Improved Neutral Density Predictions through Machine Learning Enabled Exospheric Temperature Model

Richard Joseph Licata[1], Piyush M Mehta[1], Daniel R Weimer[2], and W. Kent Tobiska[3]

[1]West Virginia University
[2]Virginia Tech
[3]Space Environment Technologies

November 23, 2022

Abstract

The community has leveraged satellite accelerometer datasets in previous years to estimate neutral mass density and subsequently exospheric temperatures. We utilize derived temperature data and optimize a nonlinear machine-learned (ML) regression model to improve upon the performance of the linear EXTEMPLAR (EXospheric TEMPeratures on a PoLyherdrAl gRid) model. The newly developed EXTEMPLAR-ML model allows for exospheric temperature predictions at any location with a single model and provides performance improvements over its predecessor. We achieve a 4.2 K reduction in mean absolute error and a 3.42 K reduction in the standard deviation of the error. Like EXTEMPLAR, our model's outputs can be utilized by the Naval Research Laboratory Mass Spectrometer and Incoherent Scatter radar Extended (NRLMSISE-00) model to more closely match satellite accelerometer-derived densities. We conducted two case studies where we compare the CHAllenging Minisatellite Payload (CHAMP) and Gravity Recovery and Climate Experiment (GRACE) accelerometer-derived temperature and density estimates to NRLMSISE-00, EXTEMPLAR, and EXTEMPALR-ML during two major storm periods. The storm-time temperature comparison showed error reductions of 7-10% and 2-5% relative to NRLMSISE-00 and EXTEMPLAR, respectively, and the density comparison showed error reductions of 20-55% and 8-12%. We use Principal Component Analysis to identify the dominant modes of variability in the model over one solar cycle. This shows the model is dominantly driven by solar activity, and there is a strong latitudinal variation related to the Summer and Winter hemispheres.

Posted on 23 Nov 2022 — CC-BY 4.0 — https://doi.org/10.1002/essoar.10507687.1 — This a preprint and has not been peer reviewed. Data may be preliminary.

Improved Neutral Density Predictions through Machine Learning Enabled Exospheric Temperature Model

Richard J. Licata[1]*, Piyush M. Mehta[1], Daniel R. Weimer[2], and W. Kent Tobiska[3]

[1]Department of Mechanical and Aerospace Engineering, West Virginia University, Morgantown, West Virginia, USA.
[2]Center for Space Science and Engineering Research, Virginia Tech, Blacksburg, VA, USA
[3]Space Environment Technologies, Pacific Palisades, California, USA.

Key Points:

- We develop a nonlinear global model for exospheric temperature prediction called EXTEMPLAR-ML.
- We leverage Principal Component Analysis to improve our understanding of the EXTEMPLAR-ML temperature formulation.
- EXTEMPLAR-ML shows increased accuracy relative to satellite observations during strong geomagnetic storms.

*1306 Evansdale Drive, Morgantown, West Virginia 26506-6106.

Corresponding author: Richard J. Licata, rjlicata@mix.wvu.edu

Abstract

The community has leveraged satellite accelerometer datasets in previous years to estimate neutral mass density and subsequently exospheric temperatures. We utilize derived temperature data and optimize a nonlinear machine-learned (ML) regression model to improve upon the performance of the linear EXTEMPLAR (EXospheric TEMPeratures on a PoLyherdrAl gRid) model. The newly developed EXTEMPLAR-ML model allows for exospheric temperature predictions at any location with a single model and provides performance improvements over its predecessor. We achieve a 4.2 K reduction in mean absolute error and a 3.42 K reduction in the standard deviation of the error. Like EXTEMPLAR, our model's outputs can be utilized by the Naval Research Laboratory Mass Spectrometer and Incoherent Scatter radar Extended (NRLMSISE-00) model to more closely match satellite accelerometer-derived densities. We conducted two case studies where we compare the CHAllenging Minisatellite Payload (CHAMP) and Gravity Recovery and Climate Experiment (GRACE) accelerometer-derived temperature and density estimates to NRLMSISE-00, EXTEMPLAR, and EXTEMPALR-ML during two major storm periods. The storm-time temperature comparison showed error reductions of 7-10% and 2-5% relative to NRLMSISE-00 and EXTEMPLAR, respectively, and the density comparison showed error reductions of 20-55% and 8-12%. We use Principal Component Analysis to identify the dominant modes of variability in the model over one solar cycle. This shows the model is dominantly driven by solar activity, and there is a strong latitudinal variation related to the Summer and Winter hemispheres.

Plain Language Summary

Density in the upper atmosphere is highly variable and difficult to model. Empirical density models often rely on temperature profile predictions to determine species and mass densities. One of three key parameters in determining the temperature profiles is the asymptotic value at the top of the thermosphere, called the exospheric temperature. By using temperatures derived from satellite acceleration measurements, we develop a machine-learned global temperature model called EXospheric TEMPeratures on a PoLyherdrAl gRid Machine Learned (EXTEMPLAR-ML). We achieve a 4.2 K reduction in mean absolute error and a 3.42 K reduction in the standard deviation of the error relative to the model's predecessor. We also look at temperatures and densities along satellite orbits during two major geomagnetic storms from the 21st century. In this study, we see major improvements over a significant empirical model called NRLMSISE-00 and the linear predecessor to EXTEMPLAR-ML. We also use a mathematical decomposition tool on the model outputs to assess its internal formulation. This shows that EXTEMPLAR-ML is most heavily driven by solar activity and the seasons.

1 Introduction

Thermospheric mass density modeling is vital to satellite tracking and orbit prediction, yet it remains a formidable task for researchers and operators. The thermosphere is highly driven by external forcings, such as space weather events, and complex internal dynamics. The primary external driver to the thermosphere is solar irradiance (*Qian and Solomon* 2011). A majority of the solar irradiance energy input to the thermosphere can be captured with various solar indices and proxies (*Bowman and Tobiska* 2006). Although, these model drivers become less effective during solar minimum where other processes (e.g. composition changes) become increasingly relevant (*Bowman et al.* 2008, *Mehta et al.* 2019). Other space weather events, such as coronal mass ejections and solar flares, can send mass and energy towards Earth. This interacts with the magnetosphere resulting in Joule heating and particle precipitation which can cause large, sudden changes in density (*Fedrizzi et al.* 2012, *Deng et al.* 2013).

Nitroc oxide (NO) is a cooling mechanism responsible for long-term cooling trends present during solar minimum (*Kockarts* 1980), and short-term temperature decreases following large geomagnetic storms (*Mlynczak et al.* 2003, *Knipp et al.* 2017). Lei et al (2012a) found that for the 2003 Halloween storm, temperature and density post-storm were appreciably lower than pre-storm levels. Many empirical models do not model this phenomena well and predict higher density in the recovery phase of major storms relative to observations (*Oliveira and Zesta* 2019, *Licata et al.* 2021b)

The Naval Research Laboratory Mass Spectrometer and Incoherent Scatter radar Extended (NRLMSISE-00 but referred to in this paper as MSIS) is a commonly used empirical thermospheric density model (*Picone et al.* 2002). As with many models (e.g. DTM (*Bruinsma* 2015) and JB2008 (*Bowman et al.* 2008)), MSIS heavily relies on temperature profiles to determine species densities and therefore mass density throughout the thermosphere. A key parameter in predicting the temperature profile is the exospheric temperature (T_∞) which is the asymptotic value that the temperature profile approaches at the top of the thermosphere, or thermopause (*Bates* 1959, *Jacchia* 1965). MSIS uses the Bates-Walker temperature profile (*Walker* 1965).

