EISSN: 2707-0425

East African Journal of Science, Technology and *Innovation, Vol. 6 (Special issue 1): December 2024*

Prediction of f2-layer height of the peak electron density (hmf2) over the Southern Africa region using artificial neural network

[1]MUSHI P E., [1]SULUNGU E., [1]PAMAIN A*

[1]Department of Physics, The University of Dodoma, Dodoma, Tanzania

***Corresponding Author**: eliaspeter089@gmail.com

Abstract

The ionospheric F2 layer is the essential layer in the propagation of high-frequency radio waves, and height of the peak electron density (hmF2) is one of the important parameter. However, ionosondes are not installed at every location on earth to allow for global measurements of hmF2, especially within the southern African region. This study therefore focuses on developing a regional model for predicting the hmF2 using Artificial Neural Network techniques. In this study, prediction model was developed using year, day of the year, time in 30 minutes intervals, Sunspot Number (SSN), Solar flux at 10.7 cm (F10.7), geomagnetic activity (Kp) and Averaged planetary Index (Ap), longitude, latitude and critical frequency (foF2) as the input. The hmF2 prediction of 2009 and 2014 obtained from proposed model its results during summer and winter was compared with AMTB-2013 and SHU-2015, which are recently parameters of the IRI-2020 model. The result showed that in both 2009 and 2014 the year of low solar activity and high solar activity respectively, ANN over perform other models with minimum RMSE and PRMSE values followed by SHU-2015 and finally AMTB-2013. Therefore, it can be concluded that the architecture and learning efficiency of ANN proposed model are as good as SHU-2015 with slightly difference between them. Although all models need some improvement to increase their accuracy.

Keywords: *Artificial neural network; F2-layer height of the peak electron density (hmF2); Ionosonde data; IRI 2020 model; Prediction*

Received: 18/06/24
Accepted: 05/12/24
Published: 20/12/24

Cite as, *Mushi et al., (2024)*. Prediction of f2-layer height of the peak electron density (hmf2) over the Southern Africa region using artificial neural network. *East African Journal of Science, Technology* and *Innovation 6 (Special issue 1).*

Introduction

The ionosphere is an extremely non-linear system and an important part near- the earth space environment that changes with time and space and it is ionized by solar radiation (Cander *et al.*, 1998). The ionosphere contains several important parameters, including F2 layer height of the peak electron density (hmF2), which varies considerably due to multiple factors, including geophysical factors, solar factors and seasonal changes. The ionization rate and, in turn, the electron density in the F2 layer are influenced by the solar activity level, which is determined by sunspots numbers and solar flares. At different latitudes, solar heating and ionization vary due to the Earth's axial tilt and its orbit around the Sun. As a result, the F2 layer and other parts of the ionospheric structure experience seasonal variations (Bothmer and Daglis, 2007). The hmF2 is a critical parameter in ionospheric studies that plays a significant role in communication and navigation systems. Accurate prediction of hmF2 is essential for

various applications, including radio wave propagation modeling, satellite communication planning, and space weather monitoring.

In order to analyze and predict the behavior of the hmF2 layer in the ionosphere, ranges of physical and empirical models are employed. Physical models of the ionosphere rely on simplifying assumptions and mathematical equations to simulate the underlying system. These assumptions may not capture all the complexities and nuances of real-world phenomena. Some physical based models include the Sheffield University Plasmasphere Ionosphere Model (SUPIM) (Bailey *et al.,* 1997), the Global Assimilation of Ionospheric Measurements - Full Physics Model (GAIM-FP) (Schunk *et al.,* 2004), Naval Research Laboratory's (NRL) SAMI3 (Huba *et al.,* 2000) and Ionospheric code SAMI3 and the atmosphere/thermosphere code WACCM-X (Huba and Liu, 2020).

Empirical models are based on large historical data sets that are often sampled irregularly over time and space (Field, 2018). This models use climatological ionospheric data from numerous sources to fit the ionospheric parameters to a particular set of control parameters. They don't necessitate a thorough comprehension of the physical processes that take place in the ionosphere. Global empirical models for the three-dimensional ionosphere that are frequently used include the International Reference Ionosphere (IRI) (Bilitza *et al.,* 2001, Bilitza *et al.,* 2008, Bilitza *et al.,* 2017 and Bilitza *et al.,* 2021) and NeQuick (Leitinge *et al.,* 2000, Radicella *et al.,* 2009). The accurate prediction of the ionospheric conditions is still a challenging task because of its complex variability due to several factors such as solar activity, geomagnetic activity, seasonal variations and diurnal variations (Anderson *et al.,* 1998). Traditional empirical models and analytical approaches have inherent limitations in adequately capturing the intricate relationships between these influential variables and the resulting hmF2 values. Given the shortcomings of existing methods, there is a pressing need to develop alternative models that can better account for the complex interplay of these factors.

In order to simulate complex, non-linear relationships between input and output variables, a promising technique has been discovered to be the Artificial Neural Network (ANN). ANN can manage the complicated and nonlinear physical variables of the environment. When compared to statistical approaches, it makes fewer assumptions and processes information quickly. Also, it operates on the parallel massive machine principle based on the simplification of biological neurons (Beale *et al.,* 2010). Because of their capacity to imitate and replicate nonlinear events, ANNs have been employed in several studies to examine the behavior of the ionosphere. These include, Tulasi Ram *et al.,* (2018), Gowtam and Ram, (2017), Hu and Zhang (2018), Kim *et al.,* (2021) and Moon *et al.,* (2020).

Gowtam and Ram, (2017) built the Immune layer Artificial Neural Network (ANNIM) model, which is an upgraded ANN model capable of replicating variations in peak electron density (NmF2) and hmF2 across both time and space, as well as the increase in equatorial hmF2 that occurs after sunset because of the pre-reversal amplification of the zonal electric field. Tulasi Ram *et al.* (2018) showed that during various solar activity times, the anticipated NmF2 and hmF2 display strong correlation with ground-based digisonde data, demonstrating the ability of controlled simulations to distinguish between the impacts of solar irradiance and recurrent geomagnetic activity. Hu and Zhang (2018) developed a novel regional hmF2 forecast model for Australia using ionosonde data and the bidirectional Long Short-Term Memory (bi-LSTM) technique. The performance of this model was compared to that of the AMTB-2013, SHU-2015, ANN, and LSTM models, among other well-known models. According to the results, the model outperformed the previous four models in the first five hours. Kim *et al.* (2021) created a deep learning model for geomagnetic storms utilizing the new training data set, which included redesigning input parameters and hyper-parameters. They employed the hmF2 and foF2 parameters from the Jeju ionosonde (33.43°N, 126.30°E) as input parameters for the LSTM model. The results showed that the LSTM storm model outperformed the LSTM quiet

model, SAMI2, and IRI-2016 models during the geomagnetic storms. Moon *et al.* (2020) developed the LSTM model, a regional ionospheric model, to predict hourly hmF2 over the Jeju station (33.43°N, 126.30°E) for up to 24 hours using a deep learning method to anticipate the F2 parameters. The outputs of the LSTM model and the IRI-2016 model were compared during geomagnetically active and inactive phases. It was discovered that the LSTM model performed best on calm days. During geomagnetic storm days, the IRI-2016 model beat the LSTM model, which they associated this with few training data available at those times.

