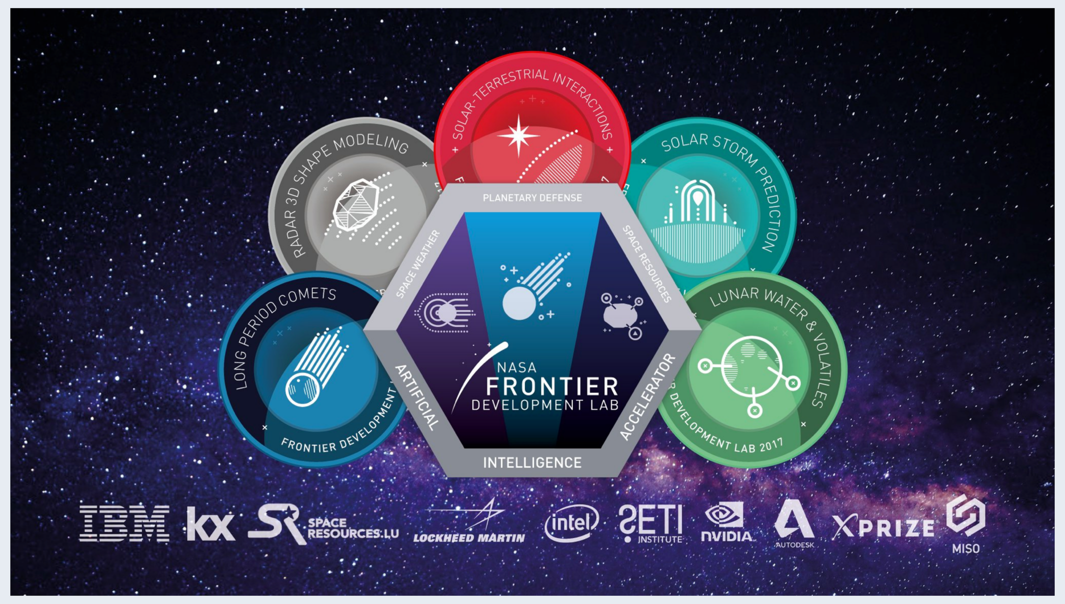
Modeling Geomagnetic Variations using a Machine Learning Framework

Mark C. M. Cheung[1,2,3] (cheung@lmsal.com), Casey Handmer[3], Burçu Kosar[4,3,*], George Gerules[3], Bala Poduval[5,3,*], Graham Mackintosh[3], Andrés Muñoz-Jaramillo[6,3], Monica Bobra[2,3], Troy Hernandez[7,3], Ryan McGranaghan[8,3]

[1]Lockheed Martin Solar & Astrophysics Lab, Palo Alto, CA, USA, [2]Stanford University, CA, USA, [3]NASA Frontier Development Lab, Mountain View, CA, USA, [4]NASA Goddard Space Flight Center, MD, USA, [5]Space Science Institute, Boulder, CO, USA, [6]Southwest Research Institute, Boulder, CO, USA, [7]IBM, Chicago, IL, USA, [8]Jet Propulsion Laboratory, Pasadena, CA, USA, *presenting authors

NASA FRONTIER DEVELOPMENT LAB 2017

- an Artificial Intelligence (AI) R &D accelerator
- to tackle important questions in the space sciences
- an intense 8–week focused study
- topics: Planetary Defense, Space Weather and Space Resources



MACHINE LEARNING/AI

Project 1	Project 2
Recurrent Neural Nets (RNNs) (Geomag + Solar Wind)	Ensemble (Geomag + Solar Wind + Kp)
Long Short Term Memory (LSTM)	Gradient Boosting AdaBoosting Random Forest Gaussian Process Bagging Extra Trees

Keras
TensorFlow
python
learn

Keras: An open source neural network (NN) library written in Python.
Scikit-Learn: A free machine learning library for Python, featuring various classification, regression and clustering algorithms including support vector machines, random forests, gradient boosting, etc.. It is designed to operate with the Python numerical and scientific libraries NumPy and SciPy.
TensorFlow: Another open source software library for machine learning, designed for building and training deep neural networks to detect and decipher patterns and correlations.

Kp INDEX

The K-indices quantify the disturbances in the horizontal component of geomagnetic field, represented by an integer in the range 0-9. It is derived from the maximum fluctuations of horizontal components during three–hour intervals. The planetary index Kp is the mean of standardized K-indices from 13 stations between 44° and 60° N/S geomagnetic latitude. NOAA/Space Weather Prediction Center (SWPC) makes use of the Kp index when issuing geomagnetic storm warnings.

