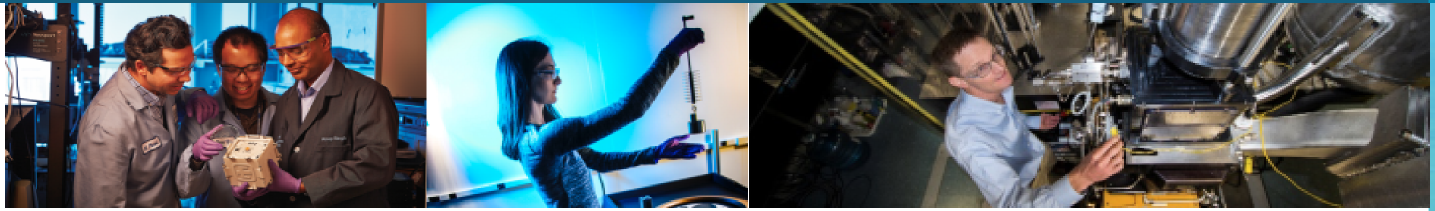
SAND2021-15599C

E3SM Atmosphere surrogate construction using machine learning and reduced order modeling methods



Kenny Chowdhary, Benjamin M. Wagman

PRESENTED BY

Kenny Chowdhary

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & 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.

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & 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.

Motivation

Tuning (calibration) state-of-the-art	Automated tuning (goal)
Time-consuming (6-12 months)	Less than 6-12 months
Deterministic: one set of "tuned" parameters per model release	Probabilistic: a distribution of parameters per model release OR observational target
Non-reproducible	Reproducible
Computationally expensive	Computationally expensive

Energy Exascale Earth System Model

Goals and motivating questions

- Goals:
  1. Build a surrogate that maps uncertain model parameters to spatial fields instead of to a cost function
     - Spatial fields build intuition and enable users to customize field-dependent cost functions
  2. Do it faster than the experts

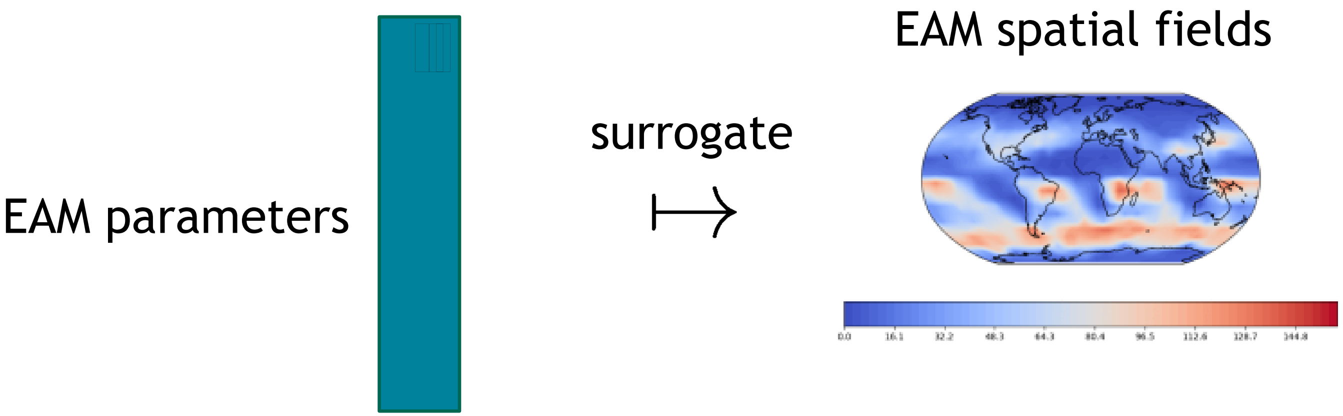
- Motivating questions

1. How many climate model simulations (samples) are needed to create a surrogate of a given accuracy?
2. What is an achievable surrogate spatial field dimensionality and resolution?
   - E.g. 1-D, coarsened 2-D, 2-D?
3. Should the multi-objective targets each get their own surrogate? Or should all targets be combined into one surrogate?
4. What can we learn from ultra-low resolution E3SM (ne4) to guide our surrogate construction for the low-resolution E3SM (ne30)?

