Simultaneously forecasting global magnetic conditions using recurrent networks

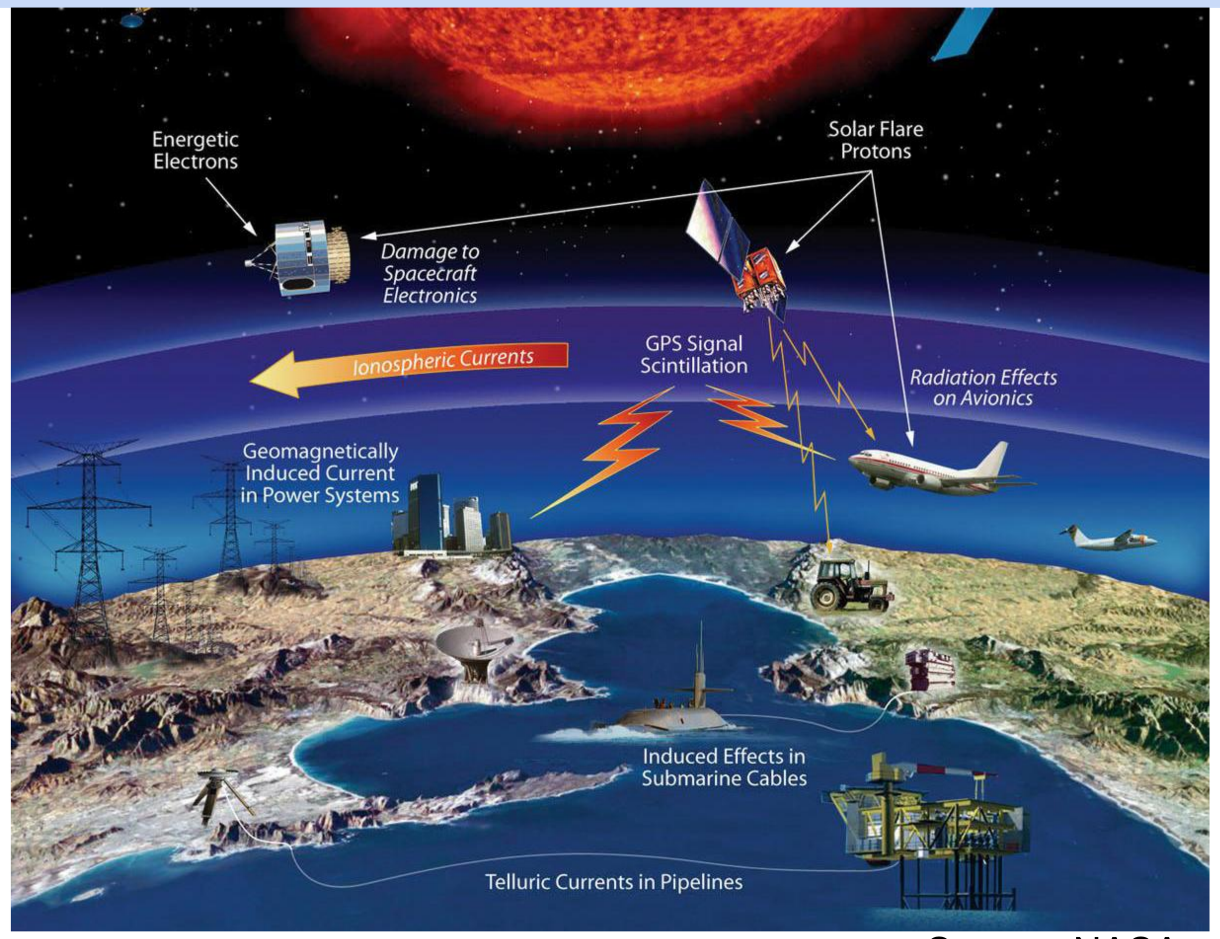
Charles Topliff*, Morris Cohen*, William Bristow†
*Georgia Institute of Technology, †Pennsylvania State University

paper: https://arxiv.org/abs/2010.06487v2

Abstract

- Many systems used by society (GPS, power grid) are vulnerable to extreme space weather
- Existing techniques exist to forecast specific phenomena, but lack global coverage
- We use recurrent neural networks to simultaneously forecast multiple magnetic indices several hours in advance using solar wind measurements
- Improvement over a persistent baseline for all experiments