G-Scale	Kp	Activity Level	Occurrence Frequency
G0	4 & lower	Below Storm	
G1	5	Minor Storm	1700 per cycle (900 days per cycle)
G2	6	Moderate Storm	600 per cycle (360 days per cycle)
G3	7	Strong Storm	200 per cycle (130 days per cycle)
G4	8	Severe Storm	100 per cycle (60 days per cycle)
G5	9	Extreme Storm	4 per cycle (4 days per cycle)

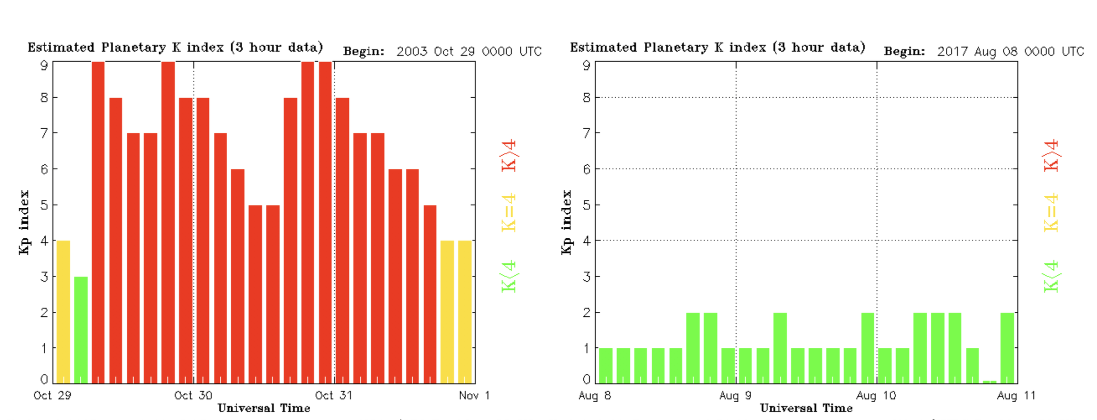


Figure : 1 Kp index during Halloween event (left) and during a very quiet period (right).

THE PROBLEM DEFINITION

(Q1) Can we apply machine learning (ML) to forecast geomagnetic variability using solar wind and ground-based measurements?

(Q2) Without imposing a priori, first-principles based, physical models of the solar wind-driven geomagnetic system, what insights can ML extract from the data?

METRIC OF ACCURACY

We obtained the mean square errors between observed and predicted Kp indices using various models. Also, we computed the p-statistics to determine the statistical significance of how well the models do compared with each other. With $> 95\%$ confidence, the models have different performance metrics.

ML method	1h ahead	3h ahead	6h ahead
Persist	0.007	0.020	0.025
Mean	0.046	0.046	0.046
Median	0.048	0.048	0.048
Gradient Boosting	0.007	0.015	0.021
Adaptive Boost	0.012	0.018	0.032
Extra Trees	0.009	0.021	0.027
Random Forest	0.015	0.015	0.026

> 95% confidence level

SPACE WEATHER EVENTS & THEIR SOCIO-ECONOMIC IMPACTS

- Space Weather: Solar-driven fluctuations in the near-Earth environment leading to disruptions and damages to our critical infrastructure and technological systems in space and on Earth.
- Space Weather events: Solar flares, coronal mass ejections (CMEs), solar energetic particle (SEP) events, solar radio bursts, geomagnetic disturbances
- Space Weather impacts: Disruptions in wireless communications, Global Positioning System (GPS), satellite operations and communication, aviation, and the electrical power grid.
- Space Weather forecast: Using physics-based and empirical models to mitigate the impacts of extreme space weather events (National Space Weather Action Plan - SWAP). Improved predictions offer better protection for space weather stakeholders.