E3SM Energy Exascale Earth System Model

Surrogate construction

- Our approaches to calibration require surrogate models that map the feature space, E3SM Atmosphere Model (EAM) parameters **x**, to a spatially varying climatological output field (a multi-output field or image).
- In this work, we explore data-driven reduced order modeling (ROM) and ML based surrogates.
  - Combine dimension reduction with classical ML techniques (unsupervised + supervised).
- These surrogates need to be **multi-target, and multi-objective (>1 field QoIs)**

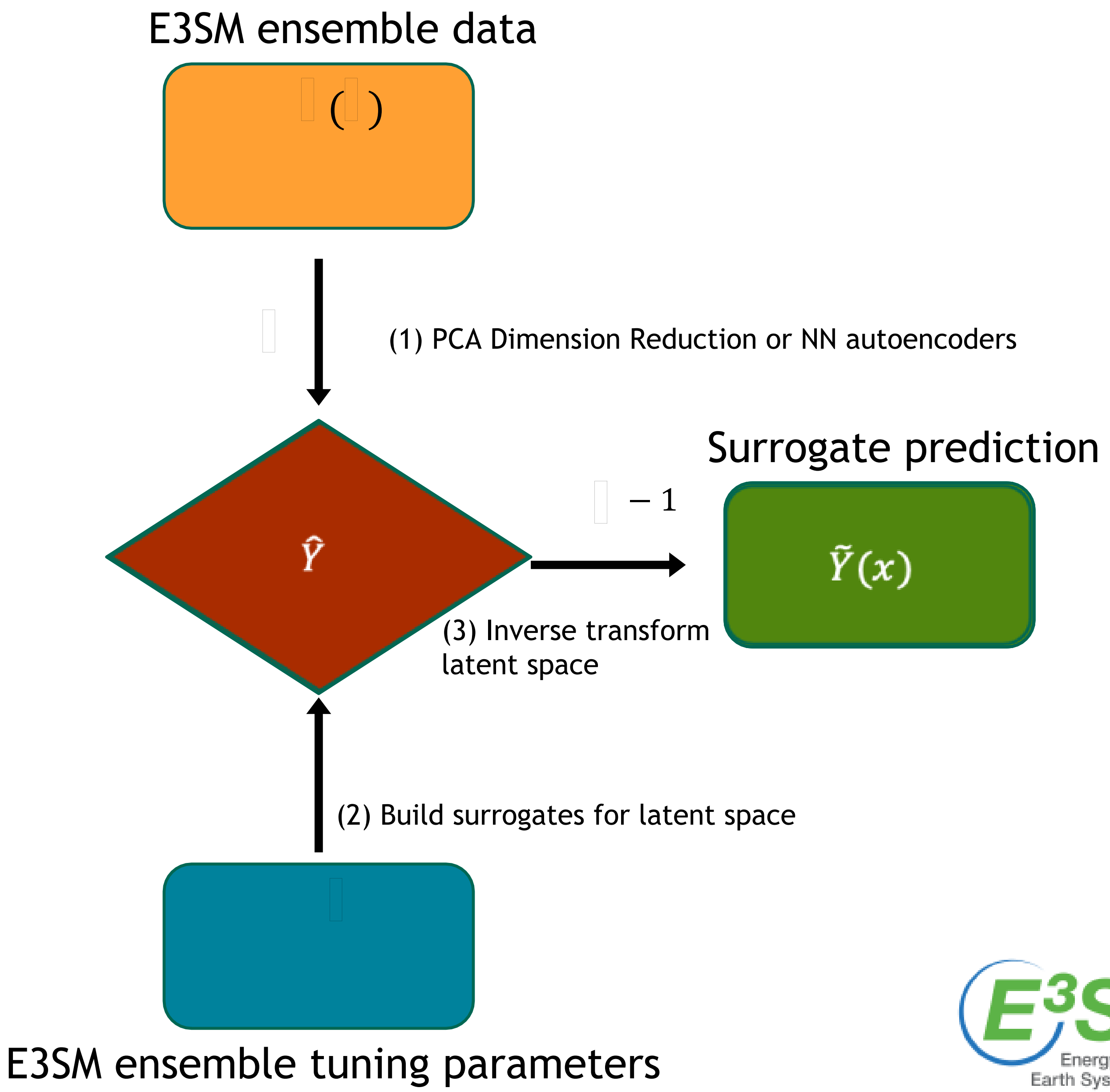


ROM based surrogate construction architecture

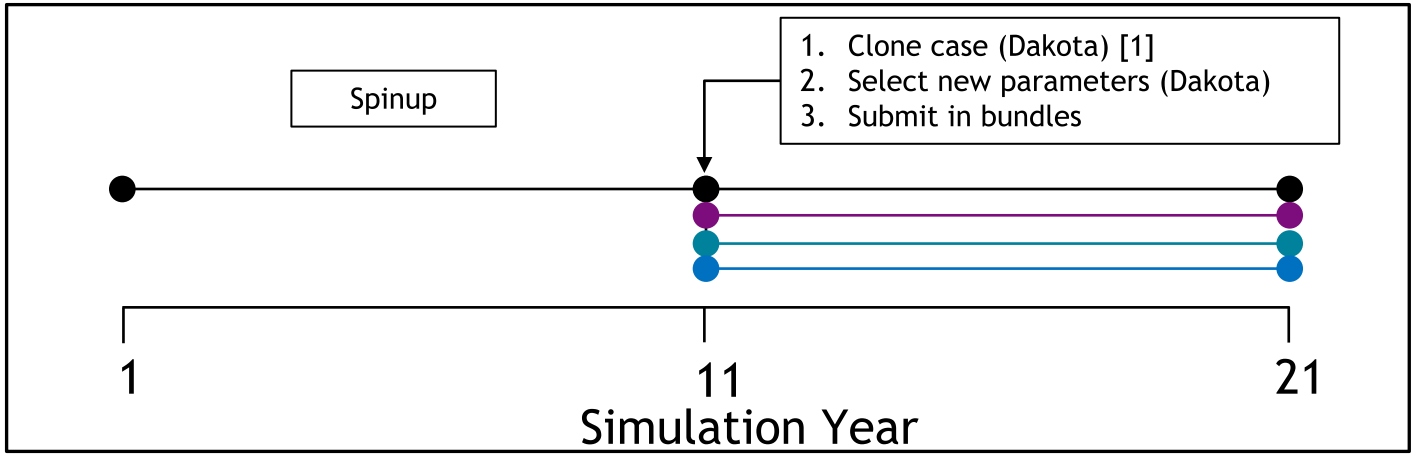
1. Dimension reduction (DR) of the target data $Y(x)$ to a low-dimensional latent space.

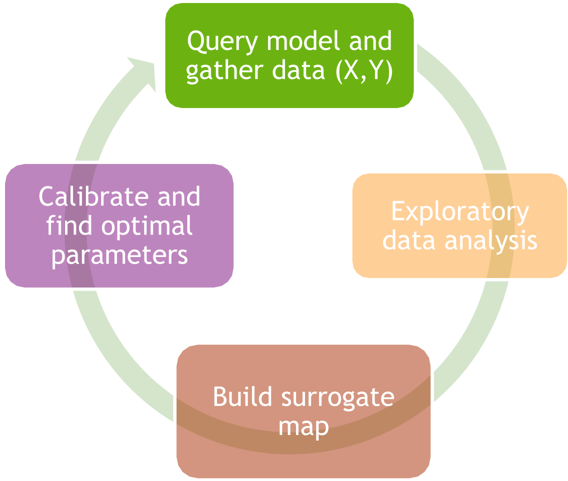
2. Build machine learning surrogates for each latent space dimension (this is where most of our effort is spent).

3. Finally, we map the latent space surrogate values back to the original space to get our surrogate prediction $\tilde{Y}(x)$.



E3SMv2 simulations





Ensemble	Res.	Config.	n	Yrs.	Nodes per bundle	Sims. per bundle	SYPD per bundle
Ultra-low resolution (ULR)*	~7.5°	F2010	500	10	100	50	5100
