34b282	Loosened major scale to 1.d6	2019-11-18
Glenn Hammond	3fed9e	Modified major failure of be a scaling of absolute/relative tolerance	2019-11-18

Prev Next

Framework

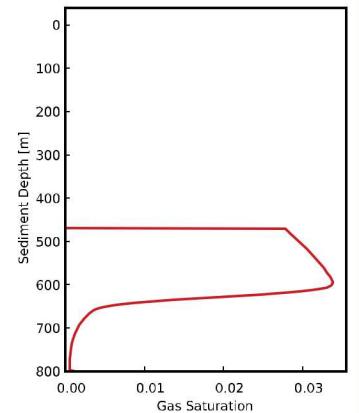
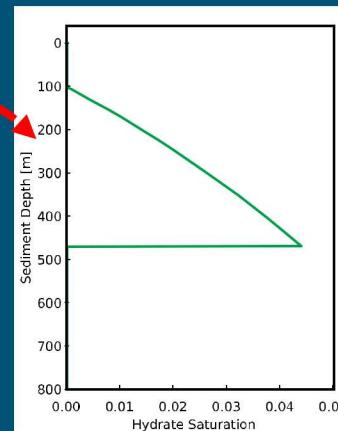
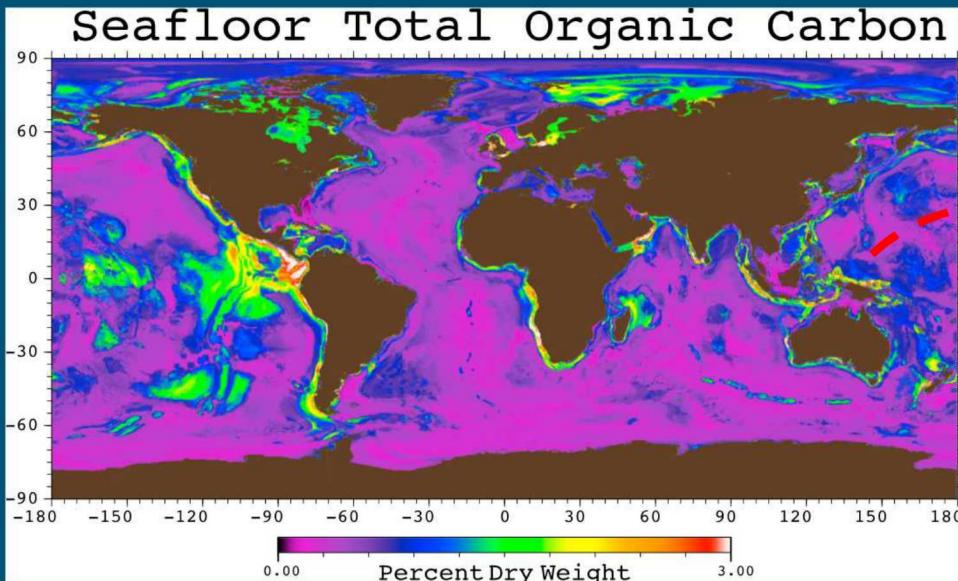


Applying the Framework

Using GPSM maps for:

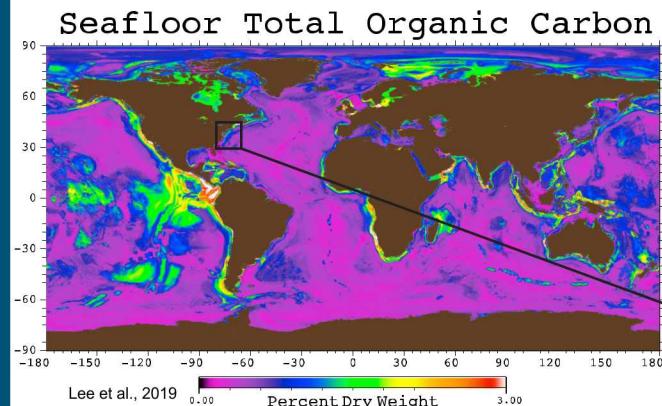
- TOC
- Sedimentation rate
- Heat flux
- Bathymetry
- Seafloor temperature