Scale	Description	Effect	Physical measure	Average Frequency (1 cycle = 11 years)
G 5	Extreme	**Power systems:** Widespread voltage control problems and protective system problems can occur, some grid systems may experience complete collapse or blackouts. Transformers may experience damage. **Spacecraft operations:** May experience extensive surface charging, problems with orientation, uplink/downlink and tracking satellites. **Other systems:** Pipeline currents can reach hundreds of amps, HF (high frequency) radio propagation may be impossible in many areas for one to two days, satellite navigation may be degraded for days, low-frequency radio navigation can be out for hours, and aurora has been seen as low as Florida and southern Texas (typically 40° geomagnetic lat.).	Kp = 9	4 per cycle (4 days per cycle)
G 4	Severe	**Power systems:** Possible widespread voltage control problems and some protective systems will mistakenly trip out key assets from the grid. **Spacecraft operations:** May experience surface charging and tracking problems, corrections may be needed for orientation problems. **Other systems:** Induced pipeline currents affect preventive measures, HF radio propagation sporadic, satellite navigation degraded for hours, low-frequency radio navigation disrupted, and aurora has been seen as low as Alabama and northern California (typically 45° geomagnetic lat.).	Kp = 8, including a 9-	100 per cycle (60 days per cycle)
G 3	Strong	**Power systems:** Voltage corrections may be required, false alarms triggered on some protection devices. **Spacecraft operations:** Surface charging may occur on satellite components, drag may increase on low-Earth-orbit satellites, and corrections may be needed for orientation problems. **Other systems:** Intermittent satellite navigation and low-frequency radio navigation problems may occur, HF radio may be intermittent, and aurora has been seen as low as Illinois and Oregon (typically 50° geomagnetic lat.).	Kp = 7	200 per cycle (130 days per cycle)
G 2	Moderate	**Power systems:** High-latitude power systems may experience voltage alarms, long-duration storms may cause transformer damage. **Spacecraft operations:** Corrective actions to orientation may be required by ground control; possible changes in drag affect orbit predictions. **Other systems:** HF radio propagation can fade at higher latitudes, and aurora has been seen as low as New York and Idaho (typically 55° geomagnetic lat.).	Kp = 6	600 per cycle (360 days per cycle)
G 1	Minor	**Power systems:** Weak power grid fluctuations can occur. **Spacecraft operations:** Minor impact on satellite operations possible. **Other systems:** Migratory animals are affected at this and higher levels; aurora is commonly visible at high latitudes (northern Michigan and Maine).	Kp = 5	1700 per cycle (900 days per cycle)

Date	Event	Level
1 Sept 1859	Carrington Event widespread disruption of telegraph	Extreme
13 March 1989	Hydro-Quebec 9 hour black out	Severe
20/21 Jan 1994	Anik-E1 and Anik-E2 failed Disrupted TV and computer transmission	Moderate
14 July 2000	Bastille Day Event	Extreme
31 October 2003	Halloween Events Affected airlines, caused power outages, damaged transformers, led astronauts on ISS to take shelter	Extreme

DATA USED

Period of Study: 2016 (descending phase of Solar cycles 24)

1. Observed solar wind properties: Multispacecraft compilation of solar wind observations at Lagrangian point 1: http://omniweb.gsfc.nasa.gov/. solar wind speed, proton density, heliospheric magnetic field (HMF) intensity, HMF B_z, etc.
2. Geomagnetic field measurements - 14 US stations operated by US Geological Survey.
3. Kp (planetary K) index.

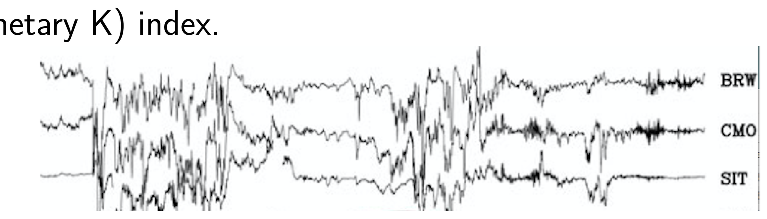


Figure : 3a Geomagnetic field data.

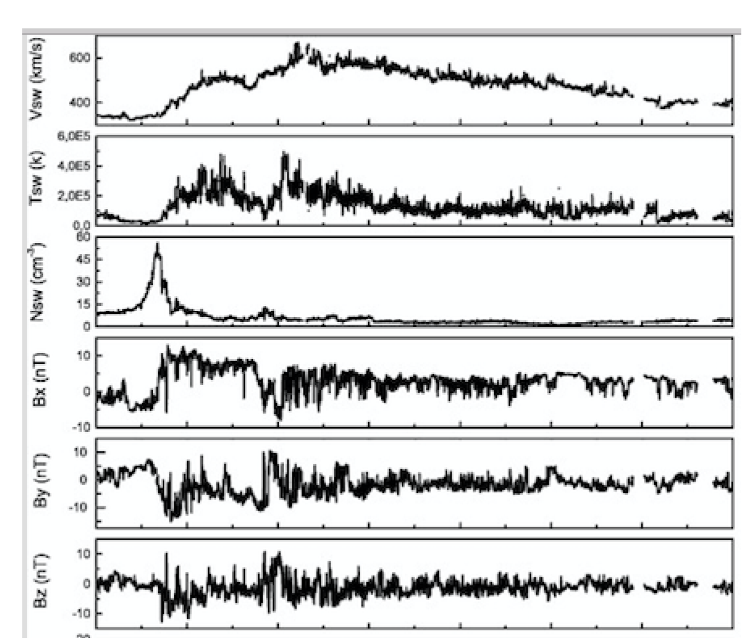


Figure : 3b Solar wind data.