The availability of accelerometer-derived density estimates, from satellite such as CHAllenging Minisatellite Payload (CHAMP) and Gravity Recovery and Climate Experiment (GRACE), has been advantageous for model development and assessment (*Luhr et al.* 2002, *Bettadpur* 2012). Over the lifetime of satellites with onboard accelerometers, we accumulate measurements over an abundance of locations and space weather conditions. Researchers have used these measurements to derive density estimates by removing accelerations from other sources (*Sutton* 2008, *Doornbos* 2012, *Calabia and Jin* 2016, *Mehta et al.* 2017). Weimer et al. (2016) used the density estimates from Mehta et al. (2017) to approximate exospheric temperatures by varying the parameter in MSIS using the bisection method until the model density closely matched that of the satellite. Weng et al. (2017) followed this methodology and used Sutton's CHAMP density estimates to create an exospheric temperature model.

Weimer et al. (2020) had used the derived exospheric temperatures to fit 1,620 linear models to make predictions on a polyhedral grid as a function of different space weather conditions over time. The model is called EXTEMPLAR (EXospheric TEMPeratures on a PoLyherdrAl gRid). In this work, we develop an improved exospheric temperature model by using a single nonlinear artificial neural network (ANN) to make predictions at any location. This global model is called EXTEMPLAR Machine Learned (EXTEMPLAR-ML).

Principal Component Analysis (PCA), also referred to as Empirical Orthogonal Function (EOF) analysis or Proper Orthogonal Decomposition (POD), is used in this work to investigate the most dominant modes of variability in EXTEMPLAR-ML. PCA has been used to analyze thermospheric density datasets previously and is often used in the development of reduced-order models (*Mehta and Linares* 2017, *Mehta et al.* 2018, *Gondelach and Linares* 2020). PCA has also been used to study satellite accelerometer datasets (*Matsuo and Forbes* 2010, *Lei et al.* 2012b, *Calabia and Jin* 2016). Sutton et al. (2012) used PCA to produce basis functions that represented the variability of temperature parameters used in an empirical Jacchia family model to improve its nominal density formulation (*Jacchia* 1970). Ruan et al. (2018) used CHAMP density estimates and a physics-based density model to develop an exospheric temperature model based in PCA. ML models tend to be ambiguous in nature, so we utilize PCA only to improve our understanding of the physical processes that drive EXTEMPLAR-ML, not for model development.

The paper is organized as follows, we start by detailing the model development. Then, we discuss the methodology for temperature and density prediction using the model. After, we look at a baseline global temperature map to compare to the preceding model. We then investigate the dominant modes and PCA coefficients across one solar cycle. As a case study, we compare the temperature and density predictions of MSIS, EXTEMPLAR and EXTEMPLAR-ML to CHAMP and GRACE-A during two major geomagnetic storms.

2 Methodology

2.1 Model Development

We had access to over 81 million exospheric temperature estimates from Weimer et al. (2020), the associated polyhedral grid locations, and different space weather indices/proxies as potential drivers. The best linear model from previous work used S_{10}, $\sqrt{M_{10}}$, Poynting flux totals (S_N and S_S), a temperature perturbation term (ΔT), day of year (doy), and universal time (UT). The cooling effect of Nitric Oxide emissions was simulated in the calculation of ΔT. We set out to use only operational indices for EXTEMPLAR-ML so that it could be used in real time. Therefore, we use S_{10}, M_{10}, S_N, S_S, local solar time (LST), geodetic latitude, doy, and UT. The S_{10} and M_{10} indices are representative of solar activity and are part of Space Environment Technologies' (SET) *SOLAR2000* algorithm (*Tobiska et al.* 2000) which has been recently benchmarked by Licata et al. (2020b). Descriptions of these indices are thoroughly explained by Tobiska et al. (2008). The Poynting flux values are calculated using an electrodynamics model (referred to as the W05 model) described by Weimer (2005a, 2005b). For doy and UT, the model uses sine and cosine functions of the fractional doy and UT, generating four temporal inputs (t_1-t_4), see Equation 1.

$$t_1 = sin\left(\frac{2\pi doy}{365.25}\right) \quad t_2 = cos\left(\frac{2\pi doy}{365.25}\right) \quad t_3 = sin\left(\frac{2\pi UT}{24}\right) \quad t_4 = cos\left(\frac{2\pi UT}{24}\right) \tag{1}$$

Upon having set up the data into inputs (described above) and labels (associated $\log_{10}T_\infty$), we leverage a tool called Keras Tuner (*O'Malley et al.* 2019). This allows us to provide a range of hyperparameters upon which the tuner searches to find the best architecture/model through a Bayesian optimization scheme. The tuner settings are shown in Table 1. The tuner is provided 1 million random training samples and 200,000 validation samples.

Table 1. Hyperparameter search space for EXTEMPLAR-ML tuner.

Parameter	**Choices**
Number of Hidden Layers	1 – 10
Neurons	min = 64, max = 1024, step = 4
Activations	relu, softplus, tanh, sigmoid, softsign, selu, elu, linear
Dropout	min = 0.01, max = 0.50, step = 0.01
Optimizer	RMSprop, Adam, Adadelta, Nadam

Once complete, the tuner returns the ten best models, which we evaluate on independent data to confirm model performance. The best architecture to come out of the tuner is displayed in Table 2. This model was trained further using 60 million random samples, with the remaining 21 million used as validation/test data. Once the final model is developed, we test its validity by first comparing its global temperature maps to that of its predecessor. This is to check for anomalies in the temperature distributions for a given condition. While EXTEMPLAR-ML is not restricted to prediction at the polyhedral grid locations, it still contains the name EXTEMPLAR, because it is trained on temperatures that are binned to those locations.

Table 2. Model architecture for the best model from the EXTEMPLAR-ML tuner. There are 10 inputs for Layer 1.

	Neurons	Activation	Dropout Rate
Layer 1	780	elu	0.23
Layer 2	584	tanh	0.13
Layer 3	336	softplus	0.13
Output	1	linear	0.00

2.2 Principal Component Analysis

Principal Component Analysis is an eigendecomposition technique that determines uncorrelated linear combinations of the data that maximize variance (*F.R.S.* 1901, *Hotelling* 1933). As mentioned in the Introduction, PCA is widely used in thermospheric density and exospheric temperature studies as both a modeling and analytical tool. We use PCA to get insight into EXTEMPLAR-ML, which requires predictions covering a vast array of conditions. To accomplish this, we evaluated the model at all 1,620 EXTEMPLAR grid locations between the solar maximums of solar cycle 23 and 24 (~2002-2014) at a three hour cadence. These predictions provide the global evolution of exospheric temperatures spannig a solar cycle. We perform PCA on the spatiotemporal temperature maps to obtain the U, Σ, and V matrices. PCA decomposes the data and separates spatial and temporal variations such that:

$$\mathbf{x}\left(\mathbf{s}, t\right) = \sum_{i=1}^{r} \alpha_i\left(t\right) U_i\left(s\right) \tag{2}$$

where $\mathbf{x} \in \mathbb{R}^n$ is the model output state (full 2D temperature maps), r is the choice of order truncation, α_i are temporal coefficients, and U_i are orthogonal modes or basis functions. The modes are the first r columns of the left singular vector derived by performing PCA on an ensemble of model output solutions such that:

$$\mathbf{X} = \begin{bmatrix} | & | & | & & | \\ \mathbf{x_1} & \mathbf{x_2} & \mathbf{x_3} & \ldots & \mathbf{x_m} \\ | & | & | & & | \end{bmatrix} \quad \text{and} \quad \mathbf{X} = U\Sigma V^T \tag{3}$$

In Equation 3, m represents the ensemble size (one solar cycle). The temperature data is denoted by $\mathbf{X}$. U is the left unitary matrix, and it is made of orthogonal vectors that represent the modes of variation. Σ is a diagonal matrix consisting of the squares of the eigenvalues that correspond to the vectors in U. We can extract temporal coefficients by performing matrix multiplication between Σ and V^T. Therefore, the signs of the modes and coefficients are important in the analysis phase.

