Physics-Informed Machine Learning for Probabilistic Space Weather Modeling and Forecasting: dB/dt and Geomagnetically Induced Currents

Piyush M. Mehta[1]; Richard Licata[1]; Daniel T. Welling[2]; Michele Cash[3]; Steven Morley[4]

1. West Virginia University, Morgantown, WV; 2. University of Texas, Arlington, TX; 3. NOAA Space Weather Prediction Center, Boulder, CO; 4. Los Alamos National Laboratory, Los Alamos, NM.

INTRODUCTION

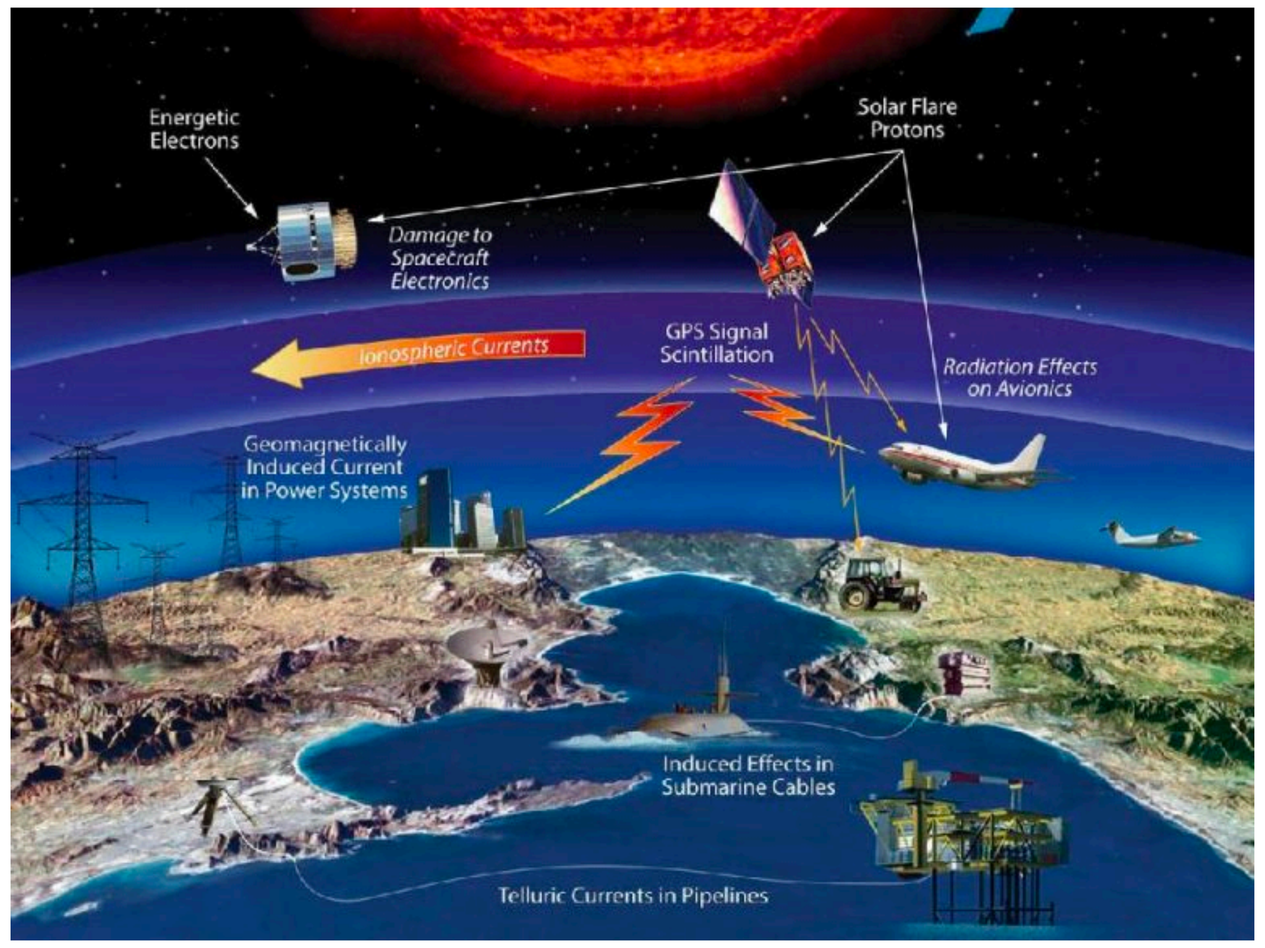
Space Weather can adversely impact technological infrastructure and humans on ground and in space.

