Physics-informed Machine Learning with Autoencoders and LSTM for Probabilistic Space Weather Modeling and Forecasting

Richard J. Licata[1]; Piyush M. Mehta[1]

1. West Virginia University, Morgantown, WV

Introduction

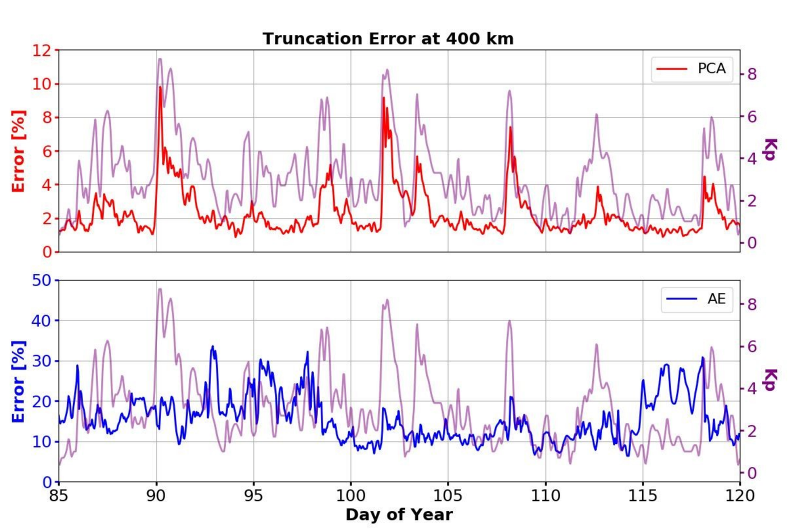
- In low Earth orbit (LEO), atmospheric drag is the largest source of uncertainty when forecasting a satellite's trajectory stemming from incomplete physical knowledge of the domain and uncertain input drivers.
- Physics-based models of the upper atmosphere are high dimensional, even at lower resolutions, inhibiting comprehensive analyses.
- Linear and nonlinear methods exist for dimensionality reduction (e.g. Proper Orthogonal Decomposition, and Deep Autoencoders).
- Low dimension representation allows for computationally inexpensive models to be developed, enabling probabilistic forecasts and supporting model-data fusion.

Methodology

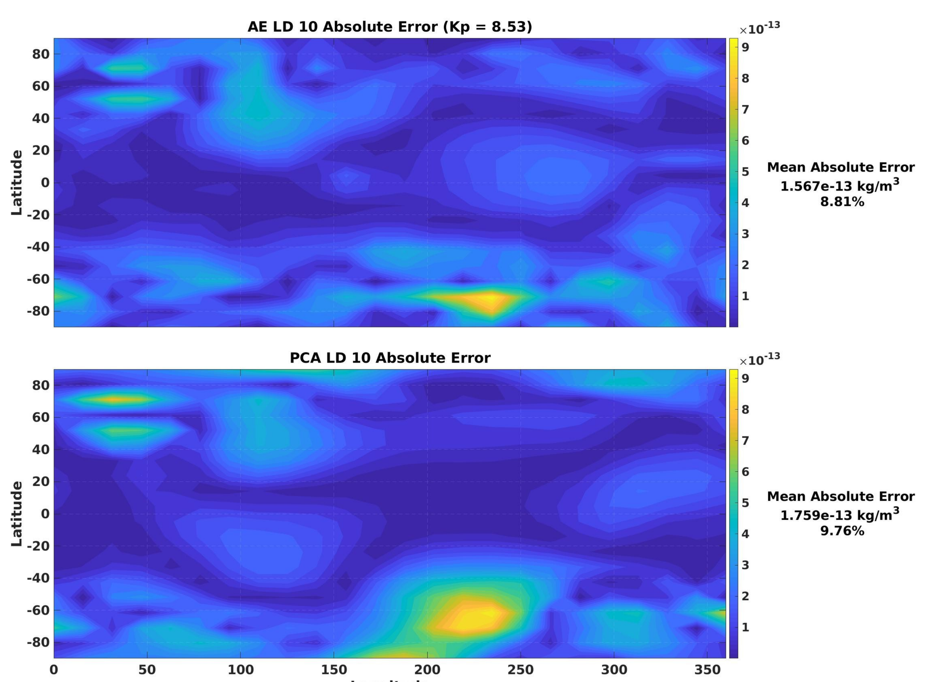
- Dimensionality reduction with Principal Component Analysis (PCA) as previously demonstrated [1].
- Linear and nonlinear Convolutional Autoencoders (CAE) (linear CAE should replicate PCA).
- Analyze truncation error between PCA and CAE.
- Investigate resulting latent space for both methods.