Low resolution (LR)	~1°	F2010	200	10	100	10	95

[1] Adams et al., 2014

*E3SM ULR is not tuned or scientifically validated

E3SMv2 sampled atm. parameters

Parameter	Description	Low-Def-High [2]
clubb_c1	Constant for dissipation of variance of mean(w'^2)	1.0__**1.335**__5.0
clubb_gamma_coef	Constant of the width of PDF in w coordinate	0.1__**0.32**__0.5
zmconv_tau	Time scale for consumption rate deep CAPE	1800__**3600**__14400
zmconv_dmpdz	Parcel fractional mass entrainment rate	-2.0e-3__**-0.7e-3**__-0.1e-3
micro_mg_ai	Fall speed parameter for cloud ice	350__**500**__1400

Field	Coordinates	Size ULR	Size LR
TREFHT	lat x lon	24x48	24x48, 129x256
PRECT	lat x lon	24x48	24x48, 129x256
SWCF	lat x lon	24x48	24x48, 129x256
LWCF	lat x lon	24x48	24x48, 129x256
PSL	lat x lon	24x48	24x48, 129x256
FLNT	lat x lon	24x48	24x48, 129x256
FSNT	lat x lon	24x48	24x48, 129x256
Z500	lat x lon	24x48	24x48, 129x256
U200	lat x lon	24x48	24x48, 129x256
U850	lat x lon	24x48	24x48, 129x256
RELHUM	**lat x lev**	**24x37**	**24x37, 129x37**
T	**lat x lev**	**24x37**	**24x37, 129x37**
U	**lat x lev**	**24x37**	**24x37, 129x37**

[2] Qian et al., 2018

E3SM ne4 results

What’s the best ML method for fitting the latent space?