This research endeavor is particularly motivated by the uneven distribution of ionosondes, the instruments used to measure hmF2, across the globe, with a notable scarcity in the southern African region. Additionally, even at the available ionosonde stations in this region, there are significant gaps in the database due to infrastructural challenges, such as power outages and maintenance requirements. In light of these limitations and data gaps, the development of a regional model for hmF2 predictions using neural network techniques emerges as a promising solution. Neural networks possess the remarkable ability to learn and model complex, nonlinear relationships from data, making them well-suited for addressing the dynamic nature of the ionosphere.

Materials and Methods

Source of data

In this work, data from ionosondes, specialized radar equipment that collects echoes from the ionosphere across a broad frequency band, were used as targets to the model. These ionosonde readings were taken from three ionosondes which were Grahamstown (33.32°S, 26.50°E), Madimbo (22.4°S, 30.90°E) and Louisvale (28.5°S, 21.2°E) located in South Africa and cover a significant record from 2005 to April 2023. The data were taken in 30-minute interval, on daily basis and on quiet days in order to uncover intricate ionospheric processes.

Identification of the input parameters of the model was carried out through searching different scientific papers and articles describing the factors that affect the hmF2. The factors that were identified were: year, day of the year, time in 30-minute intervals, sunspot number, Solar flux at 10.7cm (F10.7), longitude, latitude, horizontal wind speed, critical frequency (foF2), the planetary K-index (Kp) and the planetary A-index (Ap) (Tulasi Ram *et al.*, 2018, Moon *et al.*, 2020, Tulunay *et al.*, 2006 and Watthanasangmechai, 2012). The statistical approach was used to determine the correlations between the identified factors with the hmF2. Horizontal wind model 2014 and recorded F10.7 solar flux and Kp index were obtained from Dominion Radio Astrophysical Observatory (DRAO) through (https://www.cadc-ccda.hia-iha.nrc-cnrc.gc.ca/en/). Additionally, in order to strengthen the validity of this research, the data from the International Reference Ionosphere (IRI-2020) model which are available at https://irimodel.org/ were compared with the results of the developed model.

The following pair of function was used to define the parameters used as ANN inputs.

$$y_k = f(\sum_{k=1}^{1} w_k f(\sum_{i=1}^{m}(x_i \times w_{ki}) + b_k) + w_o) \quad (1)$$

In this equation, y_k denotes the output of the neuron, f denotes the activation function, m denotes the number of input parameters, $x_i$is the i-th input parameter, w_{ki} denotes the i-th synaptic weight, w_o denotes the original synaptic weight, and b_k denotes the bias.

The developed model was trained using the Levenberg-Marquardt algorithm because it has high speed and is better suited for variety of problems than other common algorithms (Ahmed, 2018). A higher learning rate of 0.4 was used in training the model. It was used to speed up a training process, since, if it is small tends to slow and lengthy the training process and if it is too higher may result into the saturation of output or swing of output across the desired output (Zhou *et al.*, 2022). Two stopping standards were employed during the training phase to prevent over-fitting of the models: a minimum training error and a maximum number of epochs, which were fixed at 0.00000001 and 1000, respectively. Over-fitting degrades the generalization ability of the model in prediction (Gnana and Deepa, 2013).

To design a suitable neural network architecture, five hmF2 prediction networks were trained by using thirteen years of the inputs data from 2005–2018, then tested by using data from January 2019 to December 2022, and validated by data for January 2023 to April 2023. Combinations of inputs to design the best network was done by considering the statistical techniques such as correlation analysis and mutual information as suggested by Hegland (2001). By systematically excluding or including specific inputs, the researchers can assess which variables have the greatest impact on the model's predictive performance. This helps identify the most important predictors and allows the model to focus on the key drivers of the target output.

Network 1 was designed with the aim of determining the prediction capability of the combination of inputs of year, day of the year, time in 30 minute intervals, SSN, F10.7, foF2, horizontal wind speed, Kp and Ap.

Network 2 was designed to determine if the combination of year, day of the year, time in 30 minute intervals, SSN, F10.7, foF2, longitude and latitude, can be used to generate hmF2 prediction model.

Network 3 was designed to determine if the combination of year, day of the year, time in 30 minute intervals, SSN, F10.7, longitude, latitude, horizontal wind speed, foF2, Kp and Ap can be used to generate hmF2 prediction model.

Network 4 was designed with the purpose of determining if the combination of year, day of the year, time in 30 minutes intervals, SSN, F10.7, longitude, latitude, horizontal wind speed, Kp and Ap, as inputs can produce a good result for hmF2 prediction model.

Network 5 was trained by nine input parameters, which are Year, day of the year, time in 30 minutes intervals, longitude, latitude, horizontal wind speed, foF2, Kp and Ap.

The training of the neural network model involved simulating five different network architectures by systematically varying the number of neurons in the hidden layer from 1 to 100 for each network. In order to determine if the model is appropriate, the Root Mean Squared Error (RMSE) and coefficient of determination (R^2) were employed which are given by equations (1) and (2) respectively;

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - y_i^*)^2} \tag{2}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(y_i - y_i^*\right)^2}{\sum\left(y_i - y_m^*\right)^2} \tag{3}$$

Where y_i is actual height value, y_i^* is predicted height value, y_m^* represents average height and n represents number of samples. Plots were used to compare the actual value and the predicted value. This demonstrated how well anticipated values match real values.

To test the suitability of the neural network in predicting hmF2 compared to the IRI-2020 model, the predicted hmF2 values from ANN were compared to hmF2 values from the AMTB-2013 and SHU-2015 models, which are available in the IRI-2020 model. The data used included the 2009 and 2014 data from the IRI-2020 model options, as well as ionosonde data from the Grahamstown station. The year 2009 represented a period of minimum solar activity, while 2014 represented a year of maximum solar activity. For these selected years, two seasons were analyzed - summer and winter. In the southern hemisphere where South Africa is located, the summer months are December, January, and February, while the winter months are June, July, and August.