2.3 Geomagnetic Storm Case Study

We look at a 48 hour period between 12:00 UT on July 26, 2004 to 12:00 UT on July 28, 2004. This encompasses a day-long geomagnetic storm with a_p peaking at 300 $2nT$. We predict exospheric temperatures along the orbits of CHAMP and GRACE-A and compare it to the satellite-derived temperatures, MSIS, and the best linear EXTEMPLAR model. We perform a similar comparison across the three-day period of October 29-31, 2003. This was one of the most significant geomagnetic storms of the 21st century, with a_p reaching 400 $2nT$ on two separate occasions. We then look at the CHAMP and GRACE-A densities plotted against MSIS, MSIS + EXTEMPLAR, and MSIS + EXTEMPLAR-ML for the same two periods. The density values for EXTEMPLAR and EXTEMPLAR-ML are obtained from MSIS while bypassing the normal T_∞ calculation within MSIS, using the models' temperature outputs instead, as described by Weimer et al. (2020).

3 Results

Once we obtained the model described in Section 2.1, we evaluated it on all training and validation data. This is shown in Figure 1. The top panel shows the original and predicted T_∞ values in a scatter plot with a background contour showing the absolute error. The bottom panel shows the error distributions for training and validation samples.

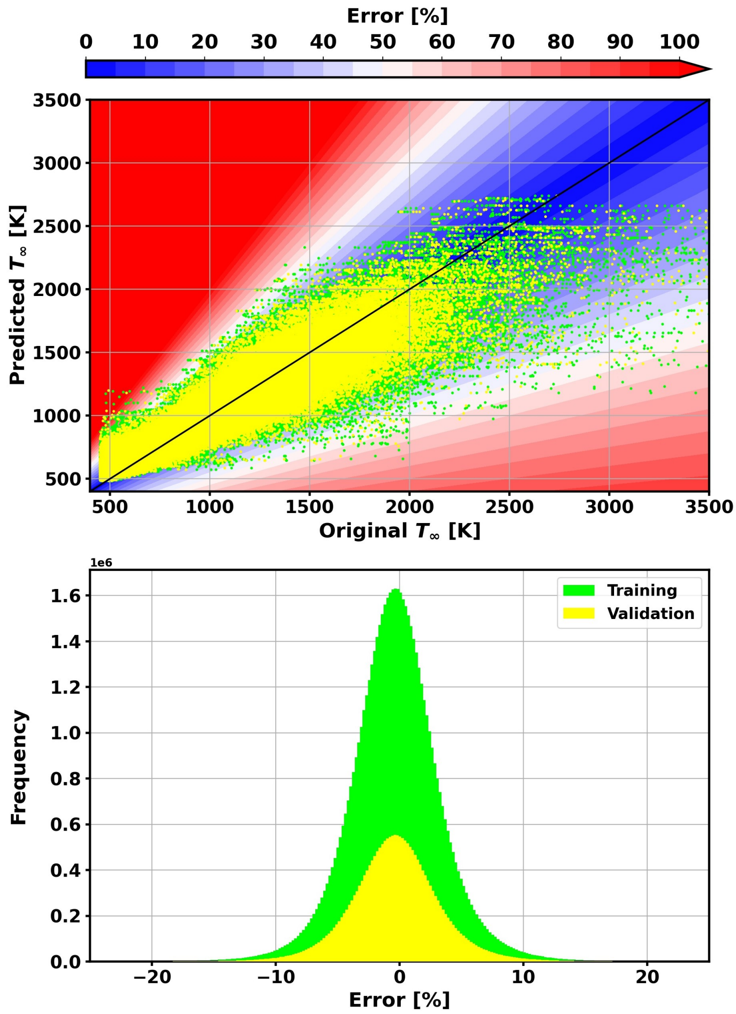

Figure 1. Comparison of observed and predicted T_∞ for training (green) and validation (yellow) sets. The background contour shows mean absolute percent error for that temperature combination. The bottom panel shows histograms of error for both sets.

The T_∞ scatter plot is fairly centered on the 1:1 line, which indicates a zero-error prediction. There is a skew towards underprediction at very high temperatures. However, some of these exospheric temperatures are not physical, due to some instances where an abnormally high temperature needs to be input to MSIS to obtain a match with the measured density. A distinguishing feature in the top panel is the similarity between training and validation performance. This indicates that the model is well-generalized and performs well on independent/new data. The bottom panel of Figure 1 shows that the error distributions have close to zero-mean, and over 98% of training and validation samples have less than 10% error. The mean absolute error for training and validation are both 2.81%, confirming the generalized behavior of the model. Table 3 shows the mean absolute error and standard deviation of the error for EXTEMPLAR-ML and the best linear EXTEMPLAR model from Weimer et al. (2020), in Kelvin.

Table 3. Model statistics for linear EXTEMPLAR (version 6) and EXTEMPLAR-ML.

Model	Mean Absolute Error	Standard Deviation
EXTEMPLAR v6	27.92 K	37.53 K
EXTEMPLAR-ML	23.72 K	34.12 K

EXTEMPLAR-ML achieves an absolute error reduction of over 4 K, and a reduction in the error standard deviation of over 3 K. As previously mentioned, the EXTEMPLAR-ML drivers are only a subset of the EXTEMPLAR version 6 (v6) drivers, in an effort to make the model operation-capable, which makes the performance improvement significant. Next, we evaluate EXTEMPLAR-ML for a global grid to compare Figure 2 (below) with Figure 4 from Weimer et al. (2020).

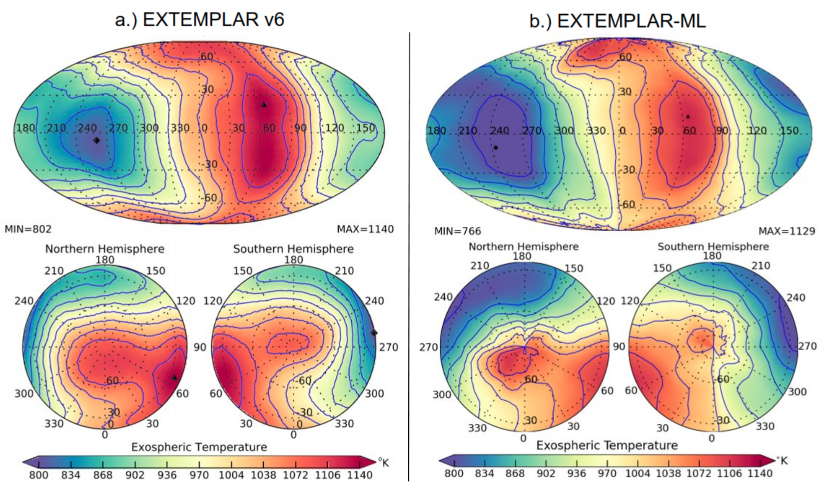

Figure 2. Global T_∞ map with following inputs: S_{10}=M_{10}=120 sfu, S_N=S_S=50 GW, doy = 80, and UT = 15 hours for EXTEMPLAR v6 (left) and EXTEMPLAR-ML (right). The black triangle and square refer to the maximum and minimum temperature locations, respectively.