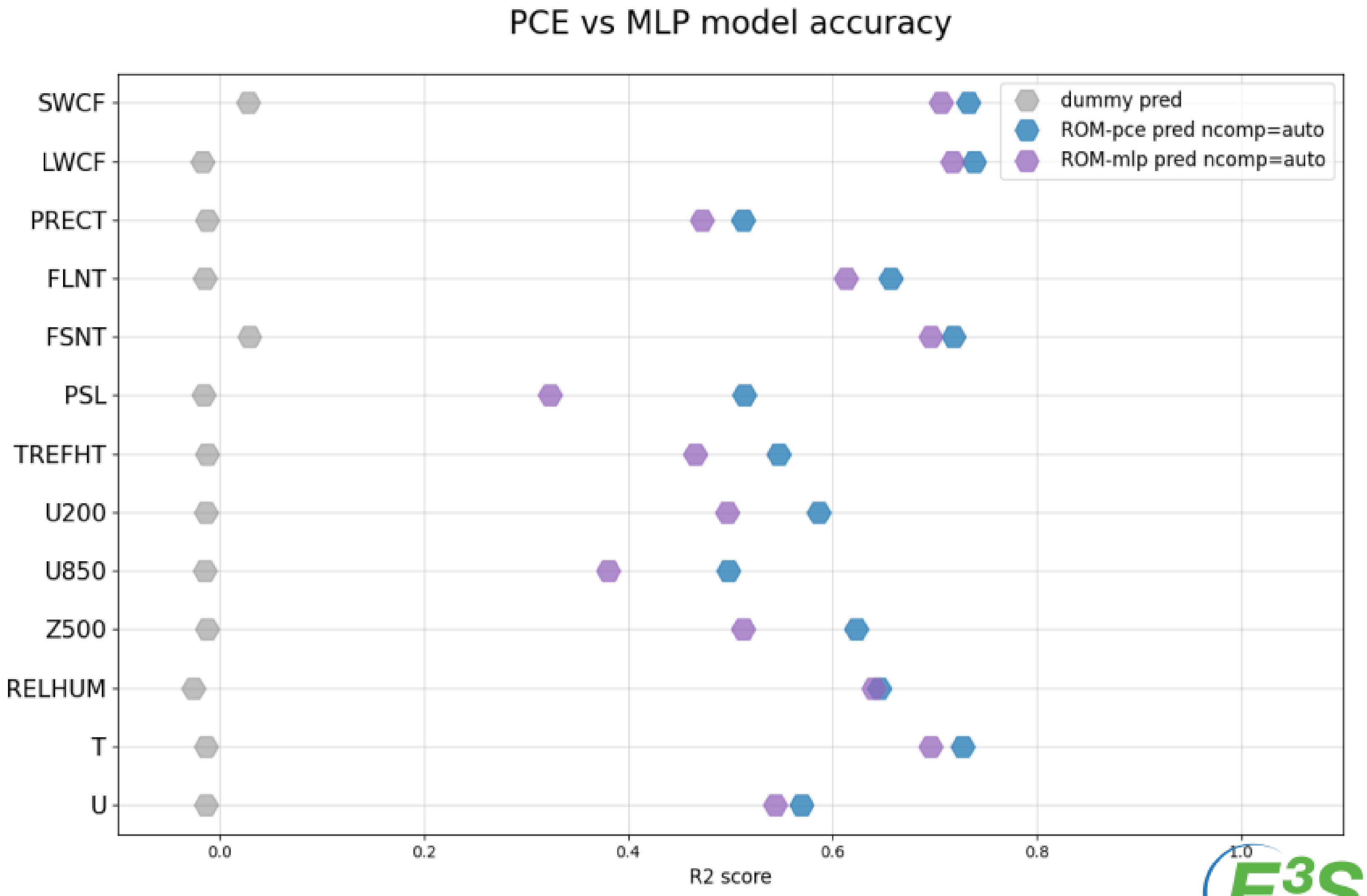
What’s the best target type – single objective (one for each climatology) or multi-objective (all of them combined)?

How many climate model samples do we need to build a good surrogate?

What can we learn from the ne4 model to guide our surrogate approach for ne30?