Geomagnetically Induced Currents (GICs) are driven by the interaction of the solar wind and interplanetary magnetic field.

They can adversely impact most long, ground-based conductors such as power grids, oil and gas pipelines, etc. They have been known to affect power gridlines causing power outages for long periods.

Robust decision making in the context of space weather and its impacts require accurate modeling and forecasting along with characterization of the associated uncertainties. This is limited by the computational cost of models.

Goal: Develop a Physics-based, Data-Driven Framework for Modeling and Forecasting Space Weather and Quantification of the Associated Uncertainties.

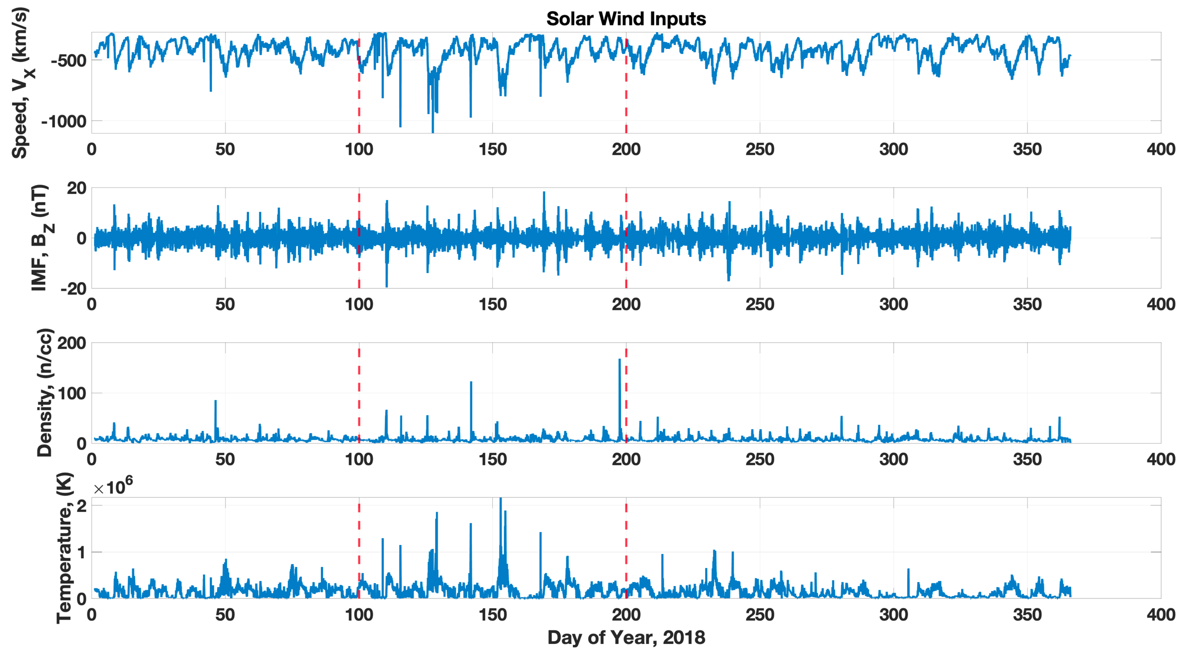


DATA PREPARATION

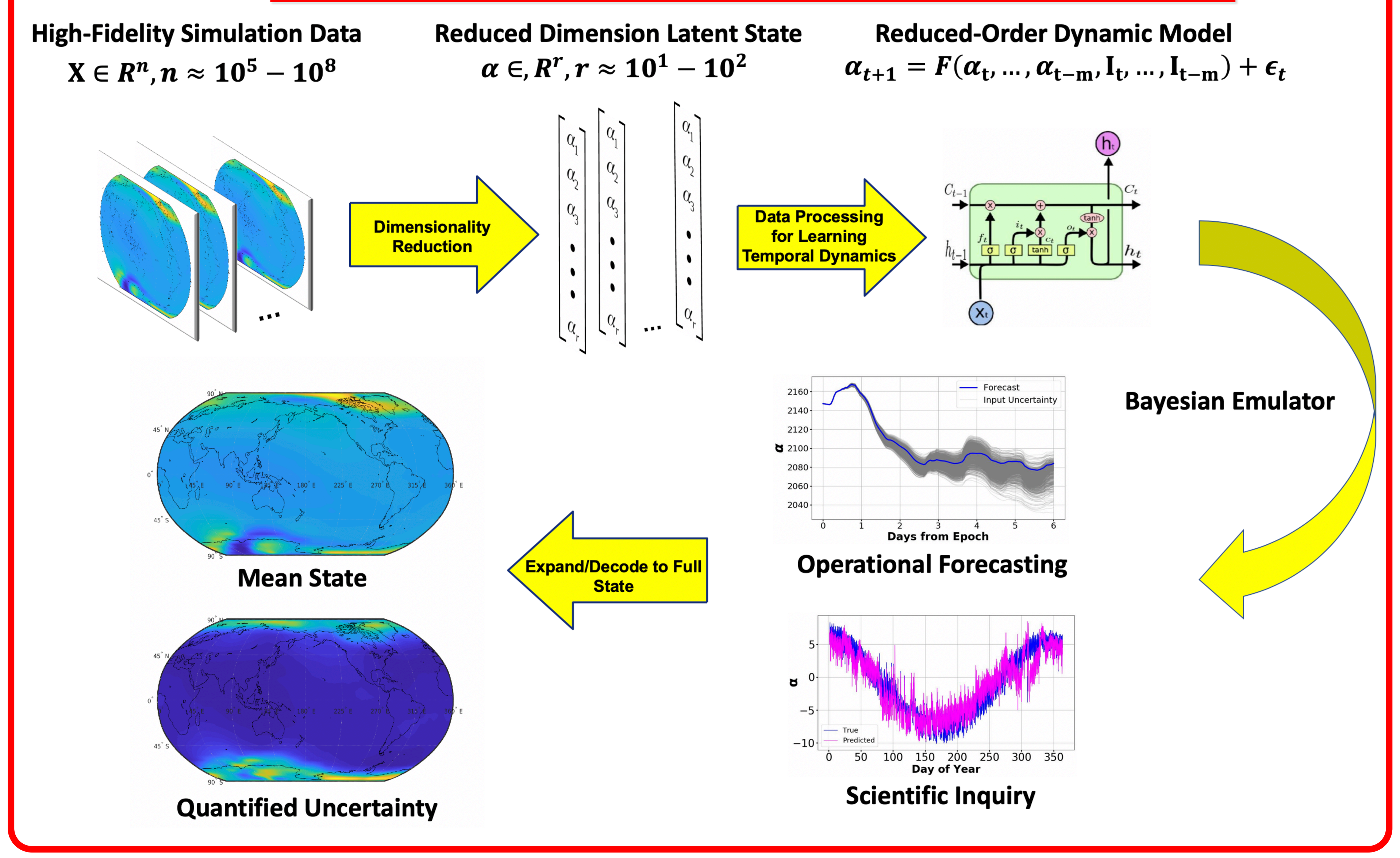
Data covering three years (part of 2016, 2017, 2018 and part of 2019) of operational output from the Space Weather Modeling Framework (SWMF) archived at NOAA was pre-processed.

Following a solar wind data gap of more than 2 hours, SWMF requires a restart. The burn-in period after a restart precludes the data from use for learning.