The global map has no clear defects and shows strong similarities to the EXTEMPLAR v6 map for the same conditions. The locations of maximum and minimum temperatures are also similar to the previous model. The main difference is in the low temperature region in the western hemisphere, where there is a larger region of < 834 K for EXTEMPLAR-ML. As there are only point estimates along orbits, there is no way to validate the global temperature maps, so we cannot confidently say which map is more accurate. However, we attribute this difference to the nonlinear temperature formulation by EXTEMPLAR-ML.

3.1 Principal Component Analysis

With the extensive EXTEMPLAR-ML prediction set described in Section 2.2, we can investigate the most dominant modes of variability and their associated PCA coefficients.

This is shown in Figure 3. It is important to note that the data was not centered prior to performing PCA, so the first mode is representative of the mean temperature distribution over this period. The longitudinal coordinates are local solar time scaled to longitude values with local noon being at 0°.

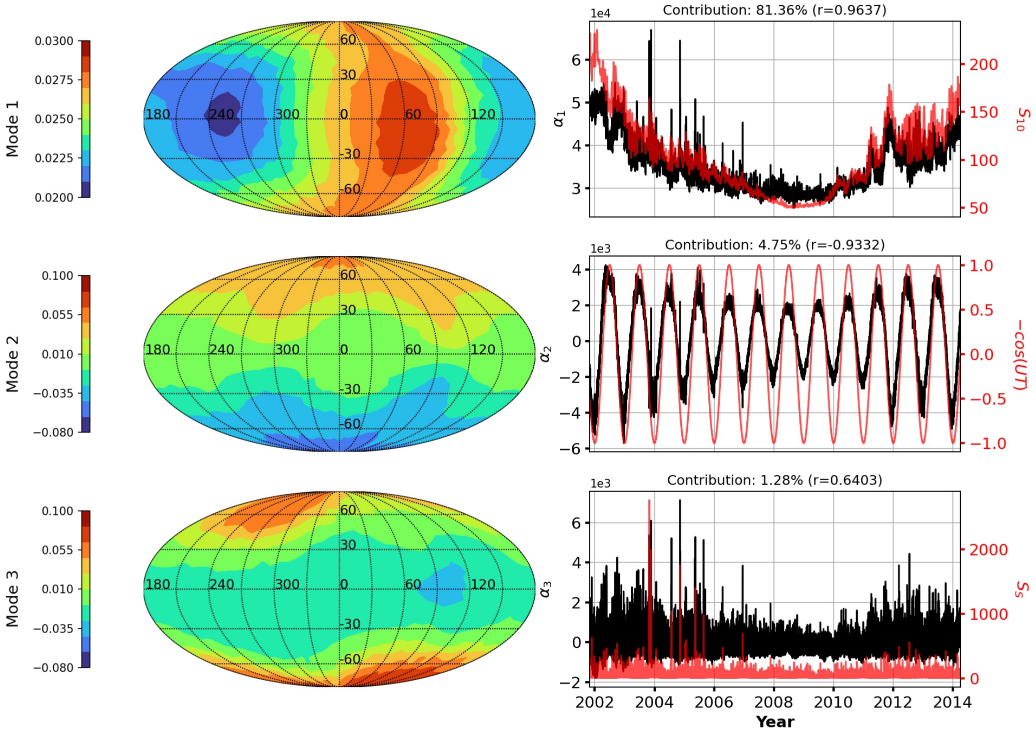

Figure 3. First three modes (left) and corresponding PCA coefficients (right) for the exospheric temperatures between November 1, 2001 and April 1, 2014. The most highly correlated drivers are plotted against the coefficients for comparison with the Pearson correlation coefficient shown in the title (*Schober et al.* 2018).

The first mode is representative of solar EUV heating denoted by the diurnal temperature map and strong correlation with S_{10}. This mode accounts for over 80% of the system's variance. Mode 2 represents a latitudinal Summer-Winter variation. There is a linear progression of the mode with latitude and α_2 oscillated about zero with a period of 365 days. It has an inverse relationship to t_2, described in Equation 1, and its amplitude is a function of the solar activity. Mode 3 resembles a map of the magnetic field with the low latitude band following the magnetic equator. There are also peaks in the poles, and α_3 most strongly correlates with the Poynting flux totals (~0.64 with S_N and S_S). We suspect this mode corresponds to the effects of high latitude heating from either Joule heating or electron precipitation.

3.2 Modeled vs Observed Temperature Along Orbits

In an effort to compare model performance between other temperature models and the observations, we evaluate the MSIS (unmodified), EXTEMPLAR v6, and EXTEMPLAR-ML exospheric temperature values along CHAMP and GRACE-A orbits for two storm periods. The first is the 48 hour period between 12:00 UT on July 26, 2004 to 12:00 UT on July 28, 2004. The results are shown in Figure 4.

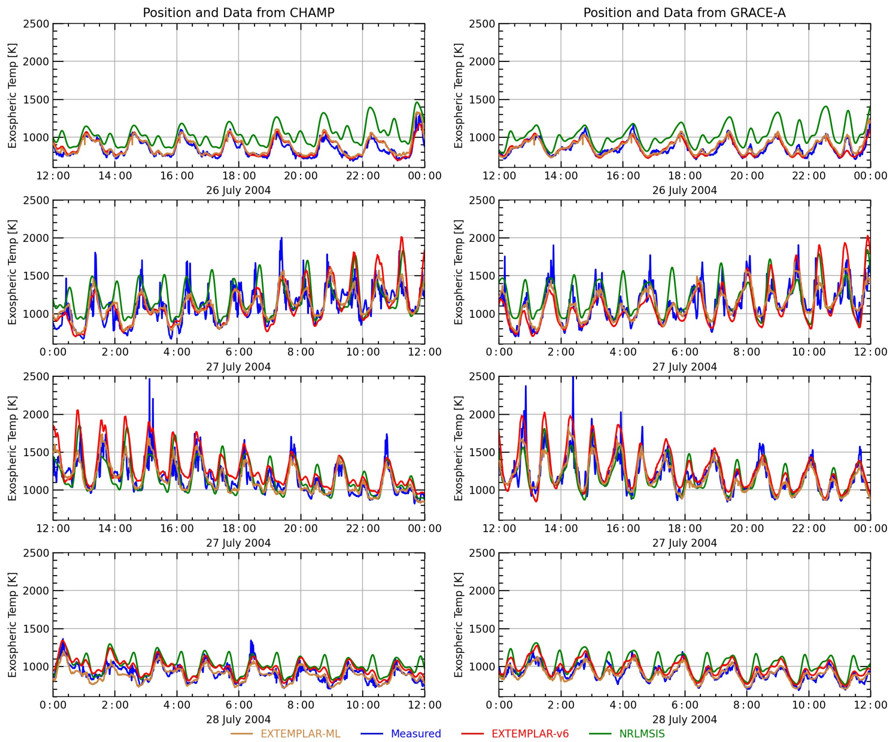

Figure 4. T_∞ for MSIS, EXTEMPLAR v6, EXTEMPLAR-ML plotted alongside CHAMP (left) and GRACE-A (right) for a July 2004 geomagnetic storm.