MLP vs Polynomial-based ROM surrogates

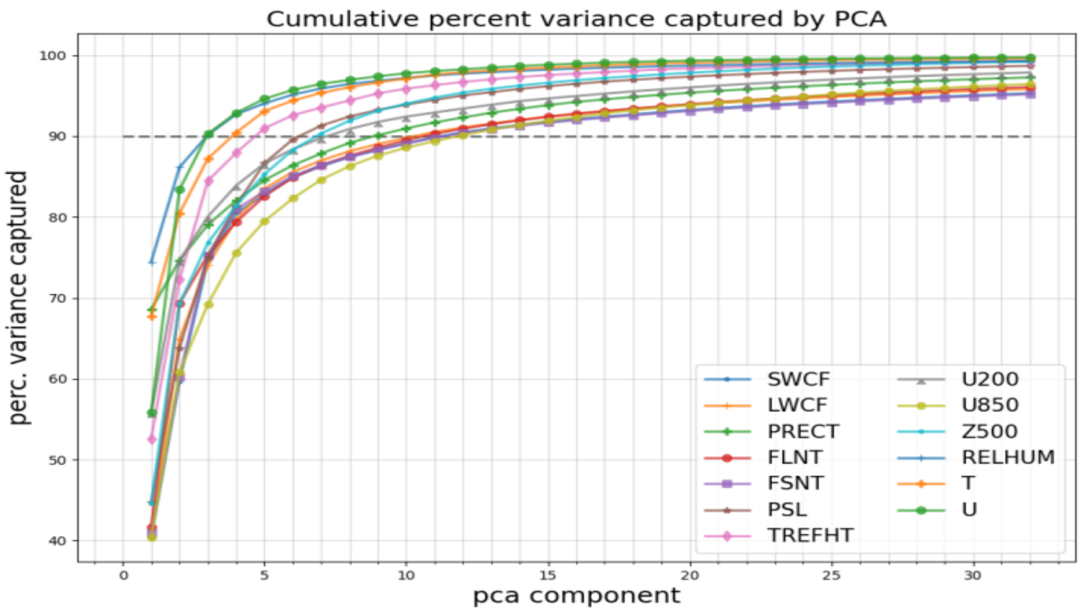
- We fit each of our 13 targets using the ROM-based ML surrogates
- We used k-fold cross-validation with hyper-parameter tuning for our model selection strategy.
- Our metric for comparison was the average R^2 scores (computed for each latitude-longitude coordinate and averaged over the globe)
- Takeaway: Polynomials performed better than MLP approaches for fitting the latent space.