PREDICTED Kp

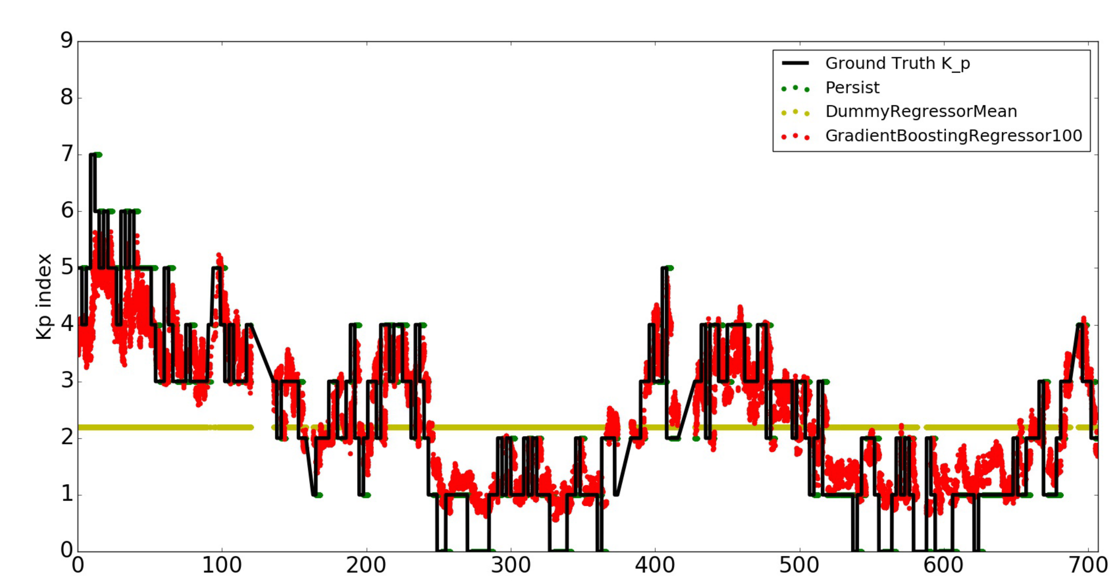
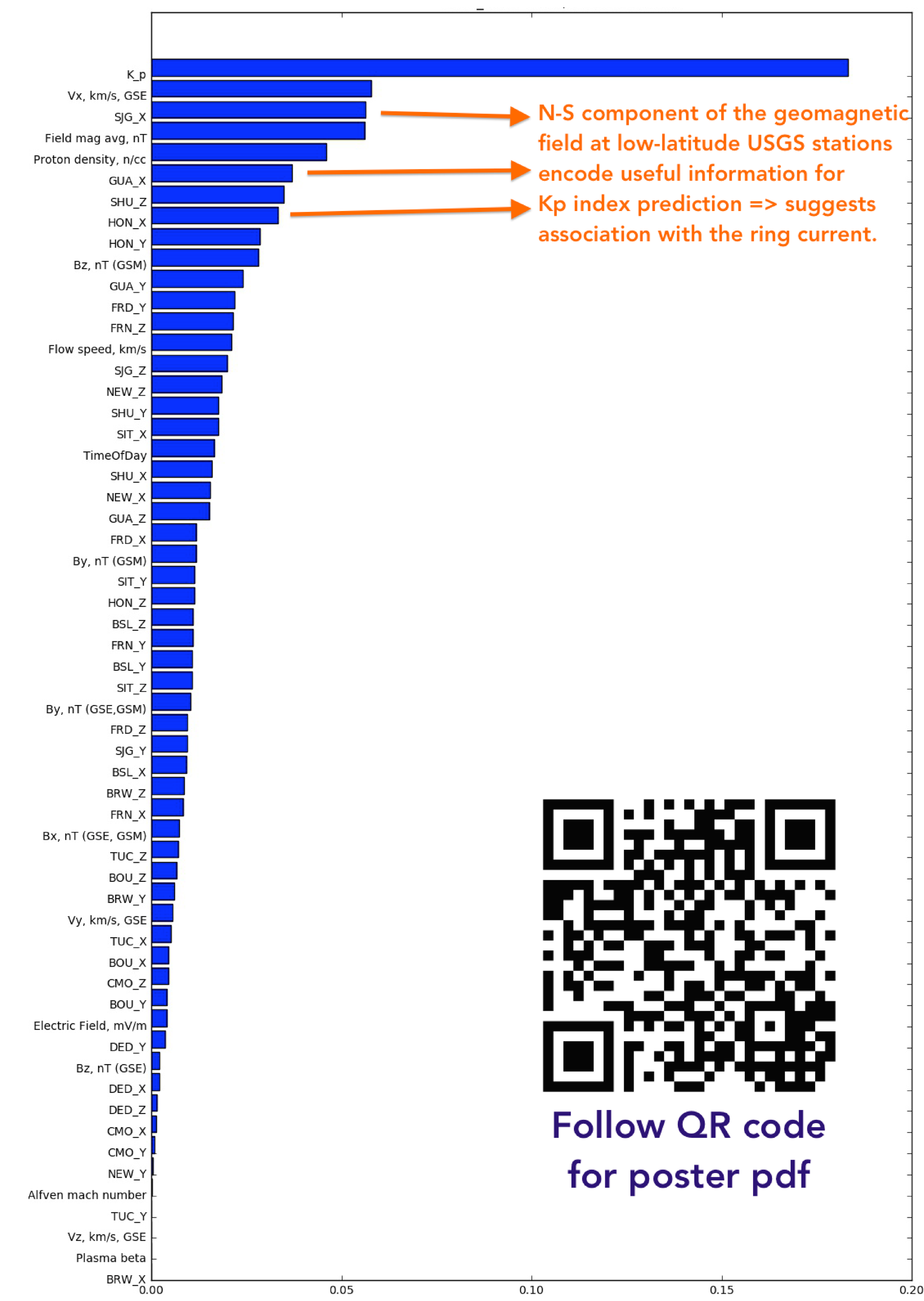