In the pre-storm period (top panel), both EXTEMPLAR models adequately track the satellites while MSIS overpredicts. During the early phase of the storm, both EXTEMPLAR models still track the observations well, but the linear model becomes more sporadic around 10:00 UT on July 27$^{\text{th}}$. Both models far outperform MSIS in the second panel. The third panel shows the storm recovery where EXTEMPLAR-ML tracks the observations most closely, and EXTEMPLAR v6 overpredicts. In the last twelve hours, both EXTEMPLAR models do well, and MSIS overpredicts. The mean absolute error for EXTEMPLAR-ML is 5.84% and 4.85% with respect to CHAMP and GRACE-A. For EXTEMPLAR v6, the mean absolute error is 8.55% and 7.46% with respect to CHAMP and GRACE-A. MSIS has the highest errors with 15.05% and 12.86% with respect to CHAMP and GRACE-A. Figure 5 shows a similar comparison for a 72-hour period encompassing the 2003 Halloween storm.

EXTEMPLAR-ML tracks the trends in both satellites for all 12-hour windows. Like the other two models, it has a more general response and does not track the abrupt peaks that are likely a result of imperfections in data processing. EXTEMPLAR struggles for brief periods and MSIS has a mixed response with respect to accuracy. The mean absolute error for EXTEMPLAR-ML is 6.26% and 5.64% with respect to CHAMP and GRACE-A. For EXTEMPLAR v6, the mean absolute error is 10.89% and 9.98% with respect to CHAMP and GRACE-A. MSIS has the highest errors with 15.91% and 13.03% with respect to CHAMP and GRACE-A.

3.3 Modeled vs Observed Density Along Orbits

We input the exospheric temperatures from Figures 4 and 5 into MSIS in order to obtain the associated mass density values along the CHAMP and GRACE-A orbits. Figure 6

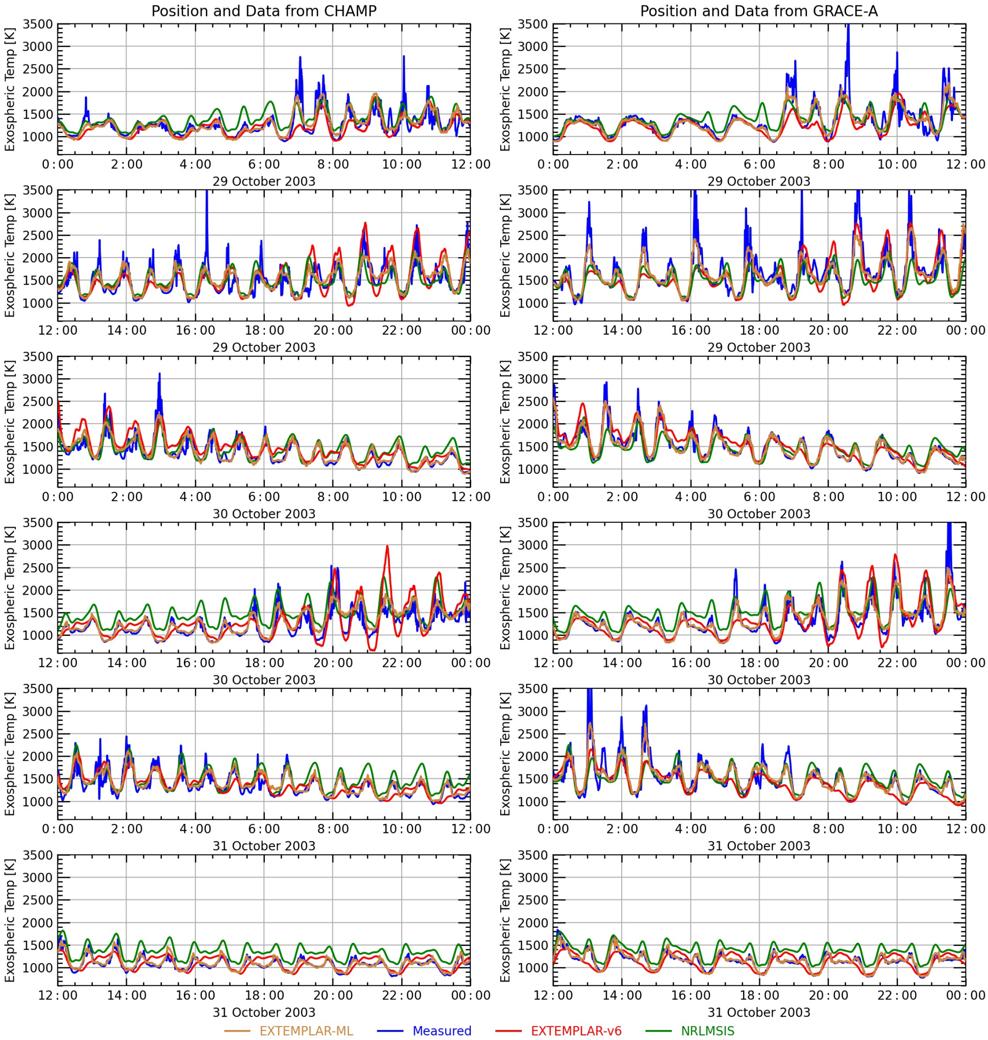

Figure 5. T_∞ for MSIS, EXTEMPLAR v6, EXTEMPLAR-ML plotted alongside CHAMP (left) and GRACE-A (right) for the 2003 Halloween Storm.

shows the modeled densities, including the unmodified MSIS values, along with the satellite density estimates for the July 2004 storm.

The overprediction of exospheric temperature by MSIS in the first 12 hours, seen in Figure 4, causes its modeled density to be notably higher than the observed values. Both EXTEMPLAR models provide similar accuracy pre-storm, but EXTEMPALR-ML more closely matches the CHAMP and GRACE-A estimates during the storm. In the recovery phase, EXTEMPLAR-ML densities are more similar to GRACE-A than to CHAMP. The mean absolute error for EXTEMPLAR-ML is 19.13% and 20.81% with respect to CHAMP and GRACE-A. For EXTEMPLAR v6, the mean absolute error is 27.78% and 32.33% with respect to CHAMP and GRACE-A. MSIS has the highest errors with 62.32% and 75.38% with respect to CHAMP and GRACE-A. Figure 7 shows the density variations resulting from the temperatures in Figure 5.