Results

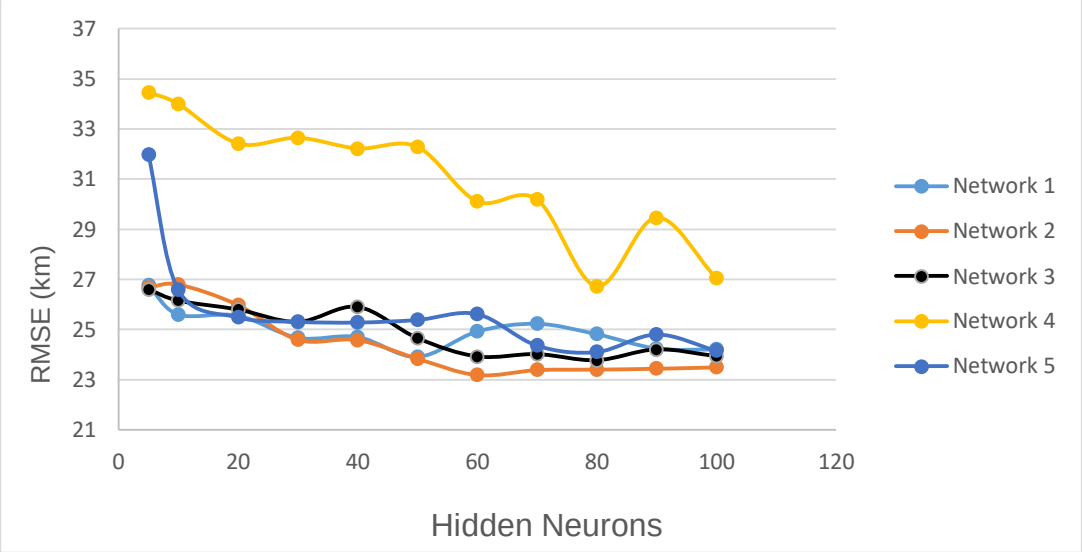
This section presents the results of the five trained neural network architectures that were developed for the purpose of predicting hmF2. The goal is to evaluate the performance of these five networks in order to identify the one that provides the most optimal results.

Figure 1 shows the performance of model using all five networks based on RMSE. It can be observed from the figure that, there is a continual

random decrease of RMSE, with the increase number of neurons in the hidden layer. As evident from the figure, network 2 exhibited the best performance among the five networks, as its RMSE values were consistently lower than those of the other four networks. This suggests that network 2 provided the closest agreement between the measured values and the values estimated by the neural network model. Network 2 portrayed a minimum prediction error at 60 neurons in the hidden layer with the RMSE of 23.197 km. Beyond 60 neurons, RMSE slightly increased, indicating the increase of prediction error of the model as the number of neurons increases beyond 60.

Figure 1

Performance of the five networks in terms of RMSE



Performance of network 2 was based on the fact that, the Kp and Ap indices may not have a strong direct relationship with hmF2 and could have introduced more noise or complexity to the model. By excluding these parameters, network 2 was able to focus on the more relevant inputs without being potentially hindered by the weaker predictive power of the Kp and Ap indices (Gulyaeva and Arikan, 2021; Martyshko and Perevalova 2011).

On the other hand, network 4 presented the poorest results with higher RMSE, where the minimum prediction error of this network occurs at 80 neurons with RMSE=26.718 km. which may have been contributed by the exclusion of hmF2 parameter as an input parameter. The foF2 has a strong direct correlation with hmF2 and its exclusion can significantly hinder the model's predictive capability (Zeng *et al.*, 2016).

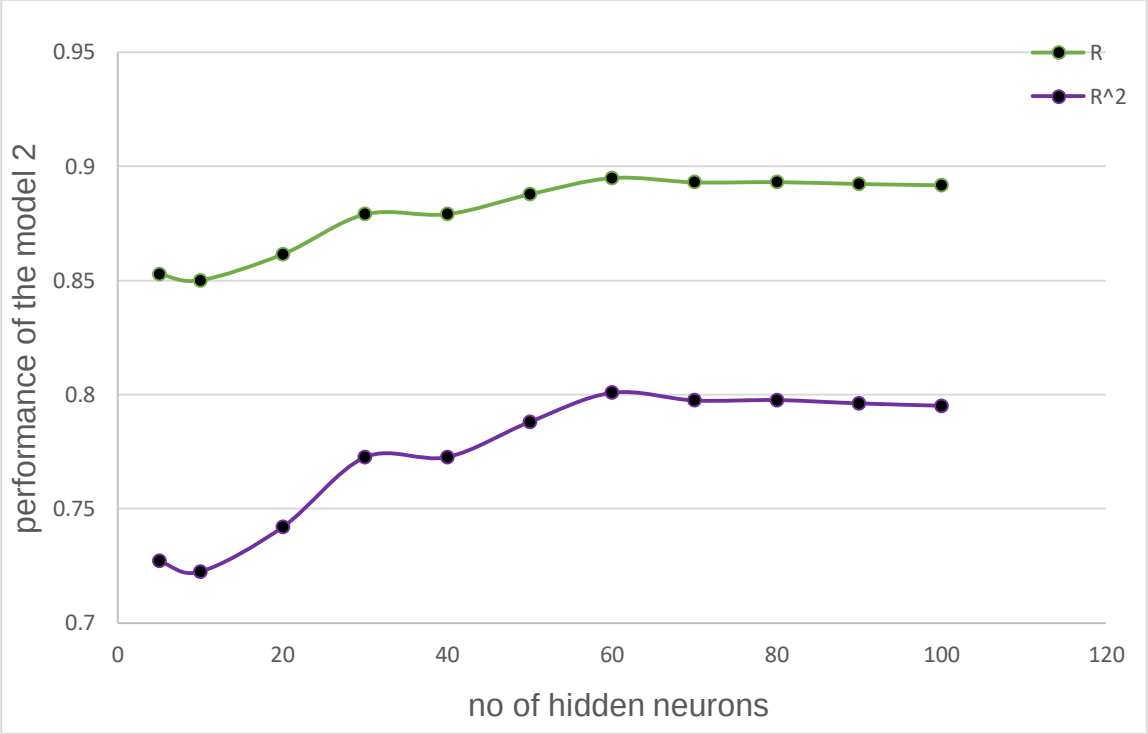
Therefore, the network with input data; year, day of the year, time in 30 minute intervals, SSN, F10.7, longitude, latitude and critical frequency (foF2), which has 60 hidden neurons is considered the best model architecture for hmF2 prediction. The inclusion of foF2 as an input parameter is crucial, as it has a strong direct correlation with hmF2 and its inclusion can significantly improve the model's predictive capability. The other input parameters, such as year, day of the year, time, SSN, F10.7, longitude, and latitude, provide relevant information for the hmF2 prediction without introducing redundancy. The 60 hidden neurons in the model architecture appear to provide an optimal balance between model complexity and predictive performance, avoiding the potential

issues of overfitting or poor generalization that can arise from excessively complex models.