PCA vs Autoencoder

- PCA and CAE are both dimensionality reduction methods aimed to separate spatial and temporal variations.
- CAEs are non-trivial to train but can achieve nonlinear dimensionality reduction, unlike PCA.
- Comparison of PCA and partially trained CAE on 2001 TIE-GCM data (figure below).



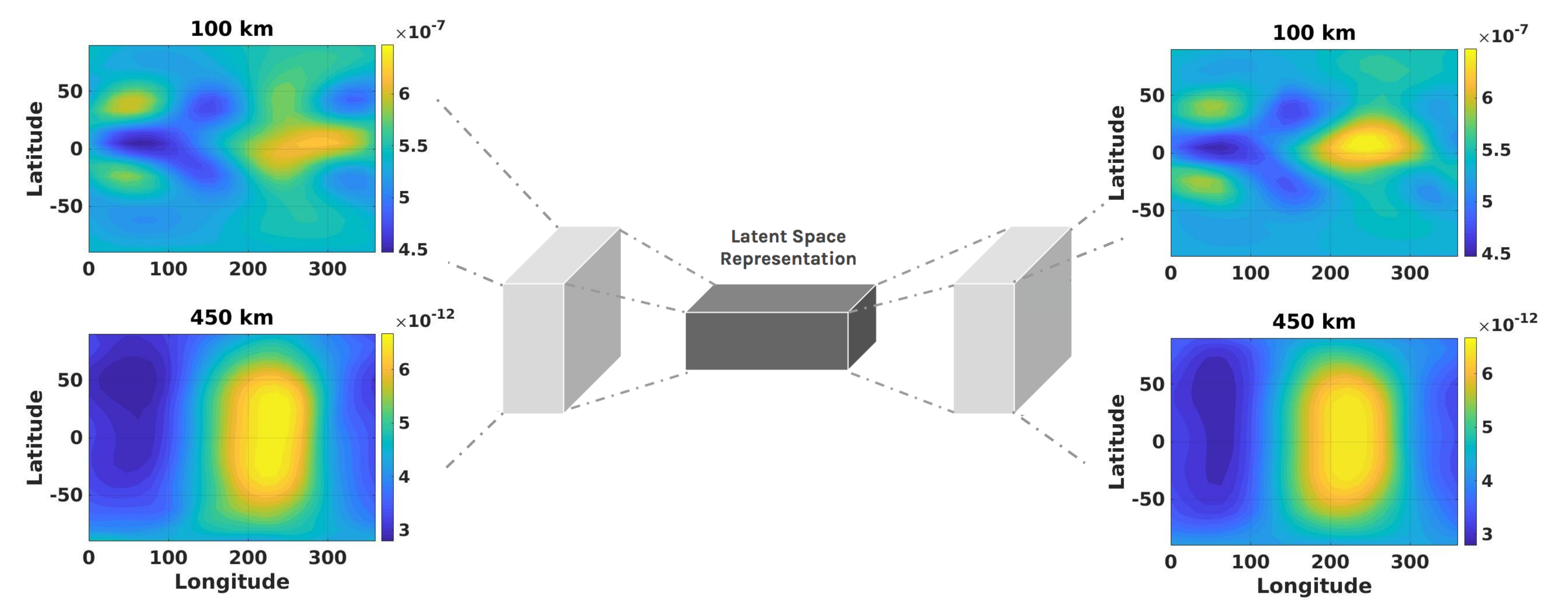
- Because PCA is linear, truncation error jumps by up to 400% during active geomagnetic conditions when the system becomes highly nonlinear.
- **Note: the AE is not currently optimized.**



- Contours show absolute truncation error at 400 km during an extreme storm.