Background



Source: NASA

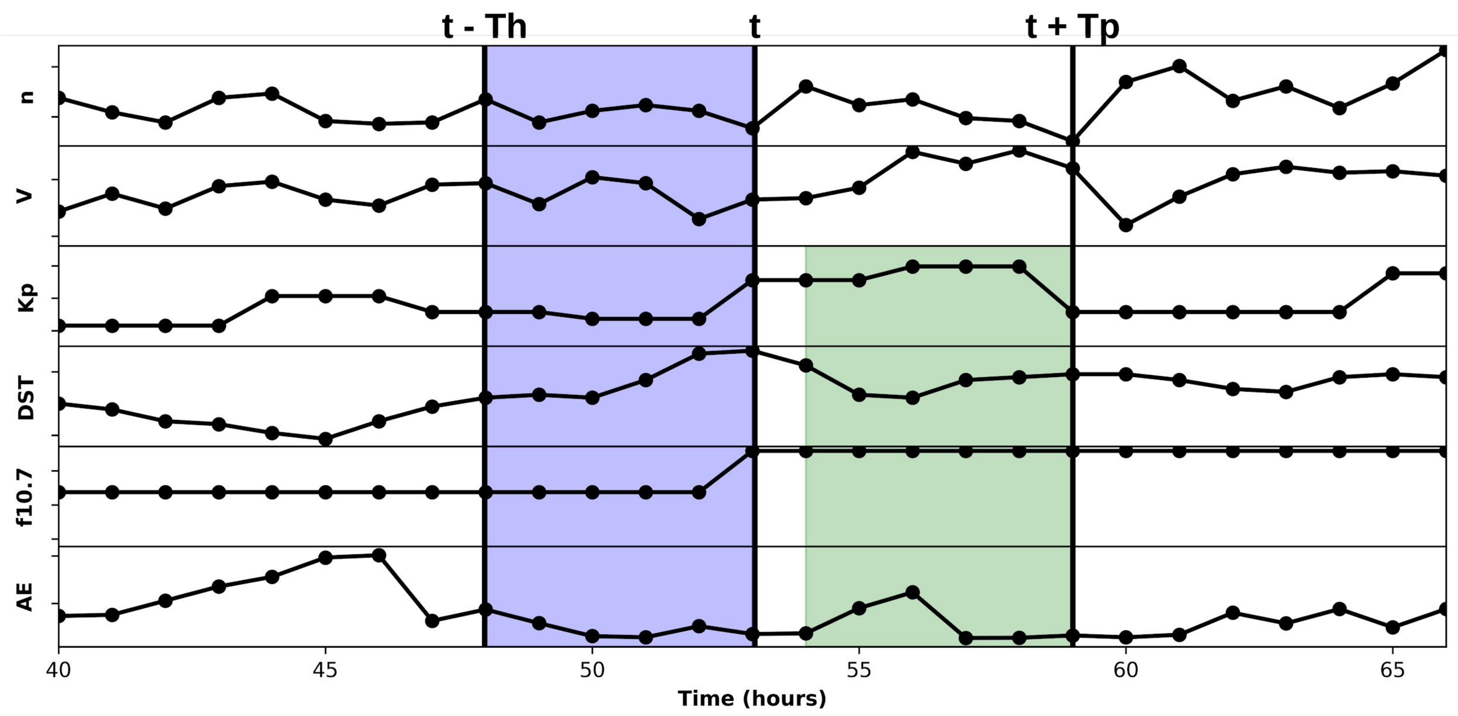
- Magnetic activity at earth is largely driven by processes that begin at the sun
- The Sun periodically emits Coronal Mass Ejections (CMEs) which intensify the solar wind and result in periods of increased magnetic activity at Earth
- Magnetic Indices were invented to summarize magnetic conditions in specific regions of earth:
  - Disturbance Time Index (DST) - summarizes equatorial magnetic activity
  - Auroral Electrojet Indices (AE, AL, AU) - summarizes polar magnetic activity
  - Planetary K (Kp) index - measures strength of magnetic storms and is used to decide on geomagnetic alerts
  - Solar Radio Flux at 10.7cm (F10.7) - measures solar activity

Magnetic Index Forecasting

Treat this as a sequence-to-sequence learning problem using recurrent neural networks

Data

- Magnetic Indices and Solar Wind Measurements downloaded from NASA OMNIWeb
  - 20 years of data dating from 2000 to 2020
  - Interplanetary Magnetic Field (IMF) Measurements, Solar wind speed V, Density n, Hour, Day, Year.
- Also use measurements derived from SuperDARN network of radars [1]



- At time t, use historical indices and solar wind measurements from $[t - T_h, t]$ (blue) to predict magnetic indices from $(t + 1, t + T_p]$ (green)

Model Training

- We use a simple neural network consisting of a multilayer LSTM followed by a linear transformation
  - Perform large scale random search [2] on Learning Rate, Weight Decay, Batch Size, Epochs, LSTM Hidden Dimension using median stopping rule using Ray [3]. All experiments done with PyTorch.

