3rd URSI AT-AP-RASC, Gran Canaria, 29 May – 3 June 2022

Machine Learning Ensemble Approach for Ionosphere and Space Weather Forecasting with Uncertainty Quantification

Randa Natras*(1), Benedikt Soja(2), and Michael Schmidt(1)
(1) Deutsches Geodätisches Forschungsinstitut der Technisches Universität München (DGFI-TUM), Department of Aerospace and Geodesy, Technical University of Munich, Munich, 80333, Germany
(2) Institute of Geodesy and Photogrammetry, ETH Zurich, Zurich, 8093, Switzerland

Abstract

This paper presents a novel Machine Learning (ML) approach to ionospheric forecasting, including forecasting the space weather impact on the ionosphere. It exploits a data-driven approach in which the models learn underlying processes and relationships from data describing solar activity, solar wind, interplanetary and Earth's magnetic fields, and the ionosphere. We applied a multi-model and multi-data ensemble forecasting approach using diverse models of different learning algorithms with different training datasets to generate 1-day VTEC forecasts. This approach improved forecasting accuracy compared to a single-model-based approach. In addition, the forecast uncertainty of the super-ensemble model was assessed by estimating an ensemble spread. The results show potential for forecasting VTEC in different ionospheric regions during quiet and storm periods while quantifying their uncertainties.

1 Introduction

In order to model the Vertical Total Electron Content (VTEC) and space weather effects in the ionosphere, a complex chain of physical, dynamic processes between the Sun, the interplanetary space, the interplanetary (IMF) and the Earth's (EMF) magnetic fields, and the ionosphere must be taken into account. However, we have a limited understanding of the underlying space weather processes to accurately model them using traditional methods. On the other hand, there is an urgent need to develop an advanced forecasting system to mitigate possible catastrophic failures of space- and ground-based technological systems, including GNSS (Global Navigation Positioning System), caused by severe space weather events. Since numerous data from satellites and observatories that monitor space weather are available nowadays, they can be exploited for improving space weather forecasting. Machine Learning (ML), a subset of artificial intelligence, is one of the most rapidly growing areas today and most promising recent technologies that provides new capabilities to learn directly from data. This approach can help us to discover the hidden relationships within the data, deepen our physical understanding and, approximate nonlinear functions in order to describe the underlying space weather processes based on the provided data [1]. Previous work on VTEC forecasting mainly included deep learning methods such as Feed-forward Neural Networks [2] and Long Short-Term Memory (LSTM) [3]. Deep learning proved to be more accurate for ionosphere modeling than traditional linear approaches of Empirical Orthogonal Functions (EOF) [2] and AutoRegressive Integrated Moving Average (ARIMA) [3]. Only a few studies so far have used other learning algorithms such as XGBoost [4]. However, the uncertainty of the forecasts has not yet been provided. The novelties of this study are primarily the application of ensemble modeling and computation of the forecast uncertainties.

2 Methodology

In this study, the ML model for VTEC forecasting is based on supervised learning, i.e. it learns from the past experience (represented as the training data) with respect to the task of the VTEC forecast and the performance measure as root mean square error (RMSE) and correlation coefficient (Corr.). Supervised learning can be seen as a function estimation or predictive learning problem [5]. Using a training sample of the input $\mathbf{x}_i$ (predictors, features, or the independent variables) and the output y_i (response or the dependent variable), for each of the N observations ($i = 1,2,...,N$), the goal is to find an approximation $\hat{F}(\mathbf{x}_i)$ of the function $F(\mathbf{x}_i)$ that maps inputs to the output (Figure 1). The ML model parameters and hyperparameters are optimized using cross-validation during the learning phase. When the iterative process of feature selection, model tuning, and training is completed, the model can be tested on previously unseen data to forecast VTEC. Publication [6] discusses the ML workflow for space weather forecasting in more detail, from problem formulation, feature engineering, learning algorithms to model training, evaluation, and deployment with challenges and open issues.