Convolutional Autoencoder

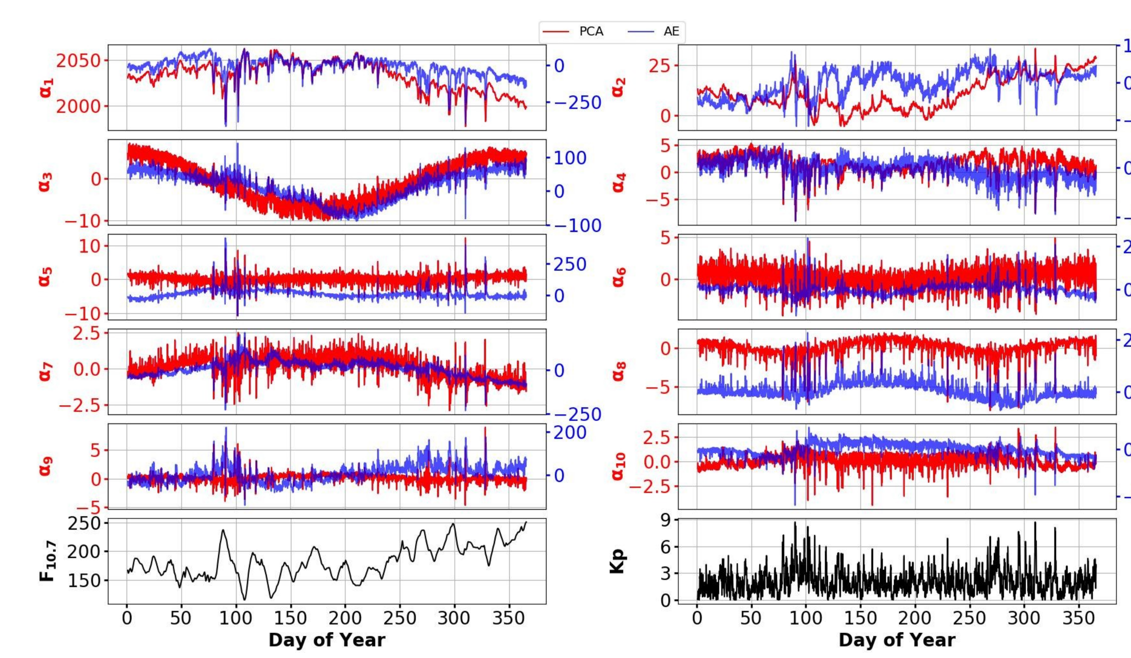
- CAEs can take large datasets (in this problem: 24 x 20 x 16 x time) and reduce it down to a latent dimension of size <10 x time.
- A benefit to using CAEs, over other autoencoders, lies in its fewer model parameters. They are not fully connected resulting in quicker training and decreased likelihood of overfitting.