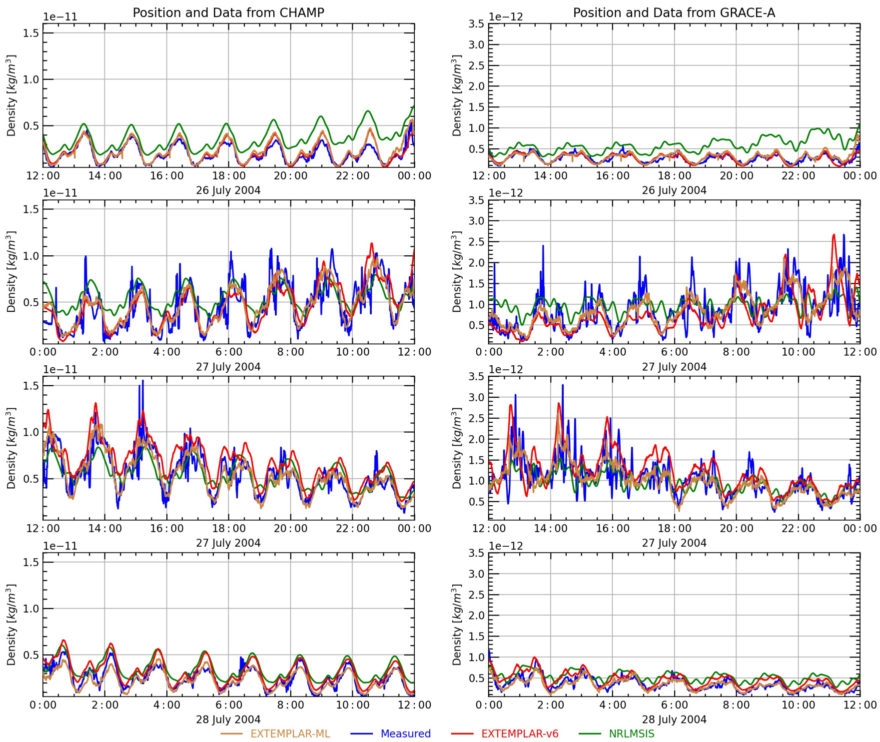

Figure 6. Density for MSIS, EXTEMPLAR v6, EXTEMPLAR-ML plotted alongside CHAMP (left) and GRACE-A (right) for a July 2004 geomagnetic storm.

EXTEMPLAR-ML tracks the satellite densities well for the entire three-day period. The performance enhancement is most notable in the recovery phase (bottom panel) where the other two models tend to overpredict. The ability to capture the anomalously low temperature and density from enhanced NO production following a shock-led geomagnetic storm is highly desired as this prolonged period of lower than expected density can result in substantially different satellite positions in the context of conjunction analyses (*Oliveira and Zesta* 2019). The mean absolute error for EXTEMPLAR-ML is 14.43% and 15.94% with respect to CHAMP and GRACE-A. For EXTEMPLAR v6, the mean absolute error is 24.62% and 25.34% with respect to CHAMP and GRACE-A. MSIS has the highest errors with 35.39% and 35.97% with respect to CHAMP and GRACE-A.

4 Summary

In this work, we developed and optimized a machine learned nonlinear regression model to predict exospheric temperatures given a set of operational Space Weather and temporal drivers. This model, called EXTEMPLAR-ML, has nearly identical training and validation/test performance with 2.81% mean absolute error across 81 million available samples. This is an extension of a linear EXTEMPLAR model developed by Weimer et al. (2020). Using fewer drivers, EXTEMPLAR-ML outperforms EXTEMPLAR with a 4.20 K decrease in absolute error and a 3.41 K decrease in the error standard deviation. An advantage of EXTEMPLAR-ML is that the single model can provide temperature predictions at any location, which was a limitation of its predecessor. The use of PCA provided insight to the temperature formulation within the "black-box" ML model. The first mode represented the effects of solar EUV heating and accounted for 81% of the system's variance. Latitudinal variations accounted for the next 4.75% of the variance and were still a function of solar ac-

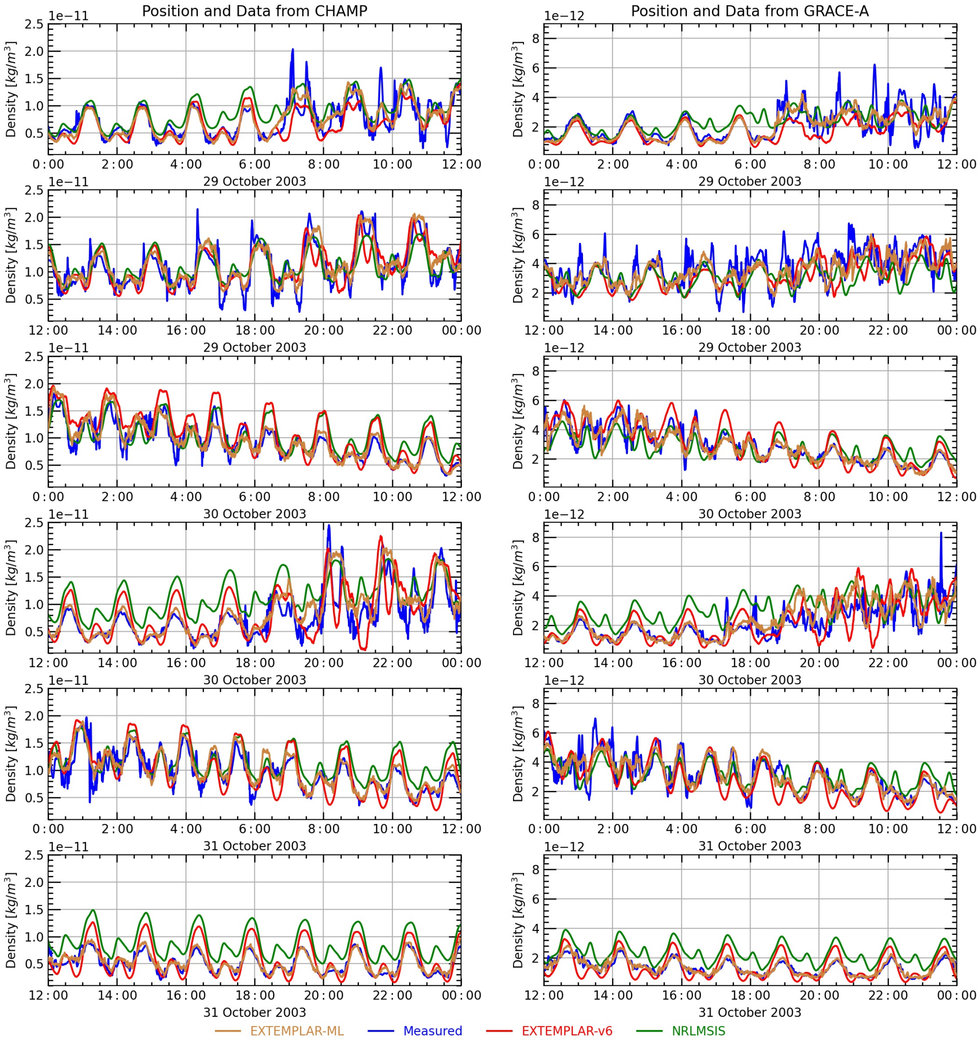

Figure 7. Density for MSIS, EXTEMPLAR v6, EXTEMPLAR-ML plotted alongside CHAMP (left) and GRACE-A (right) for the 2003 Halloween Storm.

tivity. The last mode we looked at only accounted for 1.28% of the variance but described the effects of high latitude heating caused by geomagnetic storms.