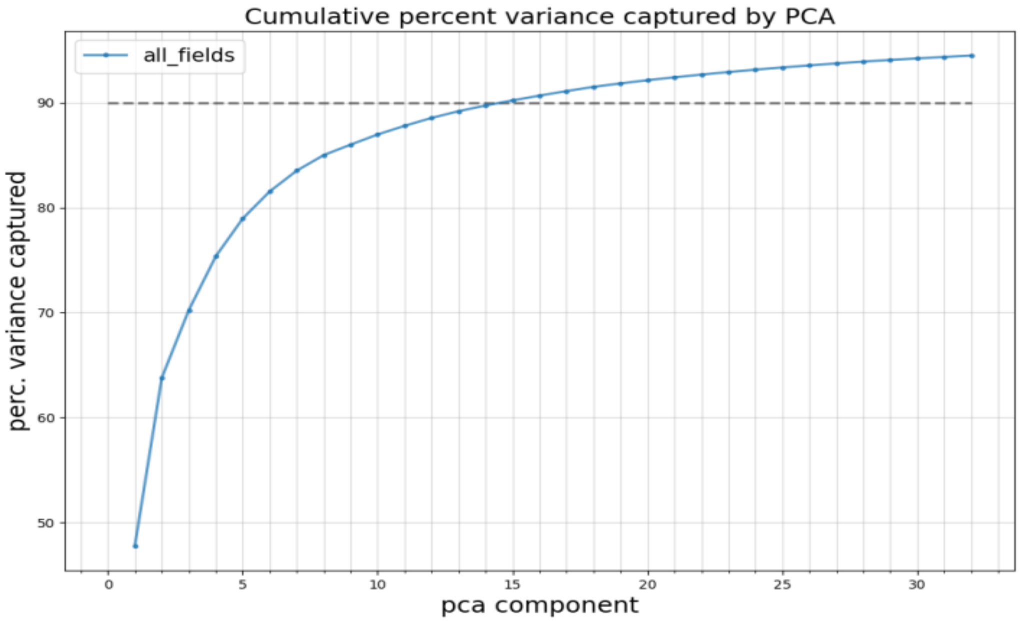


E3SM Energy Exascale Earth System Model

Single objective models vs multi-objective models

12 PCA components captures >= 90% of the variance in each field.



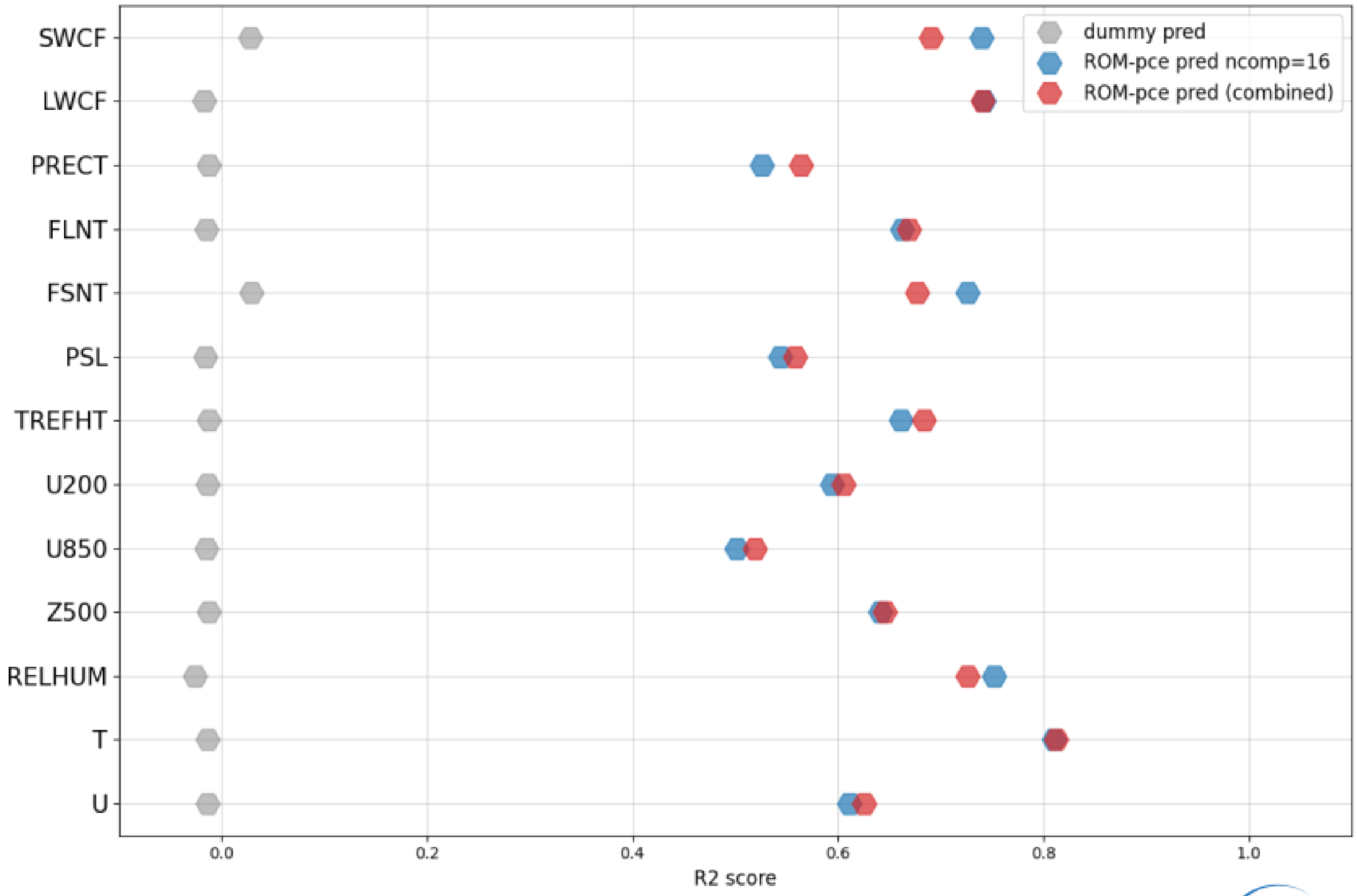


15 PCA components captures >= 90% of the variance in the multi-objective field.

- As far as surrogate accuracy for the same number of components, there is no clear winner.

- Multi-objective is preferred for efficiency and easier tuning of latent dimension dimensions. - only have to build one surrogate but need scaling to combine the vector