The model configuration of Network 2, which has the best network architecture, demonstrates a high level of performance in terms of R^2 and R as shown by Figure 2. The result showed that at optimal hidden neurons the value of R^2 and R were 0.80075 and 0.8948 respectively. According to Hinkle *et al.* (2003), if R^2 and R lies between 0.1 and 0.3, it indicates a weak correlation, between 0.4 and 0.6 it indicates a moderate correlation and when it lies between 0.7 and 1 it indicates a strong correlation. Therefore, the values of R^2 and R from the model indicated a strong correlation showing a good agreement between predicted and actual peak height of electron density (hmF2). This suggests that the model is highly effective in capturing the relationship between the input parameters and the target hmF2 variable.

Figure 2

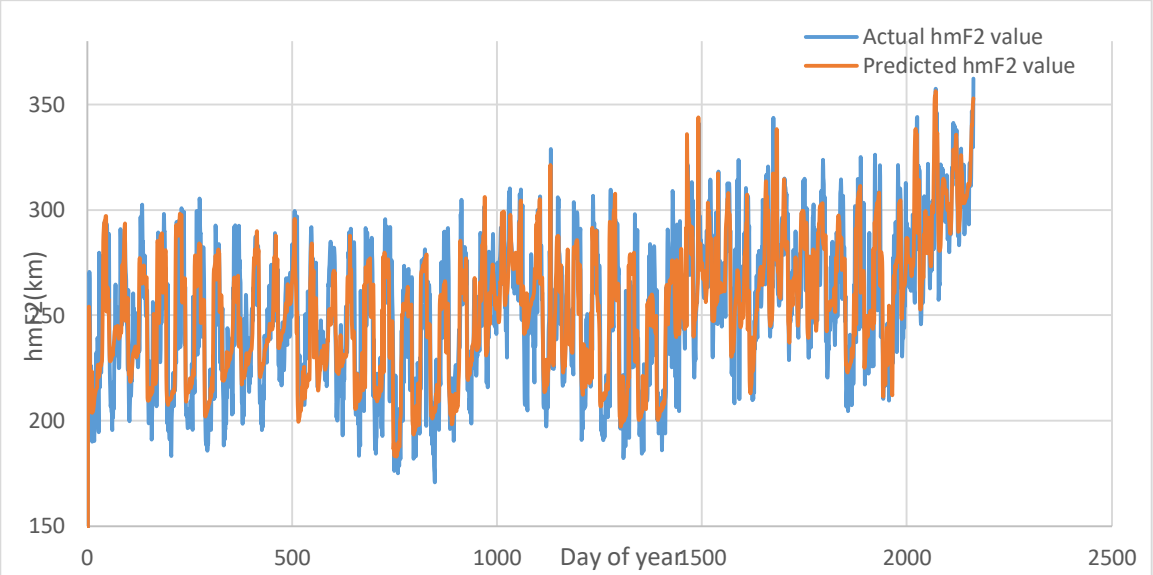
Performance results of model based on R^2 and *R*



The information presented in Figure 3 suggests that the model's predicted hmF2 values closely match the actual hmF2 values in the test dataset, with only minor variations between the two. Specifically, the figure shows that the predicted hmF2 values closely track the changes and fluctuations observed in the actual hmF2 test data. This indicates that the model has effectively captured the underlying relationships and dynamics that govern the behavior of the hmF2 parameter. The small differences or variations between the predicted and actual hmF2 values suggest that the model is able to respond and adapt to the changes in the hmF2 parameter. This level of responsiveness and accurate prediction, even in the face of fluctuations, demonstrates the model's effectiveness in modeling the complex ionospheric processes that influence the hmF2 parameter.

Figure 3

Predicted hmF2 by model as compared to the test data set



This result agrees with the study of Tulasi Ram *et al.*, (2018) who uses the improved two-dimensional Artificial Neural Network-Based Ionospheric Model (ANNIM) to predict the peak electron density (hmF2). According to their results, the RMSE for 2002, 2009, 2014, and 2016 were 29 km, 25 km, 31 km and 33 km respectively. Gowtam and Ram (2017) used an Artificial Neural Network-Based Ionospheric Model (ANNIM) to predict NmF2 and hmF2 using long-term data set of formosat-3/cosmic Radio occultation observations and the RMSE obtained was 27.9 km with R = 0.77. The improved ANNIM well captured the temporal (local time, seasonal, and solar cycle epoch) and spatial (latitude and longitude) variations of F2-layer peak electron density (hmF2).

Assessing predicting efficient of the developed model in comparison with IRI 2020 model

To assess the predicting efficient of the developed model in comparison with the widely used IRI 2020 model, the two IRI 2020 model options; AMTB-2013 and SHU2015, were used. The data used included the 2009 and 2014 data from the IRI-2020 model options, as well as ionosonde data from the Grahamstown station. The year 2009 represented a period of minimum solar activity, while 2014 represented a year of maximum solar activity. For these selected years, two seasons were analyzed - summer and winter. In the southern hemisphere where South Africa is located, the summer months are December, January, and February, while the winter months are June, July, and August. The results are presented in Figures 4 and 5.

Figure 4

Observed hmF2 compared with ANN, AMTB-2013 and *SHU-2015 hmF2 predictions for (a) January 2009 (b) February 2009 (c) December 2009 (d) June 2009 (e) July 2009 (f) August 2009*