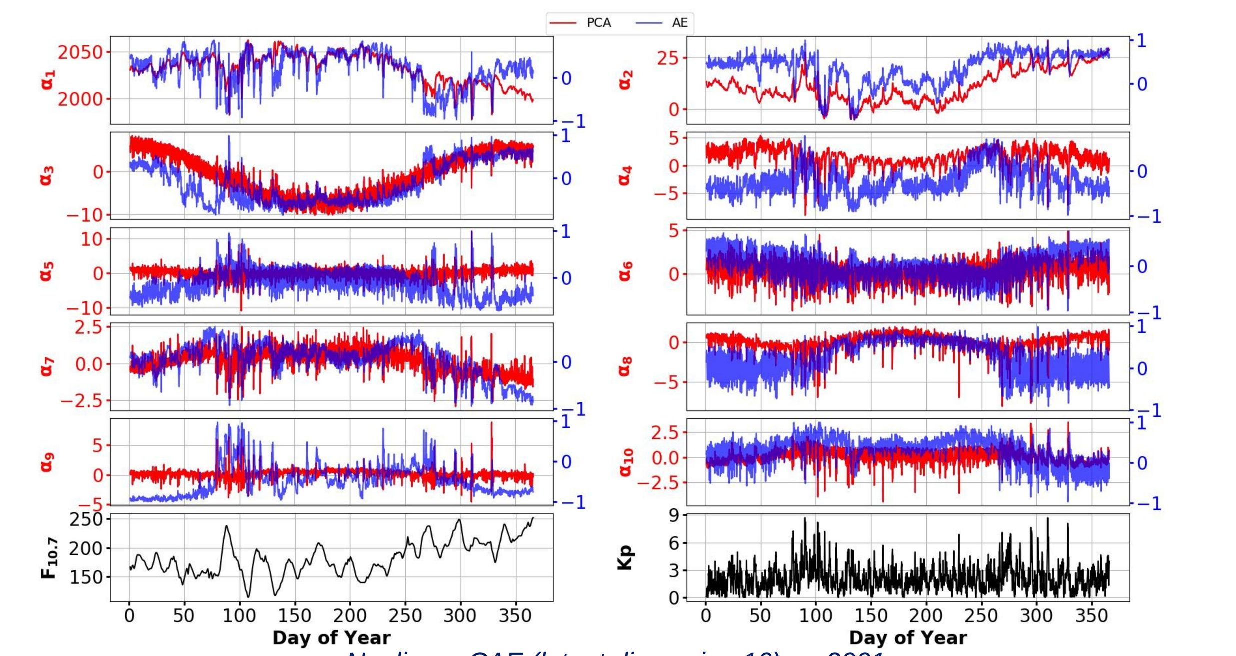


Latent Dimension Analysis

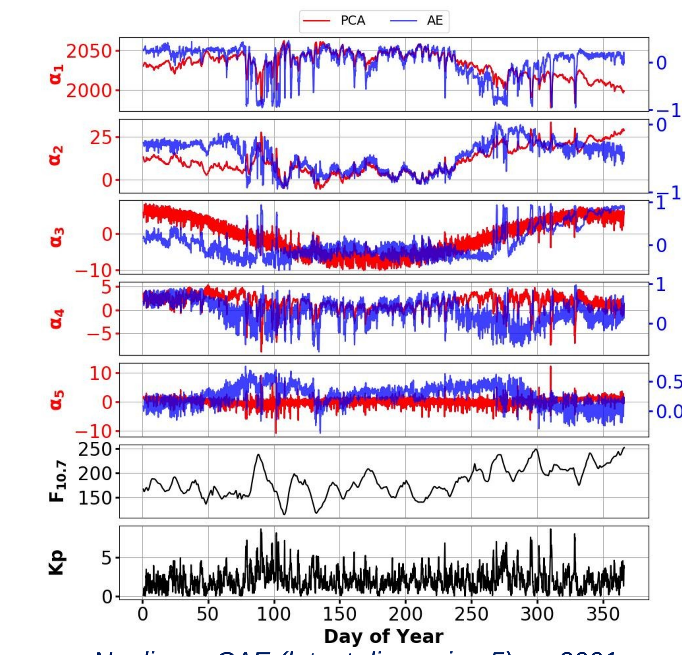
Linear Autoencoder

- The linear CAE mimics PCA.
- PCA provides the optimal order of the latent state, while CAE does not.
- However, having validated CAE for the system, we can introduce nonlinear activation functions for nonlinear dimensionality reduction.





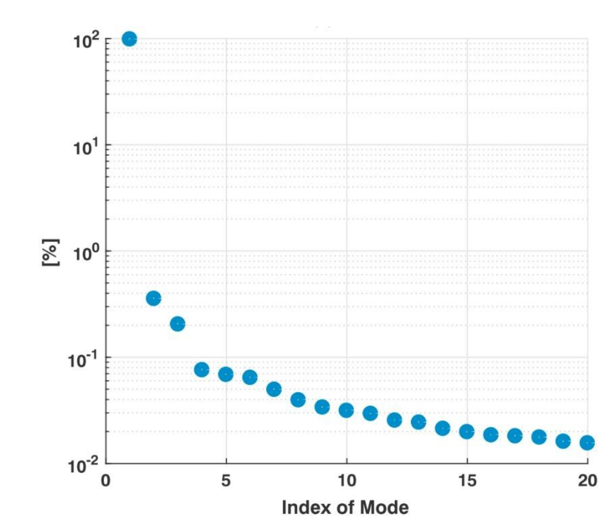
Nonlinear CAE (latent dimension 10) on 2001.



Nonlinear CAE (latent dimension 5) on 2001.

Nonlinear Autoencoder

- Trained nonlinear CAEs with latent dimensions 5 and 10.
- Both capture important physical features (e.g. semiannual trends) while reducing reconstruction error during active geomagnetic periods.
- Most of the energy from the thermosphere can be captured using first few PCA basis functions, therefore we expect the CAE with latent dimension 5 to outperform PCA once optimized.