Solar activity, solar wind, IMF, and EMF data were downloaded from NASA/GSFC OMNI-Web (https://omniweb.gsfc.nasa.gov/ow.html). The VTEC values for high-latitude (10E 70N), mid-latitude (10E 40N), and low-latitude (10E 10N) regions were extracted from the Global Ionosphere Map (GIM) of CODE (IGS AC at the University of Bern)

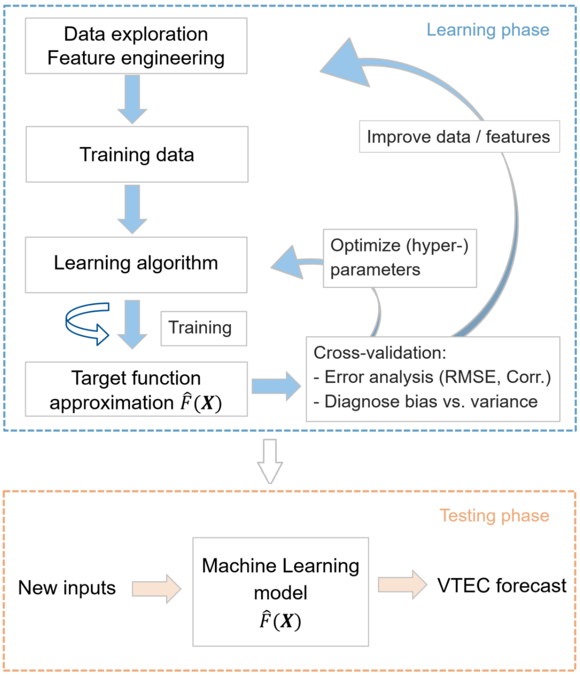

Figure 1. Workflow of the ML model development for VTEC forecasting consisting of a learning/training phase and a testing phase.

(https://cddis.nasa.gov/archive/gnss/products/ionex). In order to model the VTEC temporal and seasonal dependencies, hour of day and day of year (DOY) were added as features to the model. Exponential moving average (EMA) and time derivatives of VTEC are additional features. The exponential moving average gives higher weights on the latest data values, and is, therefore, more sensitive to recent values and recent value changes than a simple moving average. The training data cover the period from January 2015 to December 2016, while the testing is performed for the subsequent year 2017. Three data sets are prepared: 1) Original data (D1); 2) Daily differences for both inputs and outputs (D2); 3) Input data combining the first and the second data sets, while the output corresponds to the original VTEC (D3).

Learning algorithms from ensemble learning were selected for the study. Ensemble learning combines several simple models (usually based on a Decision Tree) to produce a robust model that can generalize better than a single model. We used popular ensemble learning methods of bagging (Random Forest) and boosting (AdaBoost and XGBoost). Random Forest [7] builds a large collection of de-correlated trees and then averages them to form a final outcome. In the boosting method, the trees are grown sequentially using the information from previously grown trees with a modified version of the training data. In AdaBoost (Adaptive Boosting) [8], the data are modified by applying weights to each of the training examples $(\mathbf{x}_i, y_i)$. In each successive iteration, the weights increase for wrongly estimated observations in the previous iteration, while the weights decrease for the correct ones. The final outcome represents the weighted estimations from all iterations. Gradient boosting offers a generalization of boosting to arbitrary differentiable objective functions, where a tree is fitted to the gradient. XGBoost (eXtreme Gradient Boosting) [9] is an optimized gradient boosting algorithm that uses regularization to avoid overfitting. A detailed explanation of the methods and data sets can be found in [10]. These three learning algorithms were individually trained on the three data sets, i.e. nine models were developed for each ionospheric region, resulting in a total of 27 models. The developed models were then combined into a super-ensemble model, which provides average forecasts of all models within the ensemble (Figure 2). Ensemble modeling is a standard approach for terrestrial weather forecasting. The standard deviation of ensemble members with respect to the ensemble mean, known as ensemble spread, provides an estimate of the forecasting uncertainties. A large (small) spread indicates a low (high) confidence in the forecast.