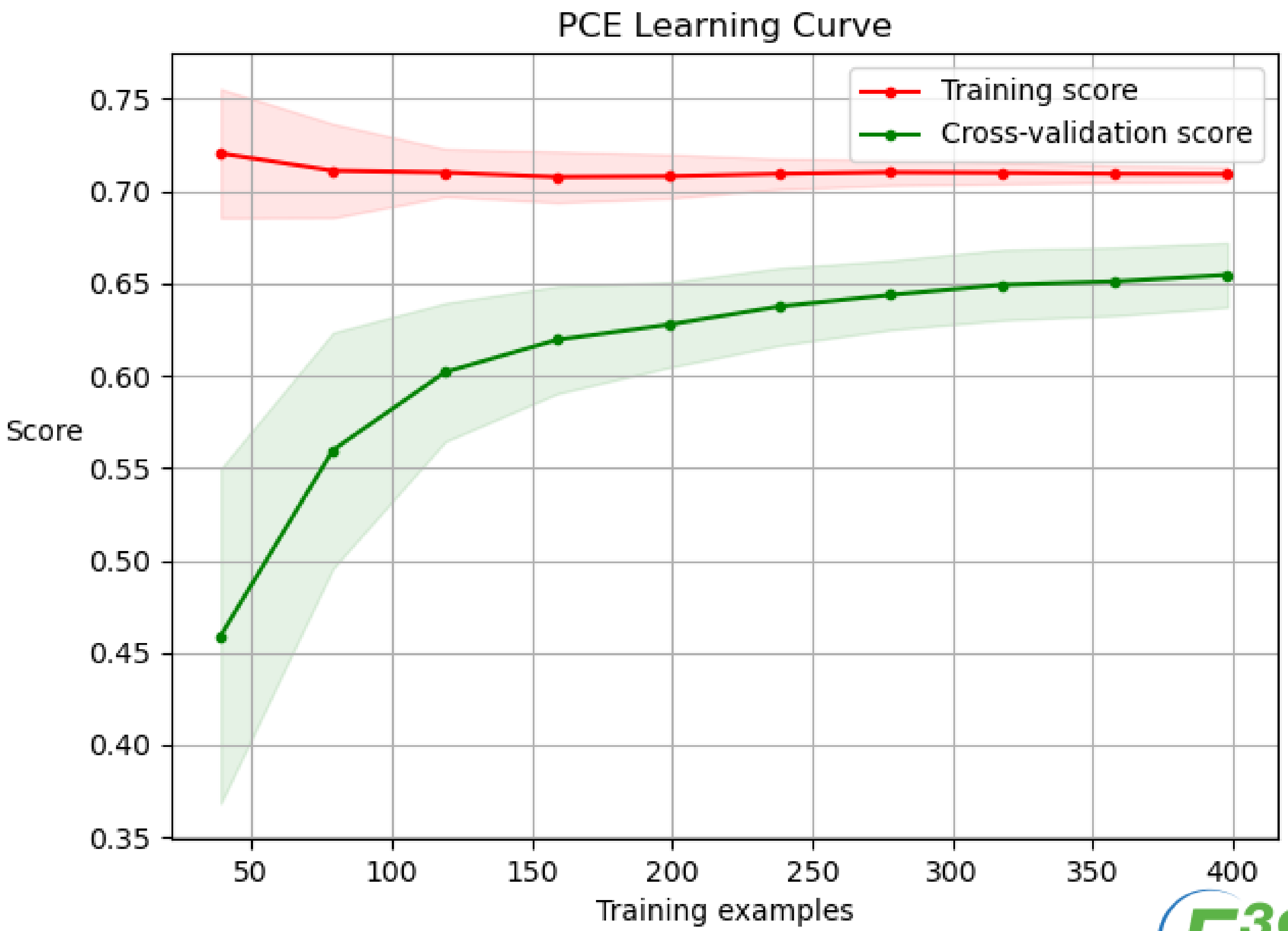
- Multi-objective is 13x faster to train since we only need a single model for all targets, but we lose some accuracy in some, while we gain in others.



E3SM Energy Exascale Earth System Model

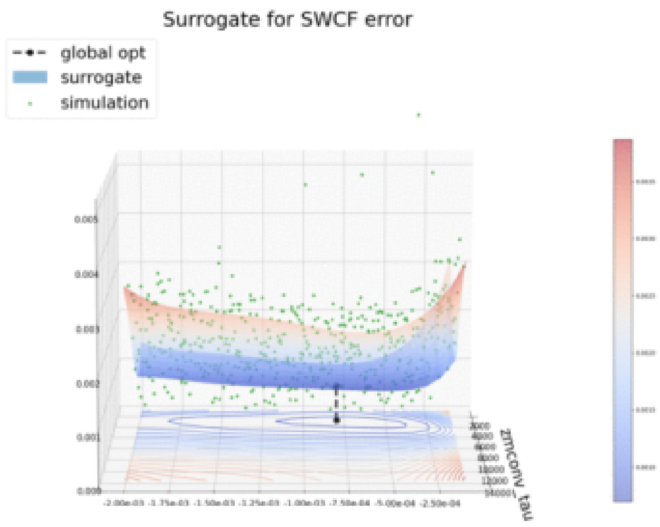
How much data is enough?