Experiments

- *20 Yr OMNIWeb Data* - We split 20 years of OMNIWeb (no superDARN) data into training, validation, and testing sets corresponding to 2000-2011, 2012-2013, and 2013-2019, respectively.
- *SuperDARN Measurements* - We include the available SuperDARN measurements to measure the added predictive capability to the existing OMNIWeb forecasts (including SuperDARN reduces our dataset size to the years 2013-2017, and we use similar training, testing, and validation proportions as above)

Results

- Baseline is persistent forecast where the prediciton for the entire sequence is the most recent available observation of the magnetic index

20 Yr OMNIWeb Data

	AE		AU		AL		Dst		f10.7		Kp	
Hrs	mNet	pers	mNet	pers	mNet	pers	mNet	pers	mNet	pers	mNet	pers
1	.901	.478	.885	.491	.874	.422	.976	.790	.995	.993	.937	.658
2	.782	.452	.789	.462	.743	.397	.949	.759	.994	.991	.882	.632
3	.696	.429	.715	.437	.656	.375	.918	.730	.993	.990	.825	.605
4	.650	.406	.665	.414	.616	.352	.892	.703	.992	.989	.790	.578
5	.622	.382	.627	.394	.594	.328	.870	.678	.991	.988	.757	.557
6	.598	.364	.598	.379	.574	.309	.850	.655	.990	.987	.726	.535

Table 1: Correlation coefficient of our model (mNet) versus persistent forecast (pers) for all indices

SuperDARN Measurements

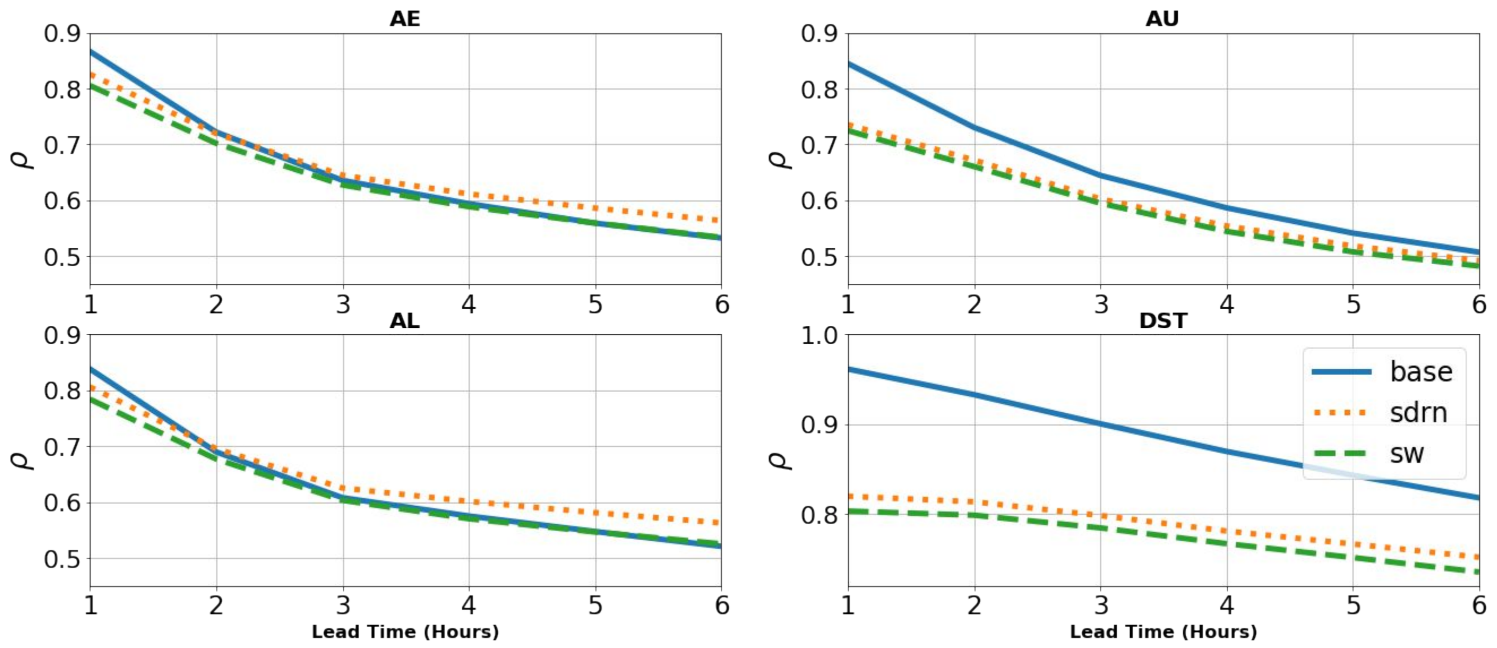


Figure 2: Correlation coefficient for all three models investigated in SuperDARN Measurement experiment (no substantial performance increase)

Future Work

- To improve forecasts, we plan to include coronal image dataset (available at lightspeed) which should contain advance information about CMEs or other activity
- Use an autoencoder for coronal images in combination with current LSTM approach
- Other architecture choices such as Attention Mechanism

[1] WA Bristow and Poul Jensen. A superposed epoch study of superdarn convection observations during substorms.Journal of Geophysical Research: Space Physics, 112(A6), 2007.
[2] James Bergstra and Yoshua Bengio. Random search for hyper-parameter optimization.The Journalof Machine Learning Research, 13(1):281–305, 2012.
[3] Moritz, Philipp, et al. "Ray: A distributed framework for emerging {AI} applications." 13th {USENIX} Symposium on Operating Systems Design and Implementation ({OSDI} 18). 2018.