We can use mechanistic simulations to project downward into sediments



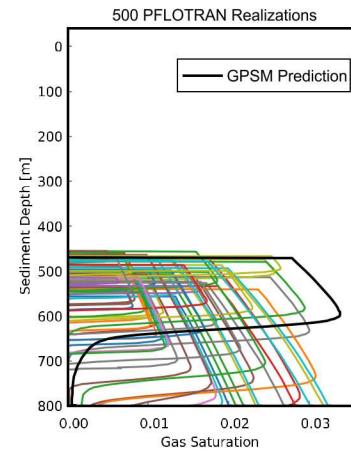
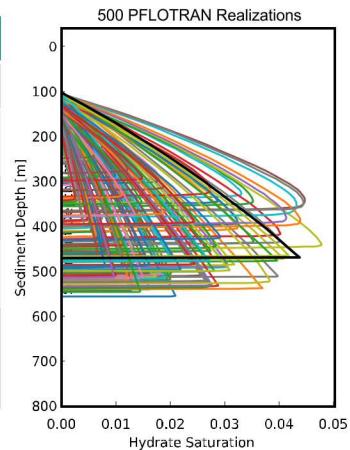
Applying the Framework

GPSM



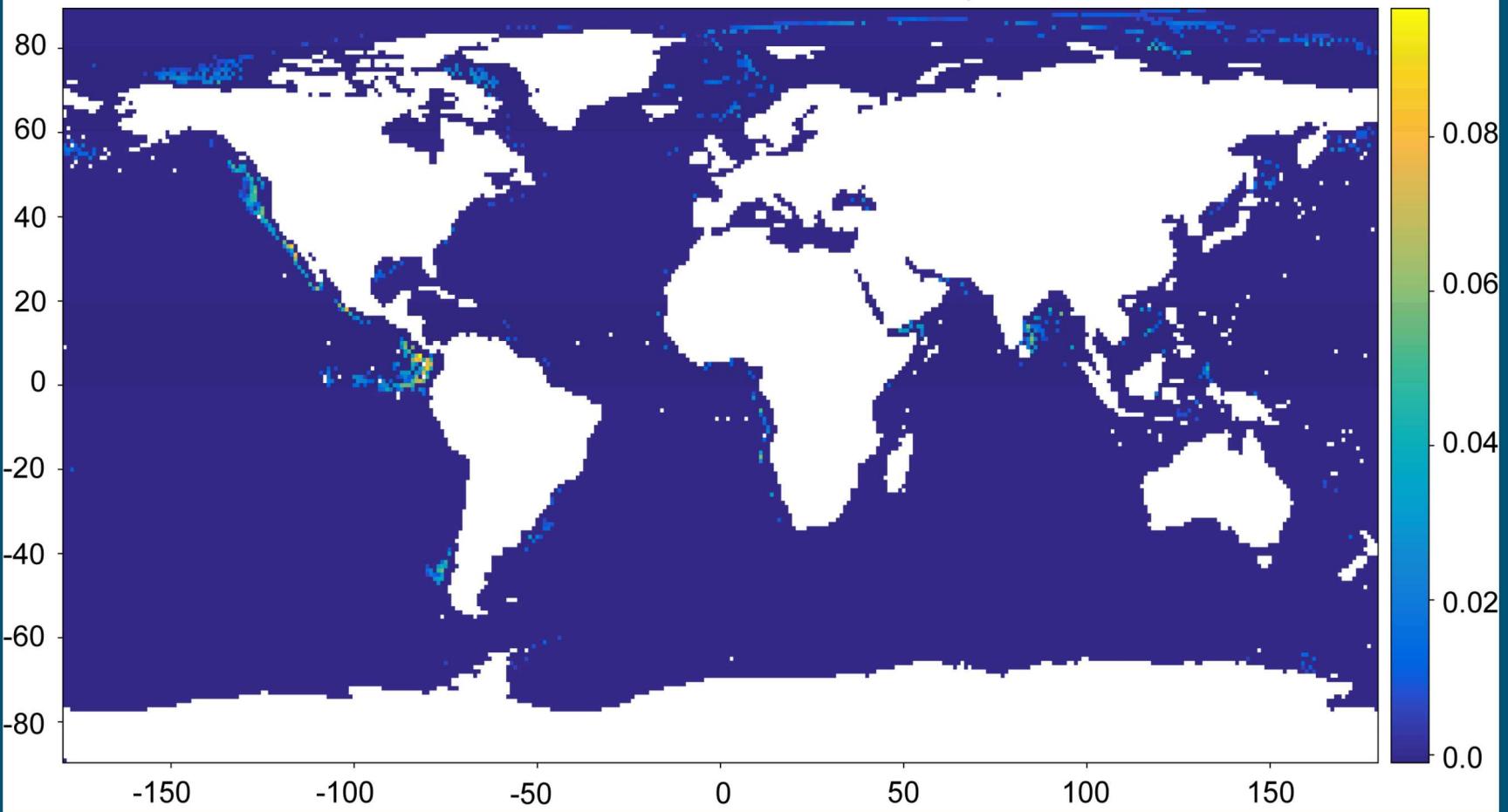
Parameter	Distribution	Min	Max	Current (GPSM)
Seafloor Depth (m)	Uniform	2760	2800	2780
Seafloor Temperature (C)	Uniform	3	5	4
Heat Flux (W/m ²)	Uniform	0.03	0.08	0.06
Sedimentation Rate (m/yr)	Log-uniform	1x10 ⁻⁴	1x10 ⁻³	2x10 ⁻⁴
Seafloor Organic Carbon Fraction	Log-Uniform	1x10 ⁻³	2x10 ⁻²	1x10 ⁻²