For demonstration here, we identify about a 3-month period with minimal data gaps to learn temporal dynamics. The figure below shows the solar wind inputs during the 100-day time period in 2018.



PHYSICS-INFORMED MACHINE LEARNING



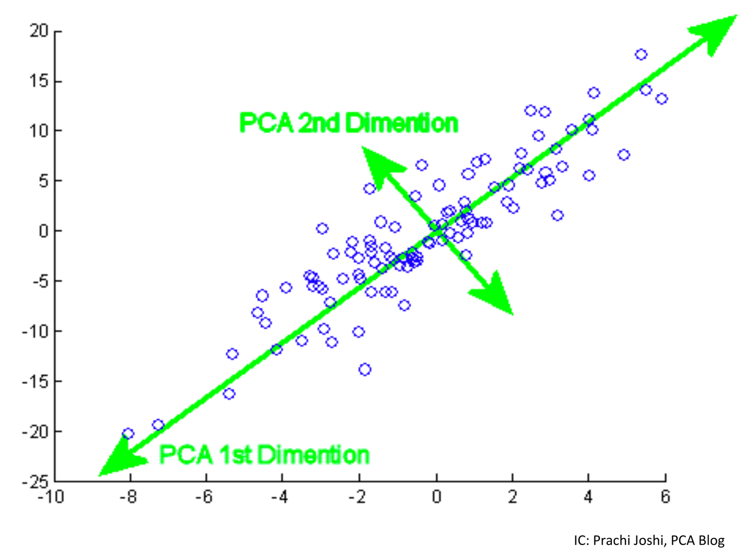
DIMENSIONALITY REDUCTION

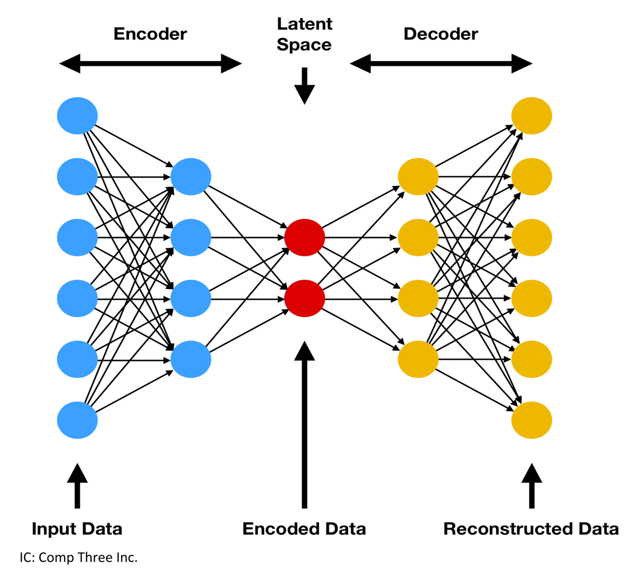
The main idea behind dimensionality reduction is to separate temporal and spatial variations.

The most common data-driven method for achieving this is the Principal Component Analysis (PCA).

$$X(s,t) \approx \sum_{i=1}^{r} c_i(t) U_i(s) \qquad \alpha_{1:r}(t) \approx U(s)_{1:r}^T X(s,t)$$

PCA identifies an optimal but linear subspace. Therefore, for highly nonlinear systems, its applicability is limited.