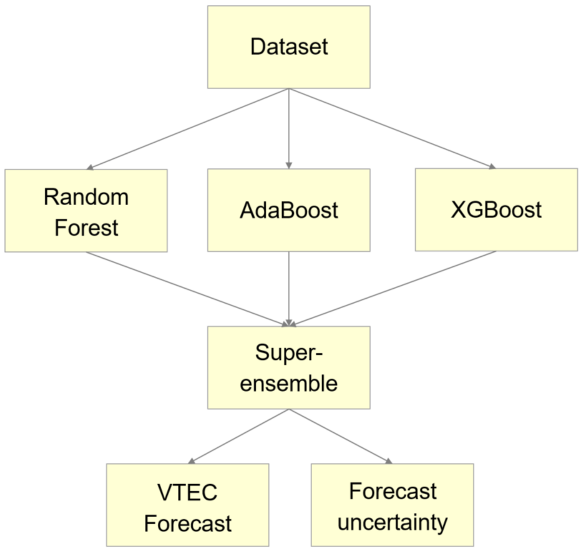

Figure 2. Overview of the developed ML models, trained on different data sets using tree-based learning algorithms. Afterward, they are combined into a super-ensemble model that provides the VTEC forecast and its uncertainty.

3 Results

An overview of solar activity, solar wind speed, and conditions in IMF and EMF for September 6 - 13, 2017 is presented in Figure 3. Two minimum Dst values can be observed: at the midnight and in the evening of September 8 (UTC), representing the main phases of geomagnetic storms. Most of the storms happened at September 8. The recovery phase followed from September 9 to 11, to less stormy conditions on September 12 to 13. On September 6, the day before the initial phase of the first storm, solar activity was increased, while EMF conditions were mostly quiet. Data from September 6 were used to forecast VTEC during the initial storm phase on September 7 in Figure 4.

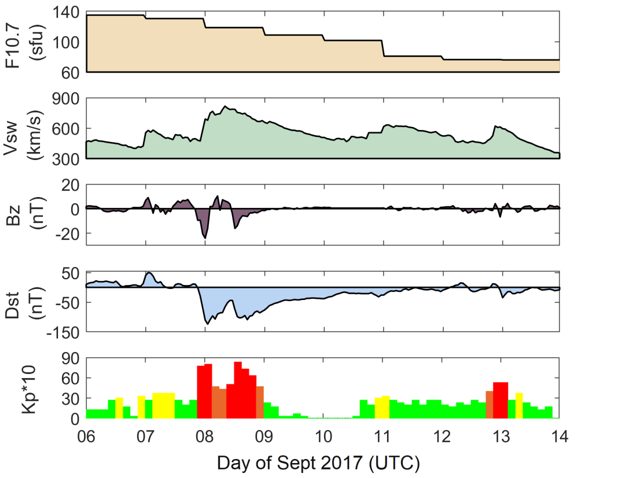

Figure 3. Solar radio flux F10.7, solar wind speed Vsw, IMF Bz, and EMF Dst and Kp data for September 6 - 13, 2017. Quiet (green): Kp < 3, moderate (yellow): $3 \leq$ Kp < 4, active (orange): $4 \leq$ Kp < 5, storm (red): Kp ≥ 5.

The VTEC forecasts of different ML models with the ensemble mean μ and a probabilistic forecast in terms of ± 2 times the standard deviations σ can be seen in Figure 4. The period September 7 to 13, 2017, covers the initial, main, and recovery phase of the severe geomagnetic storms. The forecast uncertainty is largest during the main storms phases (September 8). Afterwards, the ensemble spread gradually narrows from September 9 onwards. The GIM

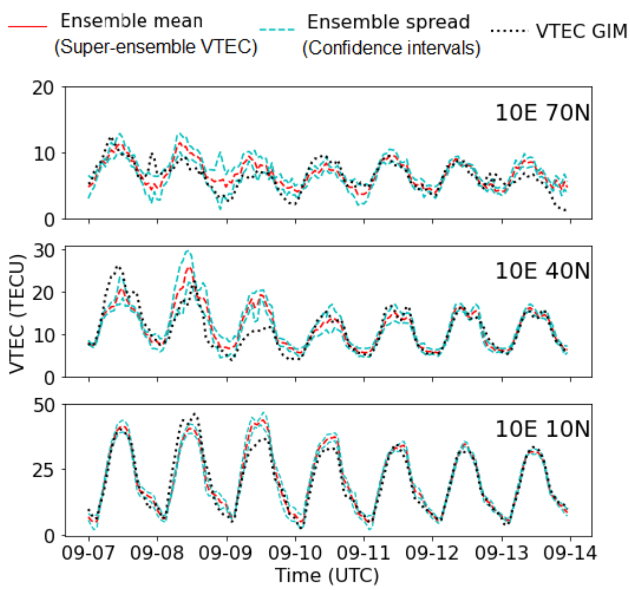

Figure 4. 1-day VTEC forecast of the Super-ensemble model (ensemble mean μ) and ensemble spread ($\mu \pm 2\sigma$) for September 7-13, 2017.