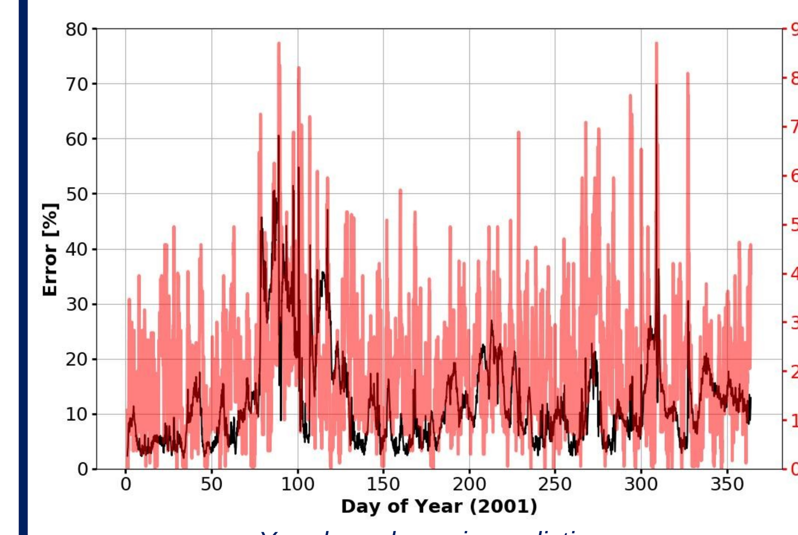
The contribution of the first 20 PCA modes to the total energy [2].

Long Short-Term Memory (LSTM) Neural Network

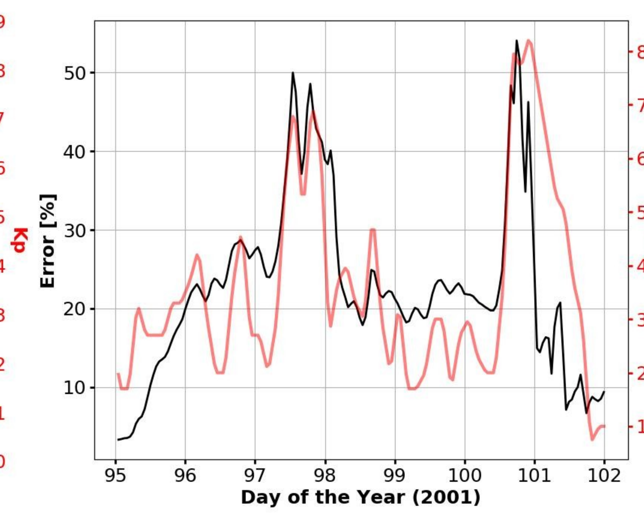
- Previous work used LSTM to learn the temporal mapping, F, for the reduced state, α [3].

$$\boldsymbol{\alpha}_{t+1} = \boldsymbol{F}(\boldsymbol{\alpha}_t, \ldots, \boldsymbol{\alpha}_{t-m}, \mathbf{I}_t, \ldots, \mathbf{I}_{t-m}) + \boldsymbol{\epsilon}_t$$

- The LSTM trained with PCA coefficients shows high errors during geomagnetically active periods when system is highly nonlinear.



Year-long dynamic prediction.



Week-long dynamic prediction.

- The best LSTM at the time could make a dynamic hourly density forecast for an entire year with a mean error of ~12%.
- We expect to improve on this by using CAE for dimensionality reduction (see section on PCA vs Autoencoder).

Conclusions

- In terms of the thermosphere, PCA has relatively low truncation errors during times of low geomagnetic activity.
- During storms, the nonlinear effects are no longer sufficiently captured by PCA.
- Linear AE essentially replicates PCA.
- CAE shows capability to effectively encode nonlinear dynamics.
- **AEs in this work are only partially trained and require further training for optimization of the hyperparameters.**

Acknowledgments

RL was supported through the American Meteorological Society Student Travel Grant. RL and PM 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.

References

1. Piyush M. Mehta and Richard Linares, "A methodology for reduced order modeling and calibration of the upper atmosphere," Space Weather, 5, doi:10.1002/2017SW001642.
2. Piyush M. Mehta, Richard Linares and Eric K. Sutton, "A Quasi-Physical Dynamic Reduced Order Model for Thermospheric Mass Density via Hermitian Space-Dynamic Mode Decomposition," Space Weather, 16, 569-588, doi:10.1029/2018SW00184.
3. Richard J. Licata and Piyush M. Mehta, "Physics-informed Machine Learning for Probabilistic Space Weather Modeling and Forecasting: Thermosphere and Satellite Drag," AGU Fall Meeting, December 2019.