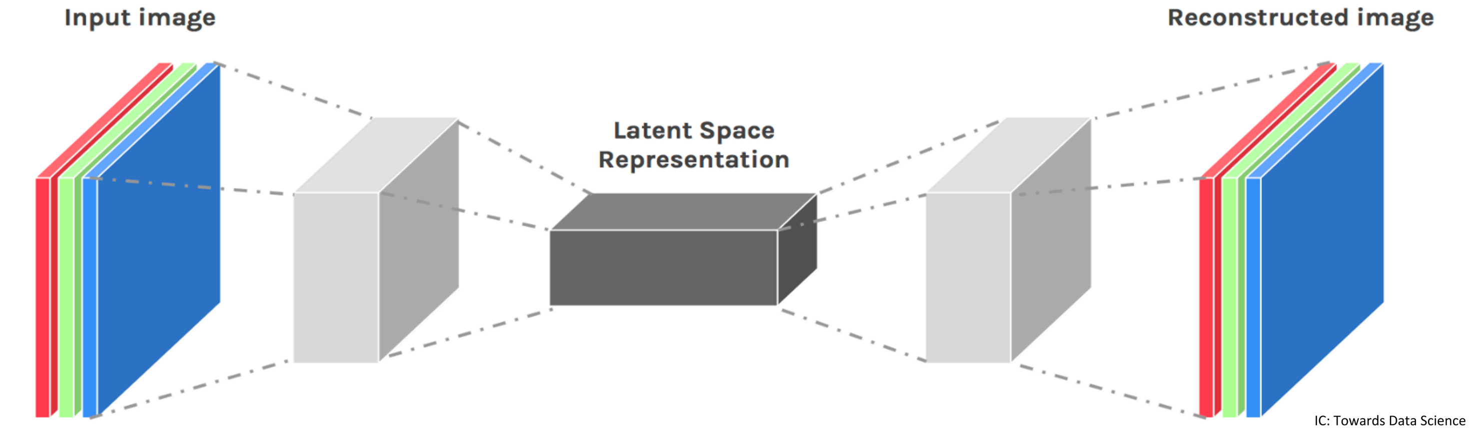


Several methods exist to overcome the linear limitation of PCA (e.g. nonlinear PCA, deep learning, etc.).

We use deep autoencoders (DAE) to achieve dimensionality reduction. The latent state, α, is given by the bottleneck layer of the DAE architecture.

The vanilla autoencoder requires fully connected layers whereby the very large number of trainable model parameters for high-dimensional systems result in intractability.

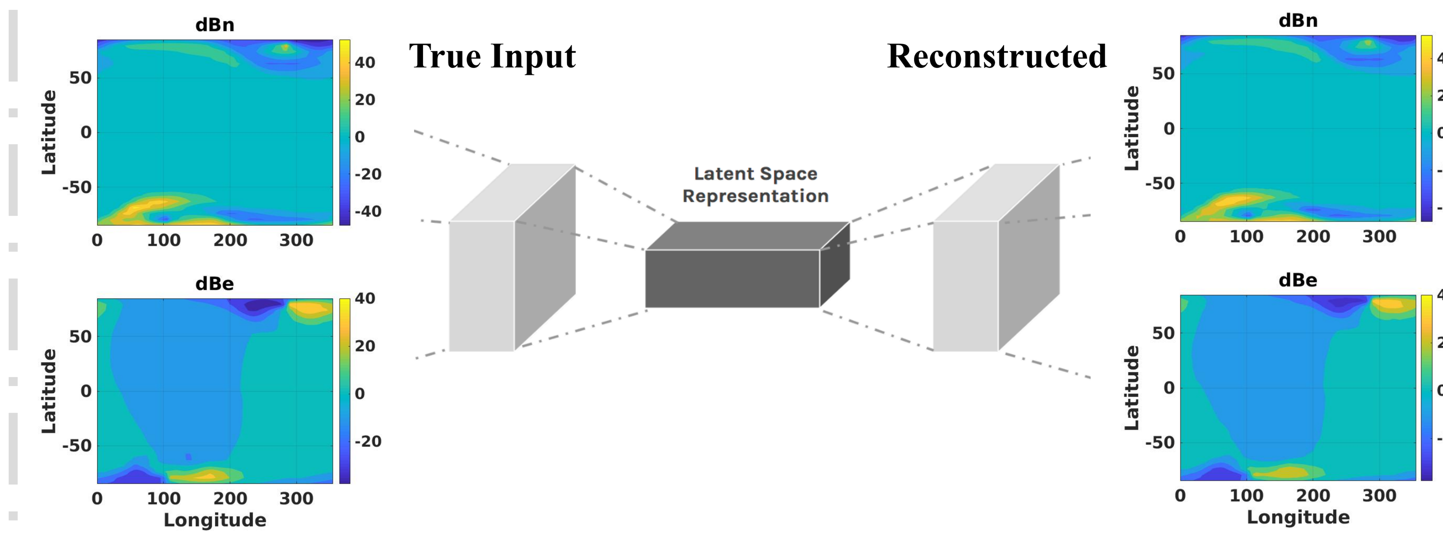
We use the convolutional autoencoder (CAE) to overcome the limitation of fully connected networks.