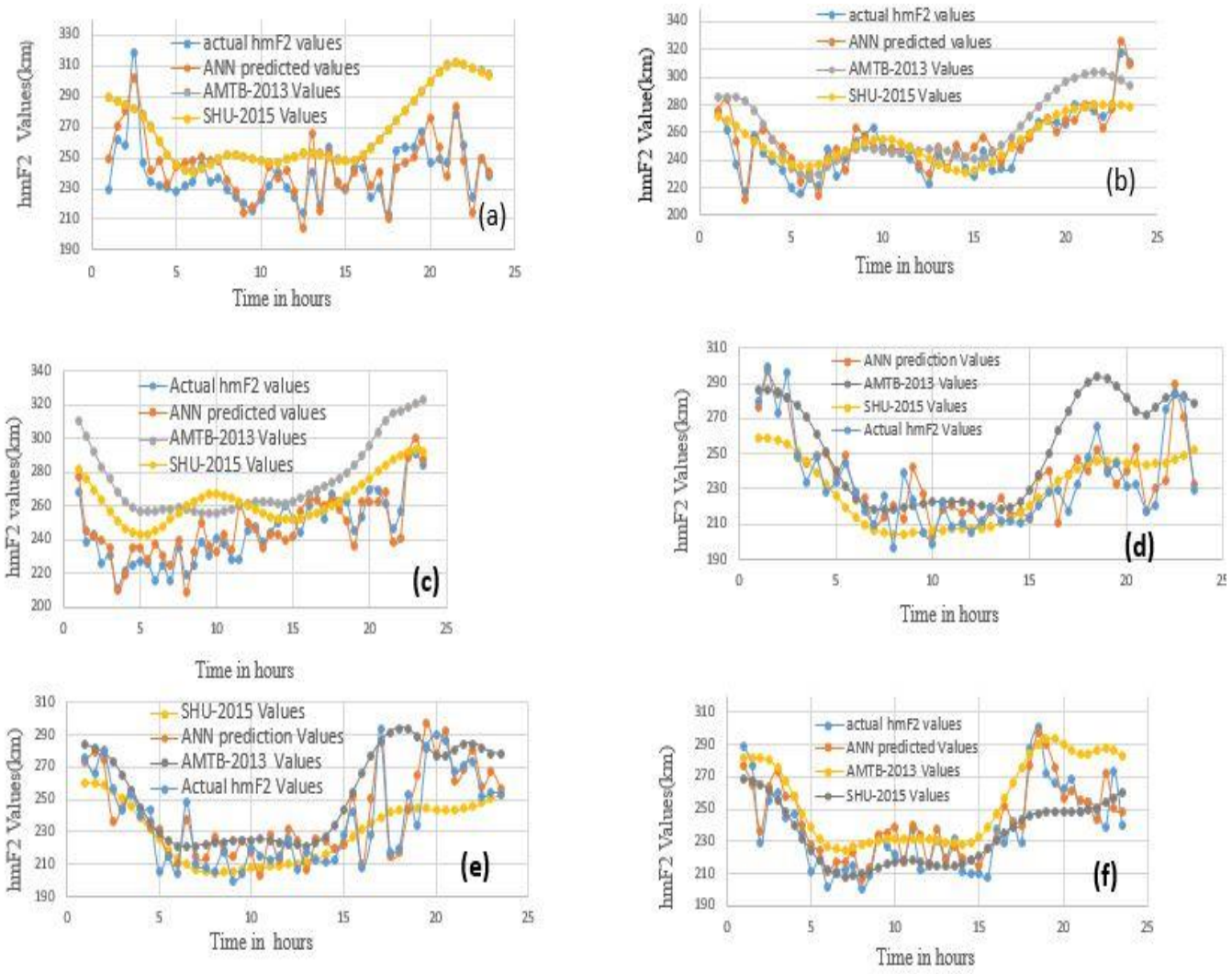


Figure 4(a) illustrates the hmF2 predictions for January 2009, displaying a positive bias in all three model options with overestimation of actual values. The RMSE of the ANN is 11.72 km, whereas the RMSE of AMTB-2013 and SHU-2015 are 18.87 km and 15.63 km, respectively. The ANN perform better compared to AMTB-2013 and SHU-2015. Even though it is summer in South Africa in January, this result suggest that ANN shows good adaptability to high temperature seasons compare to AMT-2013 and SHU-2015.

Figure 4(b) shows the hmF2 predictions for February 2009. The result showed that, all three models significantly manage to predict hmF2 values for February, with the ANN model demonstrating higher accuracy compared to the other models. However, significant differences emerge in their predictive accuracy. While AMTB-2013 had RMSE=17.74 km, the ANN and SHU-2015 display better results with the RMSE values of 10.09 km and 13.55 km respectively. ANN performed better compared to AMTB-2013 and SHU-2015. All models were managed to capture the change in hmF2 values since the temperature was decreasing at ending of summer seasons.

Figure 4(c) shows the hmF2 predictions for December 2009, which was the start of the summer season. While SHU-2015 and AMTB-2013 displayed RMSE of =14.81 km and 10.93 km respectively, ANN outperformed them, with the lowest RMSE of 9.65 km suggesting superior and accuracy, with a moderate overestimation bias. Due to hmF2 fluctuation caused by summer temperature SHU-2015, and ANN managed to capture the changes, this lead to lower RMSE values compare to AMTB-2013. These results show that ANN responded effectively in season of high temperature.

Figure 4(d) illustrates the hmF2 predictions for June 2009. All the three models showed an overestimation of actual values. The RMSE of the AMTB-2013 in predicting hmF2 was 16.49 km, whereas that of SHU-2015 and ANN were 10.06 km, and 12.35 km respectively. This result showed that ANN responded effectively than AMTB-2013 and SHU-2015 since it managed to capture the fluctuations of hmF2 in response to change in seasons.

Figure 4(e) illustrates the hmF2 predictions for July 2009. The RMSE of the ANN was 11.08 km, whereas that of AMTB-2013 and SHU-2015 were 19.48 km and 14.08 km, respectively. The ANN manage to capture the fluctuation of hmF2 values compared to AMTB-2013 and SHU-2015. Compared to June, the AMTB and SHU-2015 showed the increase in their RMSE values, showing the decrease in the accuracy, while ANN RMSE value decreased from 12.35 km in June to 11.08 km in July, showing an increase in the model accuracy.

Figure 4(f) illustrates the hmF2 predictions for August 2009. The RMSE of the ANN are 14.32 km, whereas the RMSE of AMTB-2013 and SHU-2015 are 17.30 km and 15.86 km, respectively. By comparing with June and July 2009, the models overestimated actual values due to positive biases. This implies that, these models tend to overemphasize particular input parameters such SSN during the prediction process. This shows how complicated the ionosphere is and how difficult it is to adequately capture its activity in predicting models (Pesnell, 2008). It also implies that AMTB-2013 and SHU-2015 models need to be improved or refined in order to more accurately represent the physics regulating ionospheric variability, particularly during times of decreased activity like 2009.

Figure 5(a) illustrates the hmF2 predictions for January 2014, the year of high solar activity where AMTB-2013, SHU-2015 and ANN all displaying a positive bias with SHU-2015 tend to overestimation of actual values. Significantly, the RMSE of the ANN is 9.73 km, whereas the RMSE of AMTB-2013 and SHU-2015 are 15.67 km and 12.99 km, respectively. ANN had lower value of RMSE than AMTB-2013 and SHU-2015, implying that, the ANN model managed to show high accuracy in high solar activity during the summer seasons.

Figure 5(b) illustrates the hmF2 predictions for February 2014, where AMTB-2013 and SHU-2015 displaying a positive bias with overestimation of actual values, while ANN display negative bias by underestimating the actual values. The RMSE of the ANN is 10.14 km, whereas the RMSE of

AMTB-2013 and SHU-2015 are 14.62 km and 12.20 km, respectively. ANN displayed less RMSE, showing more accuracy in predicting hmF2 compared to AMTB-2013 and SHU-2015.