VTEC at high-latitude is mainly within or at the edge of the ensemble spread. The mid-latitude GIM VTEC shows greater differences to the ensemble mean on September 7 (initial storm phase), when the VTEC forecast is underestimated, and on September 9 (beginning of the recovery phase), when the VTEC forecast is overestimated. In the main storm phase, GIM VTEC is mostly within the ensemble spread. The ensemble mean is slightly overestimated during the main phase. During the recovery phase (September 10 - 13), the mid-latitude GIM VTEC is within the ensemble spread and corresponds to the ensemble mean. For the low-latitude region, the differences between the ensemble mean and GIM VTEC are larger during daytime on September 8 and 9. An overview of the RMSE of ML models within the ensemble, the ensemble mean and the mean variance (2σ) is provided in Table 1. The super-ensemble model provides the optimal results.

Table 1. The RMSE of the ML models Random Forest (RF), AdaBoost (AB), and XGBoost (XGB) for three datasets, as well as, super-ensemble (SE) and mean variance (2σ) for high-latitude, mid-latitude and low-latitude VTEC 1-day forecasts from September 7 to 13, 2017.

		RF	**AB**	**XGB**	**SE**	2σ
	D1	1.62	1.67	1.78		
10E 70N	D2	1.75	1.69	1.66	1.63	0.90
	D3	1.66	1.76	1.77		
	D1	3.08	3.05	3.00		
10E 40N	D2	2.82	2.61	2.61	2.80	1.28
	D3	2.96	3.22	2.80		
	D1	3.14	3.38	3.20		
10E 10N	D2	3.31	3.27	3.22	3.13	1.60
	D3	3.20	3.30	3.26		

The most influential input variables (features) for the VTEC forecast are estimated using Shapley values (Figure 5). SHAP (SHapley Additive exPlanations) [11] is a widely

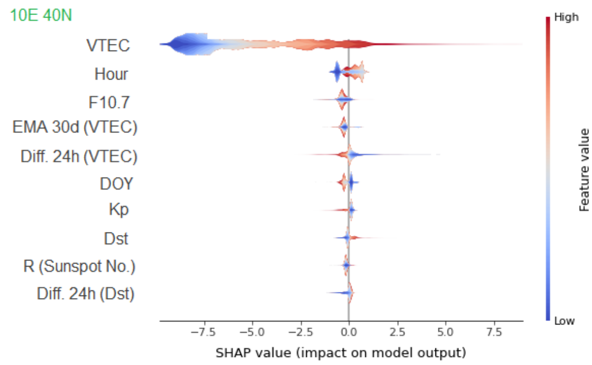

Figure 5. The ten most important features for mid-latitude VTEC 1-day forecasts from the XGBoost model estimated using SHAP for the third test data set (year 2017). "EMA" stands for the exponential moving average over the previous 30 days ("30d") and "Diff. 24h" for daily differences.

used approach from cooperative game theory to interpret the output of ML models. Features are sorted by importance, starting with the most important ones at the top. A standard violin plot illustrates the distribution of the importance for each feature. Red stands for large values of a feature, while blue stands for small ones. The x-axis corresponds to Shapley values that characterize the average impact on the model output. Positive Shapley values mean that a feature increases the VTEC forecast, while negative values represent a feature impact that decreases the VTEC forecast. The larger the absolute Shapley values, the higher impact of a feature on the VTEC forecast. The most important variables for the mid-latitude VTEC forecast are the current VTEC, hour of day, F10.7 solar activity index, exponentially moving VTEC average and VTEC daily difference. Other important variables are DOY, geomagnetic indices Kp and Dst, sunspot number and daily Dst difference.