RESULTS

We perform dimensionality reduction using both PCA and CAE. However, because of the highly nonlinear nature of the system, PCA fails to capture the dynamics. First 30 basis functions capture about 50% of the variance.

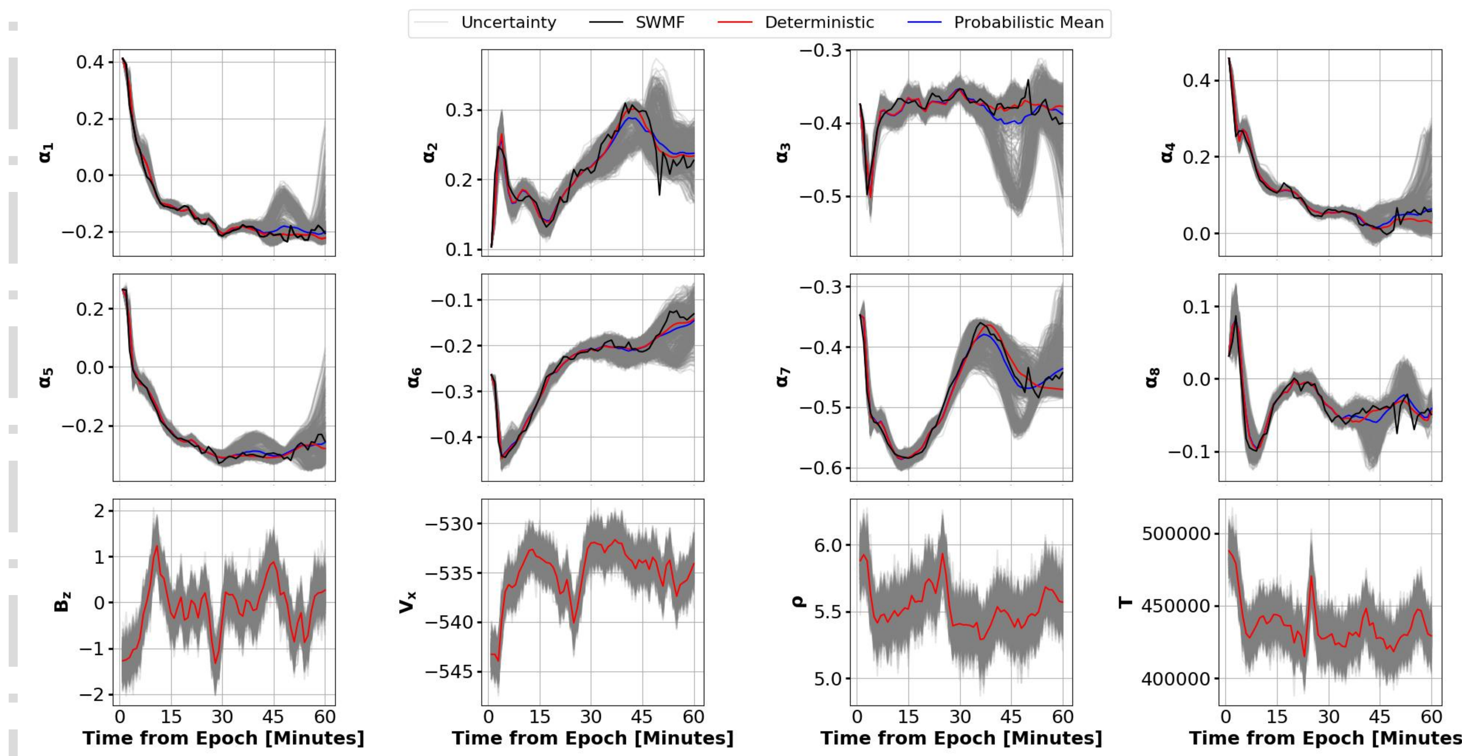
With CAE, we use a latent state dimension of 10 and 25. Preliminary results (non-optimized) show that $r = 25$ performs better (as expected) in reproducing the full state.