Figure 5

Observed hmF2 compared with ANN, AMTB-2013 and *SHU-2015 hmF2 predictions for (a) January 2014 (b) February 2014(c) December 2014 (d) June 2014 (e) July 2014 (f) August 2014*

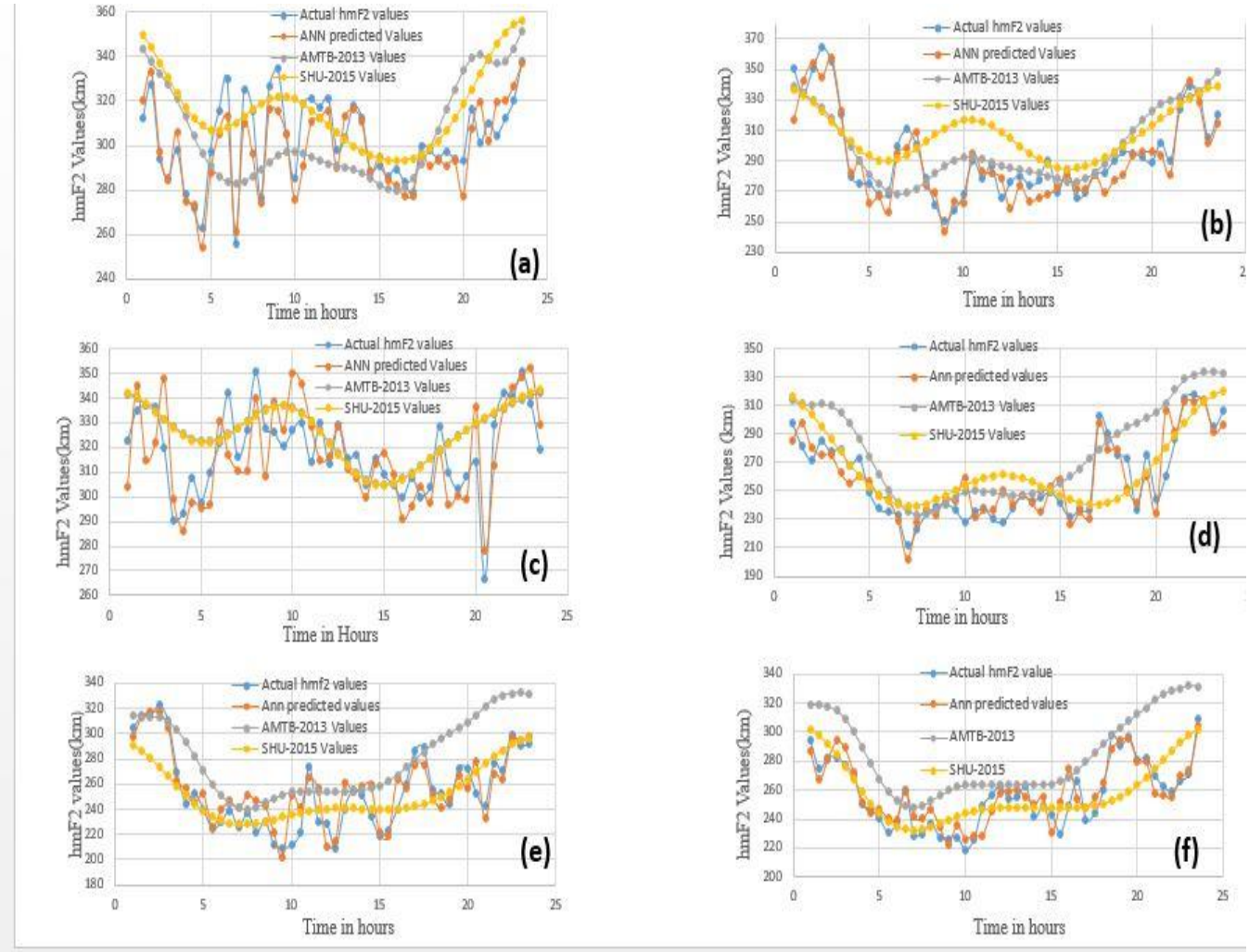


Figure 5(c) illustrates the hmF2 predictions for December 2014. The RMSE of the ANN is11.27 km, whereas the RMSE of AMTB-2013 and SHU-2015 are 14.69 km and 12.29 km, respectively. The ANN model exhibited a negative bias, indicating that it tends to underestimate the true hmF2 values. But in contrast to January and February, its RMSE for December is slightly higher, suggesting that the ANN in the beginning of summer does not manage well the change in the seasonal temperature hence it still need some improvement to adapt with the change the seasons.

Figure 5(d) illustrates the hmF2 predictions for June 2014. Among the models, the ANN model

has the lower RMSE of 11.14 km, indicating good prediction. The RMSE of AMTB-2013 and SHU-2015 were 19.05 km and 14.6 km, respectively. Comparing to June 2009, where ANN had a RMSE of 12.35 km, there is slight improvement in the model for capturing the hmF2 values, demonstrating its ability to capture ionospheric dynamics even in high solar activity periods.

Figure 5(e) illustrates the hmF2 predictions for July 2014. The RMSE of the ANN was 13.79 km, whereas the RMSE of AMTB-2013 and SHU-2015 were 18.35 km and 14.52 km, respectively. Compared to June 2014, the ANN model RMSE appears to be slightly higher, suggesting decrease in predictive accuracy during this month. On the other hand, both AMTB-2013 and SHU-2015 models demonstrated RMSE values of 19.05 km and 14.02 km, respectively, in July 2014. These results highlight how crucial it is to continuously validate and improve modeling techniques in order to improve their capacity to precisely represent ionospheric dynamics, especially in light of fluctuating solar activity levels.

Figure 5(f) illustrates the hmF2 predictions for August 2014. The RMSE of the ANN was 9.02 km, whereas the RMSE of AMTB-2013 and SHU-2015 were 14.83 km and 11.10 km, respectively. ANN continued to show lower RMSE values compared to AMTB-2013 and SHU-2015 indicating higher overall accuracy. The AMTB-2013 and SHU-2015 model's performance seems to improve accuracy with decrease in RMSE from June to August, which indicates a greater adaptivity from observed values. The dynamic character of ionospheric modeling and the possible influence of temporal changes in ionosphere circumstances on model performance are highlighted by these variations in RMSE values (Huang and Reinisch, 2006). The RMSE for 2009 and 2014 for all models during the summer for the month of December, January and February ANN display the values of 9.65 km, 11.72 km and 10.09 km respectively while AMTB-2013 values were 14.83 km, 18.87 km, and 17.74 km respectively and SHU-2015 values were 10.93 km, 15.63 km, and 13.55 km during winter for the month June, July and August ANN display the values of 10.06 km, 11.08 km and 14.32 km, respectively while AMTB-2013 values were 16.49 km, 17.48 km and 17.30 km respectively and SHU-2015 values were 12.35 km, 14.80 km, and 15.86 km.ANN over perform other models with minimum RMSE values followed by SHU-2015 and finally AMTB-2013. Again in 2014 the year of high solar Activity during the summer for the month of December, January and February ANN display the values of 11.27 km, 9.73 km and 10.14 km respectively while AMTB-2013 values were 14.69 km, 15.67 km, and 14.62 km respectively and SHU-2015 values were 12.20 km, 12.29 km, and 11.50 km during winter for the month June, July and August ANN display the values of 11.14 km, 13.79 km and 9.02 km, respectively while AMTB-2013 values were 19.05 km, 18.35 km and 14.83 km respectively and SHU-2015 values were 14.60 km, 14.52 km, and 11.10 km. ANN over perform other models with minimum RMSE values followed by SHU-2015 and finally AMTB-2013.