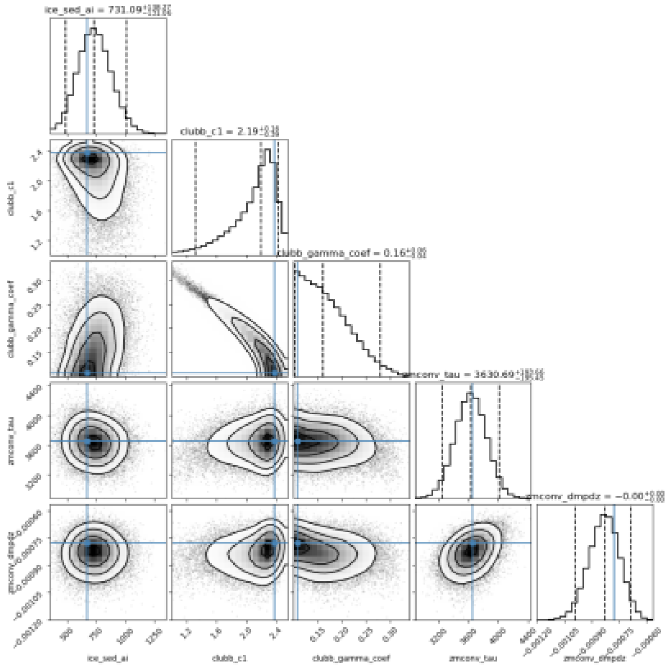
- The learning curve shows the gap between the training and testing (cross-validation) scores for different fractions of the data.
- Even with 400 training points, there is a significant gap between the training and testing scores, showing that more training data would be useful.
  - Also, the curve seems to level off after 200-300 training points indicating that one obtains a smaller return-on-investment beyond this.
- We used the ne4 study to inform us about the data requirements for the ne30.



Bayesian calibration

- Once the surrogate is constructed, we can begin performing more computationally intensive tuning and/or Bayesian calibration.
  - The top right figure shows the error surface of the surrogate which is a smooth approximation to the data shown in green. The surrogate provides a sort-of road map to the location of the most likely set(s) of tuning parameters.
- The idea is to use the surrogate and set of observations to find the optimal set(s) of EAM tuning parameters such that discrepancy between the surrogate and the observations is minimized (in some sense, e.g., squared error loss).
- With the Bayesian approach, instead of a single set of parameters, we can obtain a probability density on the parameters indicated the most likely set of parameters and their correlations (see the joint density plot to the bottom right).





Summary and future work

What's the best ML method for fitting the latent space?
- A: Polynomials seem to the winner.

What's the best target type – single objective (one for each climatology) or multi-objective (all of them combined)?
- There is no clear winner in terms of accuracy, but multi-objective is significantly more efficients (13x faster for 13 targets)

How many climate model samples do we need to build a good surrogate?
- Even 400-500 samples is probably not enough for the ne4 model. 200-300 samples seems to provide a good start for the ne30.

What can we learn from the ne4 model to guide our surrogate approach for ne30?
- What we haven't shown is that with a ROM-based PCE model, we can easily compute Sobol parameter sensitivities, for which there seems to be consistency between the ne4 and ne30 models.

Next Steps?
- Using the ne4 studies as a guide and testbed for our ROM-based ML methods, we are currently repeating the experiments with roughly 200 ne30 samples.
- Initial results are very interesting and show that there is going to be a significant hurdle going from ne4 and ne30 in terms of model complexity and achieving similar levels of accuracy.

E3SM Energy Exascale Earth System Model

References and Acknowledgments

- Adams, B.M., Bohnhoff, W.J., Dalbey, K.R., Ebeida, M.S., Eddy, J.P., Eldred, M.S., Frye, J.R., Geraci, G., Hooper, R.W., Hough, P.D., Hu, K.T., Jakeman, J.D., Khalil, M., Maupin, K.A., Monschke, J.A., Ridgway, E.M., Rushdi, A.A., Stephens, J.A., Swiler, L.P., Vigil, D.M., Wildey, T.M., and Winokur, J.G. (2014). Dakota, A Multilevel Parallel Object-Oriented Framework for Design Optimization, Parameter Estimation, Uncertainty Quantification, and Sensitivity Analysis: Version 6.8 Reference Manual. *Sandia Technical Report SAND2014-5015*.
- Energy Exascale Earth System Model v2 https://doi.org/10.11578/E3SM/dc.20210927.1
- Qian, Y., Wan, H., Yang, B., Golaz, J.-C., Harrop, B., Hou, Z., et al. (2018). Parametric sensitivity and uncertainty quantification in the version 1 of E3SM atmosphere model based on short perturbed parameter ensemble simulations. *Journal of Geophysical Research: Atmospheres*, 123, 13,046–13,073. https://doi.org/10.1029/2018JD028927
- Hoang, C., Chowdhary, K., Lee, K., Ray, J. Projection-based model reduction of dynamical systems using space-time subspace and machine learning (2021). Computer Methods in Applied Mechanics and Engineering. https://doi.org/10.1016/j.cma.2021.114341

This research was supported as part of the Energy Exascale Earth System Model (E3SM) project, funded by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research.