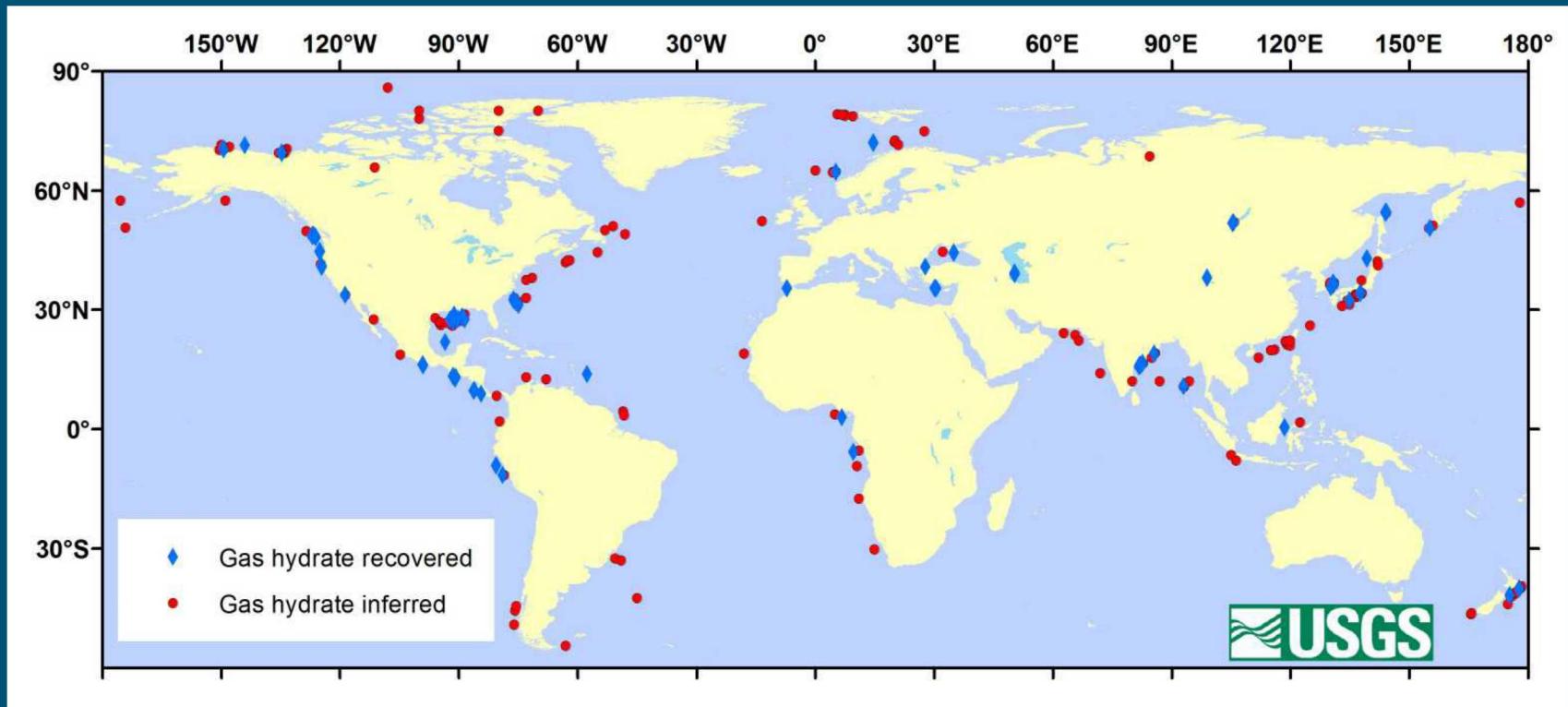
PFLOTRAN



*See Frederick et al. poster at the GRC!