Discussion

The capability of ANN to predict hmF2 has been achieved by other researchers globally, such as Gowtam and Ram (2017), developed an artificial neural network-based two-dimensional ionospheric model (ANNIM) to predict the ionospheric hmF2 using the long-term (2006–2015) and global data of F3/C radio occultation observations. The linear regression coefficients between the ANNIM results and actual F3/C data were 0.77 and the RMSE was 27.9 km for hmF2.

Habarulema *et al.* (2021) developed a global 3-D electron density reconstruction model based on radio occultation data and neural networks for predicting the foF2 and hmF2. The result showed the IRI 2016 model's performance was superior in all longitude sectors compared to 3D-NN model developed. The main reason for this is that the hmF2 option within the IRI 2016 that we have compared with was developed based on a combination of COSMIC and ionosonde data (Shubin, 2015). It is thus expected that its performance in estimating ionosonde hmF2 data will be greater than a model which used only COSMIC dataset. the developed 3D-NN model was mostly accurate in estimating ionosonde foF2 than the IRI 2016 model in the Europe, African and American sectors while the IRI 2016 model was found to be more accurate in the Asian sector.

Moon *et al.* (2020) develop a regional ionospheric model to predict hourly hmF2s over the Jeju station (33.43°N, 126.30°E) for up to 24 hours by using a deep learning method. To evaluate the model performance, root mean square error (RMSE) and the correlation coefficient (CC) were used as the performance skill scores between the model prediction and the measurement during the 338 days of 2017–2018. hmF2 model, having 41 hidden neurons and 24 batch sizes, preformed with an RMSE of 23.8 km and a CC of 0.80. The LSTM model results were compared with the results of the TIE-GCM and the IRI-2016 model for geomagnetically quiet and active periods. When the RMSEs of the hmF2 predictions were compared, the percentage improvements of the LSTM model were 45% more than those of the other models during geomagnetically quiet days.

Tulasi Ram (2018) developed ANNIM by assimilating additional ionospheric data from CHAMP, GRACE RO, worldwide ground-based Digisonde observations, and by using a modified spatial gridding approach based on the magnetic dip latitudes. The improved ANNIM better reproduces the spatial and temporal variations of hmF2, including the post sunset enhancement in equatorial hmF2 associated with the pre-reversal enhancement in the zonal electric field. The comparisons of hmF2 prediction by ANNIM and IRI-2016 prediction values with the ground-based Digisonde observations over Jicamarca showed that, in 2002 the RMSE values were 29 km and 36 km with regression coefficients of 0.89 and 0.84 respectively. In their study the result obtain in 2009 the RMSE values were 25 km and 32 km with regression coefficients of 0.81 and 0.72 respectively. In 2014, the RMSE values were 31 km and 43 km with regression coefficients of 0.91 and 0.90 respectively. In 2016 the RMSE values were 33 km and 35 km with the regression coefficients of 0.85 and 0.84 respectively.

By comparing the result obtained from other studies and the ANN performance it can be conclude that the proposed ANN model can be used to predict hmF2 layer for South Africa region.

Conclusion

The developed hmF2 prediction model based on ANN was able to detect the non-linear relationship that exists between input parameters and hmF2. It managed to develop a mapping between inputs data with the desired hmF2. Also, the model showed high degree of adaptability and consistence despite of high fluctuation of hmF2. For this reason, a model based on ANN is established as an appropriate tool for predicting hmF2 for Grahamstown, Madimbo and Louisvale ionosondes in South Africa. The RMSE values obtained by ANN exhibit good linear correlations with the ground-based ionosondes data during all the years. Further, the RMSE values in ANN are convincingly for predictions of hmF2. Therefore, from the comparisons of ANN, AMTB-2013, and SHU-2015 with the ground-based ionosondes observations, it can be concluded that the learning efficiency of the above ANN architecture is good, and the ANN predictions are as good as SHU-2015 with the slightly better difference by the ANN. This study concentrated on developing the hmF2 model for South Africa using the ANN technique. However, this study may be extended to other regions of Africa for predicting hmF2 and other ionospheric parameters. In this study, a data set of eighteen years was used. Therefore, in future, further studies may be conducted by using more long-period data set and other ionosondes in the African region to ensure a more accurate prediction of ionospheric parameters.

Recommendation

This study focused on developing an ANN model for predicting the hmF2 parameter specifically for the South African region. However, the potential of this approach could be extended to other regions across Africa for forecasting hmF2 as well as other ionospheric parameters.

In this particular study, a dataset spanning eighteen years was utilized. Moving forward, future research in this area may benefit from incorporating an even longer-duration dataset and incorporating data from additional ionosondes located throughout the African continent. This expanded data coverage could

lead to further improvements in the accuracy of predicting various ionospheric parameters.

Acknowledgements

The authors would like to acknowledge the IRI 2020 data, Horizontal Wind Model 2014, recorded F10.7 solar flux, Ap and Kp index are obtained from DRAO through (https://www.cadc-ccda.hia-iha.nrc-cnrc.gc.ca/en/).Also the foF2, hmF2 for ionosondes are abtain from Global Ionospheric Radio Observatory (GIRO) through (giro.uml.edu). Finaly the AMTB-2013 and SHU-2015 data for 2009 and 2014 were obtained from (https://irimodel.org/)

References

Ahmed, Z. R. (2018). *Rainfall prediction using machine learning techniques*. PhD diss.

Anderson, D. N., Buonsanto, M. J., Codrescu, M., Decker, D., Fesen, C. G., Fuller-Rowell, T. J., Reinisch, B. W., Richards, P. G., Roble, R. G., Schunk, R. W., & others. (1998). Intercomparison of physical models and observations of the ionosphere. *Journal of Geophysical Research: Space Physics*, *103*(A2), 2179-2192.

Bailey, G., Balan, N., & Su, Y. (1997). The Sheffield University plasmasphere ionosphere model---A review. *Journal of Atmospheric and Solar-Terrestrial Physics*, *59*(13), 1541-1552.

Beale, M. H., Hagan, M. T., & Demuth, H. B. (2010). Neural network toolbox. *User's Guide, MathWorks*, *2*, 77-81.

Biktash, L., & Azizova, Z. (2009). The equatorial scintillations and space weather effects on its generation during geomagnetic storms. *The Institution of Engineering and Technology 11th International Conference on Ionospheric radio Systems and Techniques (IRST 2009)*.

Bilitza, D. (2001). International reference ionosphere 2000. *Radio science*, *36*(2), 261-275.