4 Conclusion

Today's society relies on space- and ground-based technological infrastructures, which are not adequately protected against space weather. Therefore, accurate forecasting and early-warning systems of space weather events are urgently needed. This paper presents a novel model for forecasting VTEC and space weather effects in the ionosphere based on ML and ensemble modeling. To get the most out of the ML model, it is necessary to prepare relevant data and useful features that enhance learning, particularly of space weather events. Also, it is critical to quantify the uncertainty of weather forecasts, especially when forecasting extreme weather events. In terrestrial weather forecasting, this is typically achieved with ensemble forecasting systems. Following that approach, ML models were developed using different learning algorithms and different data sets and then combined via ensemble modeling to produce a super-ensemble model with more reliable results that can quantify their uncertainties. During the severe geomagnetic storm, the forecast of the super-ensemble ML model is less confident than after the storm. Most of the time, the mean and the spread of the ensemble corresponds to the reference VTEC values. The largest differences occur during the initial storm phase and at the beginning of the recovery phase. The changes in the interplanetary and geomagnetic fields were short-term and insufficient information on these changes was provided in the input data from a day before. However, during the main storm phase, the reference VTEC does not deviate much from the ensemble spread. After the storm, the ensemble spread is narrower and the ensemble mean agrees with the reference VTEC.

5 Acknowledgments

We acknowledge the use of NASA/GSFC OMNIWeb service for OMNI data, and the University of Bern for GIM data. This study was funded by Research Grants - Doctoral Programmes in Germany from the German Academic Exchange Service (DAAD).

References

[1] E. Camporeale, S. Wing, and J. Johnson, "Machine learning techniques for space weather," *Elsevier*, Amsterdam, Netherlands, 22, May 2018, ISBN: 9780128117880.

[2] J. C. Uwamahoro, J. B. Habarulema "Modelling total electron content during geomagnetic storm conditions using empirical orthogonal functions and neural networks," *Journal of Geophysical Research: Space Physics*, **120**, 3, December 2015, pp. 794–795, doi:10.1002/2015JA021961.

[3] R. Tang, F. Zeng, Z. Chen "The comparison of predicting storm-time ionospheric tec by three methods: Arima, lstm, and seq2seq," *Atmosphere*, **11**, 25, March 2020, doi:10.3390/atmos11040316.

[4] A. V. Zhukov, Y. V. Yasyukevich, A. E. Bykov "Correction to: Gimli: Global ionospheric total electron content model based on machine learning," *GPS Solutions*, **25**, 21, 22, November 2020, pp. 794–795, doi:10.1007/s10291-020-01063-1.

[5] J. H. Friedman "Greedy function approximation: A gradient boosting machine," *The Annals of Statistics*, **29**, 5, 1, October 2001, pp. 1189–1232, doi:10.1214/aos/1013203451.

[6] R. Natras, M. Schmidt "Machine Learning Model Development for Space Weather Forecasting in the Ionosphere," *Proceedings of the CIKM 2021 Workshops: 1st Workshop on Complex Data Challenges in Earth Observation (CDCEO'21)*, *CEUR-WS.org*, **3052**, 12, December 2021, http://ceur-ws.org/Vol-3052/short10.pdf

[7] L. Breiman "Random forests," *Machine Learning*, **45**, 1, 1, October 2001, pp. 5–32, doi:10.1023/A:1010933404324.

[8] Y. Freund, R. E. Schapire "A decision-theoretic generalization of on-line learning and an application to boosting," *Journal of Computer and System Sciences*, **55**, 1, 1, August 1997, pp. 119–139, doi:10.1006/jcss.1997.1504.

[9] T. Chen, C. Guestrin "Xgboost: A scalable tree boosting system," *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, 1, August 2016, pp. 785–794, doi:10.1145/2939672.2939785.

[10] R. Natras, B. Soja, M. Schmidt "Ensemble Machine Learning of Random Forest, AdaBoost and XGBoost for Forecasting Vertical Total Electron Content of the Ionosphere," *J Geod*, submitted 13, December 2021 (under review).

[11] S. M. Lundberg, S. Lee "A Unified Approach to Interpreting Model Predictions," *Advances in Neural Information Processing Systems*, **30**, 2017, pp. 4765-4774, http://papers.nips.cc/paper/7062-a-unified-approach-to-interpreting-model-predictions.pdf.