We performed two case studies where EXTEMPLAR-ML along with EXTEMPLAR v6 and MSIS predicted T_∞ along CHAMP and GRACE-A orbits during two major geomagnetic storms. In the July 2004 storm, EXTEMPLAR-ML achieved error reductions along both CHAMP and GRACE-A's orbits ranging from 2.07-9.35% compared to EXTEMPLAR v6 and MSIS. In the 2003 Halloween storm, EXTEMPLAR v6 struggled during periods, leading to a higher error reduction of 4.57% and 3.97% with respect to CHAMP and GRACE-A. When these temperatures were used for density prediction, the relative accuracy of EXTEMPLAR-ML became more pronounced. The error reduction from EXTEMPLAR in terms of the resulting density ranged from 8-11% and 9-12% with respect to CHAMP and GRACE-A, respectively for the two storms. In the future, we plan to incorporate model uncertainty into EXTEMPLAR-ML (*Licata et al.* 2020a, *Licata and Mehta* 2021, *Licata et al.* 2021a). We plan to develop a newer model using temperatures derived with NRLMSIS 2.0

(*Emmert et al.* 2021). With the successive model, we also plan to use the exact observation locations to reduce errors associated with binning measurements to the polyhedral grid.

Data Availability Statement

CHAMP and GRACE density estimates from (*Mehta et al.* 2017) can be found at `http://tinyurl.com/densitysets`. A data archive containing the supplemental graphs of neutral density predictions can be accessed online (at `https://doi.org/10.5281/zenodo.3525166`). Also contained here are the adjustments to the NRLMSISE-00 model supplied by J. Emmert; the total Poynting flux into both Northern and Southern Hemispheres from the Weimer 2005 model, for years 2002–2017; the derived ΔT values; and EXTEMPLAR model code with the required files.

Acknowledgements

This work was supported by NASA grant 80NSSC20K1362 to Virginia Tech under the Space Weather Operations 2 Research Program, with subcontracts to WVU and SET. The authors would like to thank Douglas Drob for his insight into the MSIS model. The authors also appreciate the work of the anonymous reviewers for all of their time and effort in helping improve this manuscript.

References

Bates, D. R. (1959), Some Problems concerning the Terrestrial Atmosphere above about the 100 km Level, *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, *253*(1275), 451–462.

Bettadpur, S. (2012), Gravity Recovery and Climate Experiment: Product Specification Document, *GRACE 327-720, CSR-GR-03-02*, cent. for Space Res., The Univ. of Texas, Austin, TX.

Bowman, B., W. K. Tobiska, F. Marcos, C. Huang, C. Lin, and W. Burke (2008), A New Empirical Thermospheric Density Model JB2008 Using New Solar and Geomagnetic Indices, in *AIAA/AAS Astrodynamics Specialist Conference*, AIAA 2008-6438.

Bowman, B. R., and W. K. Tobiska (2006), Improvements in Modeling Thermospheric Densities Using New EUV and FUV Solar Indices, in *AIAA/AAS Astrodynamics Specialist Conference*, AAS 06-237.

Bruinsma, S. (2015), The DTM-2013 thermosphere model, *J. Space Weather Space Clim.*, *5*, A1, doi:10.1051/swsc/2015001.

Calabia, A., and S. Jin (2016), New modes and mechanisms of thermospheric mass density variations from GRACE accelerometers, *Journal of Geophysical Research: Space Physics*, *121*(11), 11,191–11,212, doi:https://doi.org/10.1002/2016JA022594.

Deng, Y., T. J. Fuller-Rowell, A. J. Ridley, D. Knipp, and R. E. Lopez (2013), Theoretical study: Influence of different energy sources on the cusp neutral density enhancement, *Journal of Geophysical Research: Space Physics*, *118*(5), 2340–2349, doi:https://doi.org/10.1002/jgra.50197.

Doornbos, E. (2012), *Producing Density and Crosswind Data from Satellite Dynamics Observations*, pp. 91–126, Springer Berlin Heidelberg, Berlin, Heidelberg, doi:10.1007/978-3-642-25129-0_4.

Emmert, J. T., D. P. Drob, J. M. Picone, D. E. Siskind, M. Jones Jr., M. G. Mlynczak, P. F. Bernath, X. Chu, E. Doornbos, B. Funke, L. P. Goncharenko, M. E. Hervig, M. J. Schwartz, P. E. Sheese, F. Vargas, B. P. Williams, and T. Yuan (2021), NRLMSIS 2.0: A Whole-Atmosphere Empirical Model of Temperature and Neutral Species Densities, *Earth and Space Science*, *8*(3), e2020EA001,321, doi:https://doi.org/10.1029/2020EA001321, e2020EA001321 2020EA001321.

Fedrizzi, M., T. J. Fuller-Rowell, and M. V. Codrescu (2012), Global Joule heating index derived from thermospheric density physics-based modeling and observations, *Space Weather*, *10*(3), doi:https://doi.org/10.1029/2011SW000724.

F.R.S., K. P. (1901), LIII. On lines and planes of closest fit to systems of points in space, *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, *2*(11), 559–572, doi:10.1080/14786440109462720.

Gondelach, D. J., and R. Linares (2020), Real-Time Thermospheric Density Estimation via Two-Line Element Data Assimilation, *Space Weather*, *18*(2), e2019SW002,356, doi:https://doi.org/10.1029/2019SW002356, e2019SW002356 10.1029/2019SW002356.

Hotelling, H. (1933), Analysis of a complex of statistical variables into principal components, *Journal of Educational Psychology*, *24*(6), 417–441, doi:10.1037/h0071325.

Jacchia, L. G. (1965), Static Diffusion Models of the Upper Atmosphere with Empirical Temperature Profiles, *Smithsonian Contributions to Astrophysics*, *8*, 215.

Jacchia, L. G. (1970), New Static Models of the Thermosphere and Exosphere with Empirical Temperature Profiles, *SAO Special Report*, *313*.

Knipp, D. J., D. V. Pette, L. M. Kilcommons, T. L. Isaacs, A. A. Cruz, M. G. Mlynczak, L. A. Hunt, and C. Y. Lin (2017), Thermospheric nitric oxide response to shock-led storms, *Space Weather*, *15*(2), 325–342, doi:https://doi.org/10.1002/2016SW001567.

Kockarts, G. (1980), Nitric oxide cooling in the terrestrial thermosphere, *Geophysical Research Letters*, *7*(2), 137–140, doi:https://doi.org/10.1029/GL007i002p00137.

Lei, J., A. G. Burns, J. P. Thayer, W. Wang, M. G. Mlynczak, L. A. Hunt, X. Dou, and E. Sutton (2012a), Overcooling in the upper thermosphere during the recovery phase of the 2003 October storms, *Journal of Geophysical Research: Space Physics*, *117*(A3), doi:https://doi.org/10.1029/2011JA016994.

Lei, J., T. Matsuo, X. Dou, E. Sutton, and X. Luan (2012b), Annual and semiannual variations of thermospheric density: EOF analysis of CHAMP and GRACE data, *Journal of Geophysical Research: Space Physics*, *117*(A1), doi:https://doi.org/10.1029/2011JA017324.

Licata, R., and P. Mehta (2021), Physics-informed Bayesian Deep Learning for Space Weather Science and Operations, in *Proceedings of the 18th Conference on Space Weather - 101st AMS Annual Meeting*, doi:10.13140/RG.2.2.13248.76807.

Licata, R., P. Mehta, and W. K. Tobiska (2020a), Data-Driven HASDM Density Model using Machine Learning, in *Proceedings of the AGU Fall Meeting 2020*, doi:10.1002/essoar.10505213.1.