Bilitza, D., Altadill, D., Truhlik, V., Shubin, V., Galkin, I., Reinisch, B., & Huang, X. (2017). International Reference Ionosphere 2016: From ionospheric climate to real-time weather predictions. *Space weather*, *15*(2), 418-429.

Bilitza, D., & Xiong, C. (2021). A solar activity correction term for the IRI topside electron density model. *Advances in Space Research*, *68*(5), 2124-2137.

Bothmer, V., & Daglis, I. A. (2007). *Space weather: physics and effects*. Springer Science & Business Media.

Cander, L. R., Milosavljevic, M., Stankovic, S., & Tomasevic, S. (1998). Ionospheric forecasting technique by artificial neural network. *Electronics Letters*, *34*(16), 1573-1574.

Cesaroni, C., Spogli, L., Aragon-Angel, A., Fiocca, M., Dear, V., De Franceschi, G., & Romano, V. (2020). Neural network based model for global Total Electron Content forecasting. *Journal of Space Weather and Space Climate*, *10*, 11.

Du, K.-L. (2010). Clustering: A neural network approach. *Neural networks*, *23*(1), 89-107.

Field, D. (2018). *A new empirical climatological model of ionospheric foF2 and hmF2 and review of the International Reference Ionosphere*.

Gnana, S. K., & Deepa, S. N. (2013). Review on methods to x number of hidden neurons in neural networks. mathematical problems in engineering. *Hindawi Publish Corp*.

Goodarzi, M. (2020). Application and performance evaluation of time series, neural networks and HARTT models in predicting groundwater level changes, Najafabad Plain, Iran. *Sustainable Water Resources Management*, *6*(4), 67.

Habarulema, J. B., Okoh, D., Burešová, D., Rabiu, B., Tshisaphungo, M., Kosch, M., Häggström, I., Erickson, P. J., & Milla, M. A. (2021). A global 3-D electron density reconstruction model based on radio occultation data and neural networks. *Journal of Atmospheric and Solar-Terrestrial Physics*, *221*, 105702.

Hegland, M. (2001). Data mining techniques. *Acta numerica*, *10*, 313-355.

Hu, A., & Zhang, K. (2018). Using bidirectional long short-term memory method for the height of F2 peak forecasting from ionosonde measurements in the Australian region. *Remote Sensing*, *10*(10), 1658.

Huba, J., Joyce, G., & Fedder, J. (2000). Sami2 is Another Model of the Ionosphere (SAMI2): A new low-latitude ionosphere model.

Journal of Geophysical Research, Space Phys, 23035-23053.

Huba, J., & Liu, H. L. (2020). Global modeling of equatorial spread F with SAMI3/WACCM-X. *Geophysical Research Letters*, *47*(14), e2020G.

Khoshgoftaar, T. M., Seiffert, C., Van Hulse, J., Napolitano, A., & Folleco, A. (2007). Learning with limited minority class data. *Sixth International Conference on Machine Learning and Applications (ICMLA 2007)*, 348-353.

Kim, J. H., Kwak, Y. S., Kim, Y., Moon, S. I., Jeong, S. H., & Yun, J. (2021). Potential of Regional Ionosphere Prediction Using a Long Short-Term Memory Deep-Learning Algorithm Specialized for Geomagnetic Storm Period. *Space weather*, *19*(9), e2021S.

Leitinger, R., Hochegger, G., Hafner, J., Radicella, S., & Nava, B. (2000). NeQuick, COSTprof, NeUoG-plas- A family of ionospheric models. *Kleinheubacher Berichte*, *43*, 20-25.

Moon, S., Kim, Y. H., Kim, J.-H., Kwak, Y.-S., & Yoon, J.-Y. (2020). Forecasting the ionospheric F2 Parameters over Jeju Station (33.43° N, 126.30° E) by Using Long Short-Term Memory. *Journal of the Korean Physical Society*, *77*(12), 1265-1273.

Nkoana, R. (2011). Artificial neural network modelling of flood prediction and early warning. *Master Degree. Bloemfontein: University of the Free State*.

Pedroni, N., Zio, E., & Apostolakis, G. E. (2010). Comparison of bootstrapped artificial neural networks and quadratic response surfaces for the estimation of the functional failure probability of a thermal--hydraulic passive system. *Reliability Engineering \& System Safety*, *95*(4), 386-395.

Radicella, S. M. (2009). The NeQuick model genesis, uses and evolution. *Annals of Geophysics*, *52*(3-4), 417-422.

Saba, T., Rehman, A., & AlGhamdi, J. S. (2017). Weather forecasting based on hybrid neural model. *Applied Water Science*, *7*, 3869-3874.

Sai Gowtam, V., & Tulasi Ram, S. (2017). (2017). No TitleSai Gowtam, V., & Tulasi Ram, S. (2017). An Artificial Neural Network based Ionospheric Model to predict NmF2 and hmF2 using long-term data set of FORMOSAT-3/COSMIC radio occultation observations: Preliminary results. *Journal of Geophysical Research: Space Physics, 122, 11,743–11,755. https://doi.org/10.1002/2017JA024795*.

Schunk, R. W., Scherliess, L., Sojka, J. J., Thompson, D. C., Anderson, D. N., Codrescu, M., Minter, C., Fuller-Rowell, T. J., Heelis, R. A., & Hairston, M. (2004). Global assimilation of ionospheric measurements (GAIM). *Radio science*, *39*(1), 1-11.

Shubin, V. N. (2015). Global median model of the F2-layer peak height based on ionospheric radio-occultation and ground-based Digisonde observations. *Advances in Space Research*, *56*(5), 916-928.

Suzuki, K. (2013). *Artificial neural networks: Architectures and applications*. BoD--Books on Demand.

Tulasi Ram, S., Sai Gowtam, V., Mitra, A., & Reinisch, B. (2018). The improved two-dimensional artificial neural network-based ionospheric model (ANNIM). *Journal of Geophysical Research: Space Physics*, *123*(7), 5807-5820.

Tulunay, E., Senalp, E. T., Radicella, S. M., & Tulunay, Y. (2006). Forecasting total electron content maps by neural network technique. *Radio science*, *41*(04), 1-12.

Watthanasangmechai, K., Supnithi, P., Lerkvaranyu, S., Tsugawa, T., Nagatsuma, T., & Maruyama, T. (2012). TEC prediction with neural network for equatorial latitude station in Thailand. *Earth, Planets and Space*, *64*(6), 473-483.

Yegnanarayana, B. (2009). *Artificial neural networks*. PHI Learning Pvt. Ltd.

Zhou, C., Yang, L., Su, X., & Li, B. (2022). Neural network-based ionospheric modeling and predicting---To enhance high accuracy GNSS positioning and navigation. *Advances in Space Research*, *70*(10), 2878-2893.