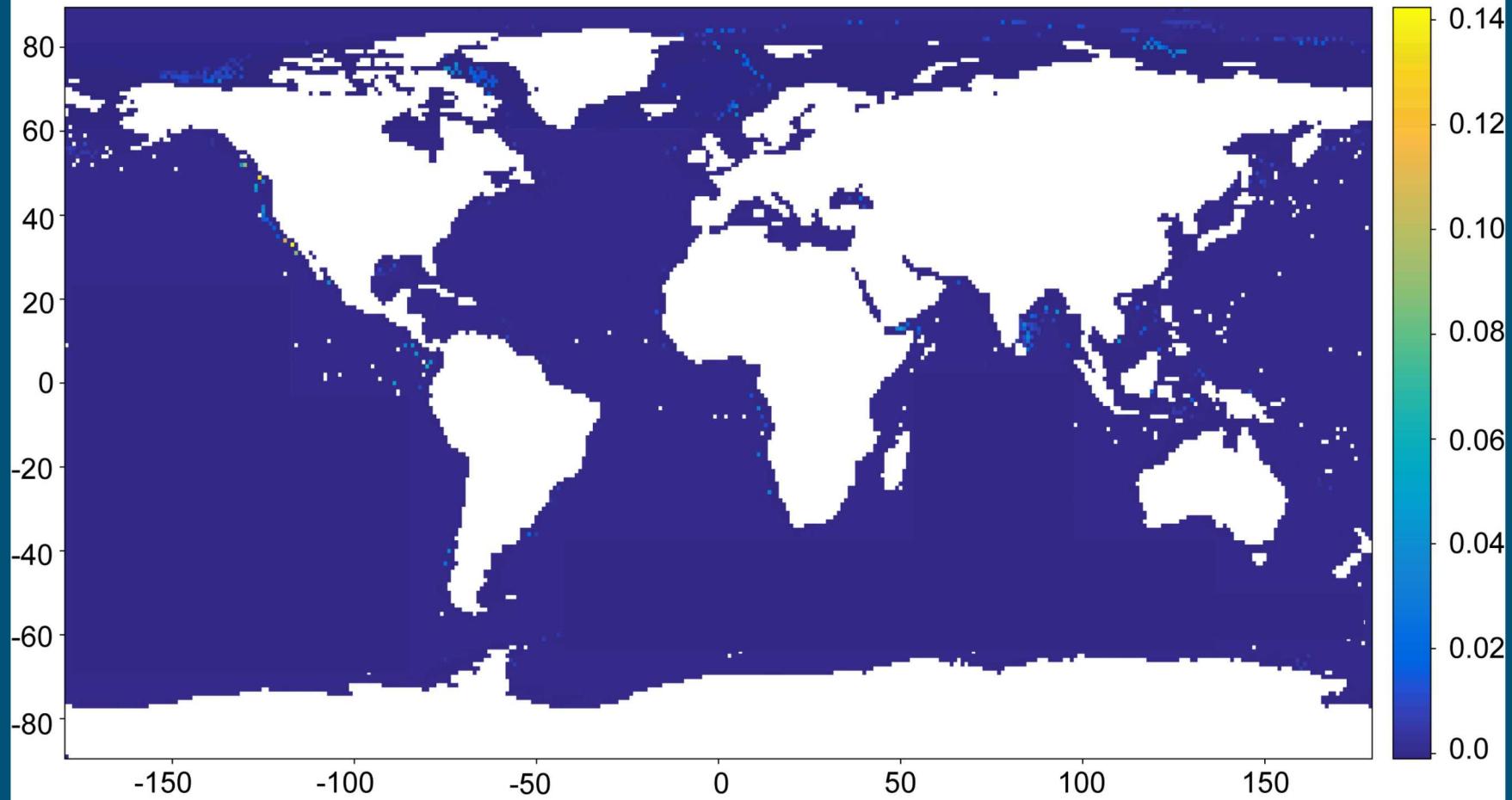
GPSM-PFLOTRAN Predicted Maximum Gas Hydrate Saturation







GPSM-PFLOTRAN Predicted Maximum Gas Saturation



Future Work

- Higher-resolution maps, zoomed in sub-sampling
- Quantify uncertainty
- Sub-sample, do 3D
- Mixed migration mechanisms
- Expanded capabilities (mixed hydrates [with UT Austin], geomechanics/slope stability with CSM, etc)

Affiliated posters you should check out

- Jenn Frederick
- David Fukuyama