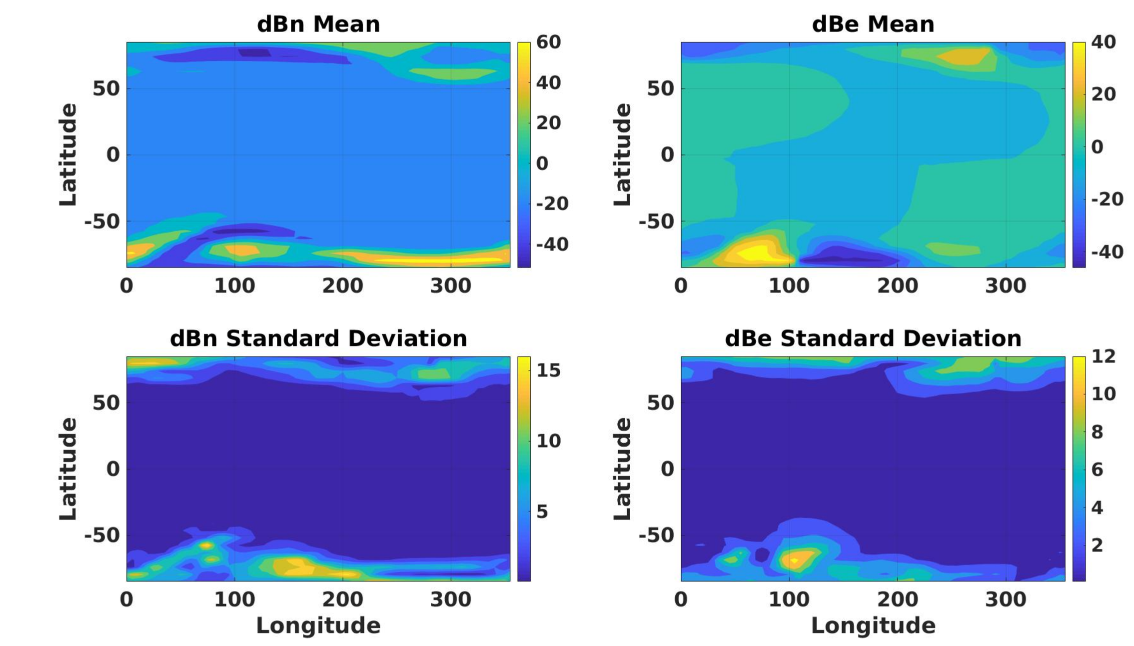


We then use the Long-Short Term Memory (LSTM) network to model the temporal dynamics in the latent state space.

The LSTM model performs well in providing a one hour ahead dynamic prediction accounting for the effects of input uncertainty. Only first eight components are shown to save on space.

Highly nonlinear response of the system to the various input structures highlights the importance of input uncertainty quantification.





The mean grid error at one hour ahead prediction is about 11% (partially optimized). The framework provides the ability to quantify the mean state along with the uncertainty at any given time in the prediction.

ACKNOWLEDGEMENTS

PM and RL acknowledge the support from NSF under grant #1929127. This research was made possible by NASA West Virginia Space Grant Consortium, Training Grant #NNX15AI01H, Training Grant #NNX15AI01H, and NASA Established Program to Stimulate Competitive Research, Grant #80NSSC19M0054. The authors also acknowledge the use of Thorny Flat Super Computing System at WVU, funded in part by the NSF MRI Award #1726534 and WVU.