Figure : 4 Actual observed Kp (calculated from ground observations) (black dots) and corresponding values from 3-hr ahead forecast using a persistence (dark green dots), global mean (light green dots) and gradient boosting (yellow dots) models.

We used nearly 7 months of data to train the model and then tested the model by predicting the Kp indices for 3 months (Figure 4 shows a subset of the test data). The training and testing data were partitioned such that the models have not seen data with any overlap between the two sets.
The Gradient Boosting Regressor model provided the best results, consistently beating a persistence model (i.e. the current Kp index predicted not to change in the future) and various machine learning models in scikit-learn, with a confidence level $> 95\%$.
Also, the Gradient Boosting model ranks the input features by their relative importance for creating a good prediction (Figure 5).

RELATIVE IMPORTANCE OF INPUT PARAMETERS



Follow QR code for poster pdf

Figure : 5 Relative importance of input parameters in the prediction of Kp using the Gradient Boosting Regressor model.

SUMMARY & CONCLUDING REMARKS

Without prior domain knowledge, the model learned that the most important precursor is the current Kp index.
Other important factors:
- Solar wind speed and proton density.
- Solar wind magnetic field strength and Bz.

Moreover, the model suggested that the N-S component of the geomagnetic field at low latitude stations - Guam (GUA), Hawaii (HON), Puerto Rico (SJG), are also important precursors. These quantities are largely influenced by ring current and therefore, this finding implies the importance of considering the effects of ring current in the prediction of geomagnetic storm. This result came as a total surprise since the machine learning algorithm was not expected to be capable of learning such heuristics without prior knowledge!
Scope: Based on the results we feel confident that the method can be applied to address other aspects of the socio-economic impact of space weather by predicting the appropriate variable if sufficient data exist.
Ultimate goal: To couple the complex and dynamic solar-terrestrial system using AI.

ACKNOWLEDGEMENTS

Thanks are also due to IBM for providing the computing resources for the project and NVIDIA, Intel, Lockheed Martin and Kx for their partnership in the FDL 2017.
B. Poduval wishes to acknowledge Dr. J. Love and colleagues, USGS; Dr. Jennifer Ganon, CPI and Robert Steenburgh, NOAA, for the many discussions helpful for this project.

ACRONYMS

AI	Artificial Intelligence	GB	Gradient Boosting Regressor
ML	Machine Learning	NN	Neural Network
RMSE	Root Mean Square Error	Kp index	Planetary K index
K	'Kennziffer' for 'characteristic digit.'		