Licata, R., P. Mehta, and W. K. Tobiska (2021a), Impact of Driver and Model Uncertainty on Drag and Orbit Prediction, in *Proceedings of the 31st AAS/AIAA Space Flight Mechanics Meeting*.

Licata, R. J., W. K. Tobiska, and P. M. Mehta (2020b), Benchmarking Forecasting Models for Space Weather Drivers, *Space Weather*, *18*(10), e2020SW002,496, doi:https://doi.org/10.1029/2020SW002496, e2020SW002496 10.1029/2020SW002496.

Licata, R. J., P. M. Mehta, W. K. Tobiska, B. R. Bowman, and M. D. Pilinski (2021b), Qualitative and Quantitative Assessment of the SET HASDM Database, *Earth and Space Science Open Archive*, p. 21, doi:10.1002/essoar.10506516.2.

Luhr, H., L. Grunwaldt, and C. Forste (2002), CHAMP Reference Systems, Transformations and Standards, *Tech. rep.*, CH-GFZ-RS-002, GFZ-Potsdam, Postdam, Germany.

Matsuo, T., and J. M. Forbes (2010), Principal modes of thermospheric density variability: Empirical orthogonal function analysis of CHAMP 2001–2008 data, *Journal of Geophysical Research: Space Physics*, *115*(A7), doi:https://doi.org/10.1029/2009JA015109.

Mehta, P. M., and R. Linares (2017), A methodology for reduced order modeling and calibration of the upper atmosphere, *Space Weather*, *15*(10), 1270–1287, doi:10.1002/2017SW001642.

Mehta, P. M., A. C. Walker, E. K. Sutton, and H. C. Godinez (2017), New density estimates derived using accelerometers on board the CHAMP and GRACE satellites, *Space Weather*, *15*(4), 558–576, doi:https://doi.org/10.1002/2016SW001562.

Mehta, P. M., R. Linares, and E. K. Sutton (2018), A Quasi-Physical Dynamic Reduced Order Model for Thermospheric Mass Density via Hermitian Space-Dynamic Mode Decomposition, *Space Weather*, *16*(5), 569–588, doi:10.1029/2018SW001840.

Mehta, P. M., R. Linares, and E. K. Sutton (2019), Data-Driven Inference of Thermosphere Composition During Solar Minimum Conditions, *Space Weather*, *17*(9), 1364–1379, doi: https://doi.org/10.1029/2019SW002264.

Mlynczak, M., F. J. Martin-Torres, J. Russell, K. Beaumont, S. Jacobson, J. Kozyra, M. Lopez-Puertas, B. Funke, C. Mertens, L. Gordley, R. Picard, J. Winick, P. Wintersteiner, and L. Paxton (2003), The natural thermostat of nitric oxide emission at 5.3 μm in the thermosphere observed during the solar storms of April 2002, *Geophysical Research Letters*, *30*(21), doi:https://doi.org/10.1029/2003GL017693.

Oliveira, D. M., and E. Zesta (2019), Satellite Orbital Drag During Magnetic Storms, *Space Weather*, *17*(11), 1510–1533, doi:https://doi.org/10.1029/2019SW002287.

O'Malley, T., E. Bursztein, J. Long, F. Chollet, H. Jin, L. Invernizzi, et al. (2019), Keras Tuner, https://github.com/keras-team/keras-tuner.

Picone, J. M., A. E. Hedin, D. P. Drob, and A. C. Aikin (2002), NRLMSISE-00 empirical model of the atmosphere: Statistical comparisons and scientific issues, *Journal of Geophysical Research: Space Physics*, *107*(A12), SIA 15–1–SIA 15–16, doi:10.1029/2002JA009430.

Qian, L., and S. Solomon (2011), Thermospheric Density: An Overview of Temporal and Spatial Variations, *Space Science Reviews - SPACE SCI REV*, *168*, 1–27, doi:10.1007/s11214-011-9810-z.

Ruan, H., J. Lei, X. Dou, S. Liu, and E. Aa (2018), An Exospheric Temperature Model Based On CHAMP Observations and TIEGCM Simulations, *Space Weather*, *16*(2), 147–156, doi:https://doi.org/10.1002/2017SW001759.

Schober, P., C. Boer, and L. Schwarte (2018), Correlation Coefficients: Appropriate Use and Interpretation, *Anesthesia & Analgesia*, *126*, 1763–1768, doi:10.1213/ANE.0000000000002864.

Sutton, E. K. (2008), Effects of solar disturbances on the thermosphere densities and winds from CHAMP and GRACE satellite accelerometer data, Ph.D. thesis, University of Colorado at Boulder.

Sutton, E. K., S. B. Cable, C. S. Lin, L. Qian, and D. R. Weimer (2012), Thermospheric basis functions for improved dynamic calibration of semi-empirical models, *Space Weather*, *10*(10), doi:https://doi.org/10.1029/2012SW000827.

Tobiska, W., T. Woods, F. Eparvier, R. Viereck, L. Floyd, D. Bouwer, G. Rottman, and O. White (2000), The SOLAR2000 empirical solar irradiance model and forecast tool, *Journal of Atmospheric and Solar-Terrestrial Physics*, *62*(14), 1233 – 1250, doi:https://doi.org/10.1016/S1364-6826(00)00070-5.

Tobiska, W. K., S. D. Bouwer, and B. R. Bowman (2008), The development of new solar indices for use in thermospheric density modeling, *Journal of Atmospheric and Solar-Terrestrial Physics*, *70*(5), 803–819, doi:https://doi.org/10.1016/j.jastp.2007.11.001.

Walker, J. C. G. (1965), Analytic Representation of Upper Atmosphere Densities Based on Jacchia's Static Diffusion Models, *Journal of Atmospheric Sciences*, *22*(4), 462 – 463, doi:10.1175/1520-0469(1965)022<0462:AROUAD>2.0.CO;2.

Weimer, D. R. (2005a), Improved ionospheric electrodynamic models and application to calculating Joule heating rates, *Journal of Geophysical Research: Space Physics*, *110*(A5), doi:https://doi.org/10.1029/2004JA010884.

Weimer, D. R. (2005b), Predicting surface geomagnetic variations using ionospheric electrodynamic models, *Journal of Geophysical Research: Space Physics*, *110*(A12), doi:https://doi.org/10.1029/2005JA011270.

Weimer, D. R., E. K. Sutton, M. G. Mlynczak, and L. A. Hunt (2016), Intercalibration of neutral density measurements for mapping the thermosphere, *Journal of Geophysical Research: Space Physics*, *121*(6), 5975–5990, doi:https://doi.org/10.1002/2016JA022691.

Weimer, D. R., P. M. Mehta, W. K. Tobiska, E. Doornbos, M. G. Mlynczak, D. P. Drob, and J. T. Emmert (2020), Improving Neutral Density Predictions Using Exospheric Temperatures Calculated on a Geodesic, Polyhedral Grid, *Space Weather*, *18*(1), e2019SW002,355, doi:https://doi.org/10.1029/2019SW002355.

Weng, L., J. Lei, E. Sutton, X. Dou, and H. Fang (2017), An exospheric temperature model from CHAMP thermospheric density, *Space Weather*, *15*(2), 343–351, doi:https://doi.org/10.1002/2016SW001577.