Modeling ionograms with deep neural networks: Training small datasets

J. Aricoché[1], E. Rojas[2], R.Rojas[4] , M. Milla[3]

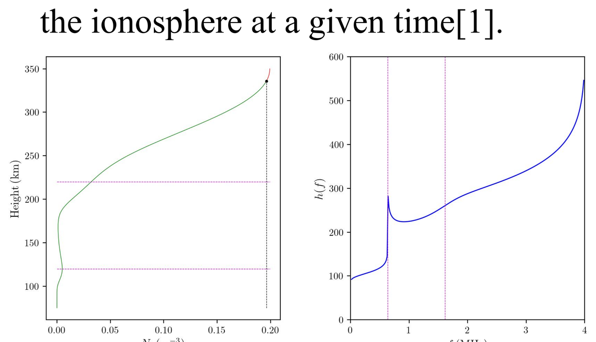
[1,2]Cornell University,New York, USA, [3]Pontificia Universidad Católica del Perú, Lima, Perú, [4]Universidad de Ingeniería y tecnología, Lima, Perú

Abstract

We employed deep neural networks to forecast ionograms across different solar activity periods and database sizes. In order to estimate foF2, we employed two distinct models. The first model identified the last frequency of each ionogram, allowing the neural network to extrapolate virtual heights for all frequencies provided. We carefully optimized the hyperparameters of these models and compared their accuracy against estimates obtained from the IRI and Sami2 models. Additionally, we explored the temporal variability of our predictions by training on consecutive datasets and observing how the results evolved over time. In this study, we will present our findings for three different analysis to train the models: the solstices and equinoxes analysis, the climatological and the rolling window average analysis.

Scientific Problem

- Initially, we developed this work as part of our main research project, which aims to estimate electron densities while forecasting ionograms. Ionograms are states of representation of the ionosphere at a given time[1].



$$h(f) = \int_0^{z_r} \frac{dz}{n(z)} = \int_0^{z_r} \frac{dz}{\sqrt{1 - \frac{fp(z)^2}{f^2}}}$$

Figure 1. Electron density profile and the corresponding ionogram

- We propose to train DNNs with Jicamarca Radio Observatory's digisonde data to reconstruct this unknown function sss, which would give us ionogram forecasting capabilities. To evaluate the performance of our models, we conducted training using several datasets.

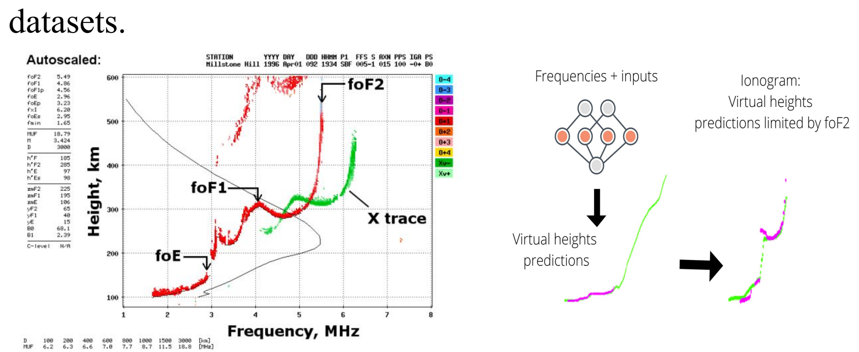


Figure 2. Ionograms predictions limited by foF2.

- Given that regression DNN will estimate a virtual height for every possible frequency, a separate estimate of fof2 has to be provided. Several approaches have been used to estimate foF2 by training neural networks with foF2 and geophysical data and, as presented in [2].

Datasets

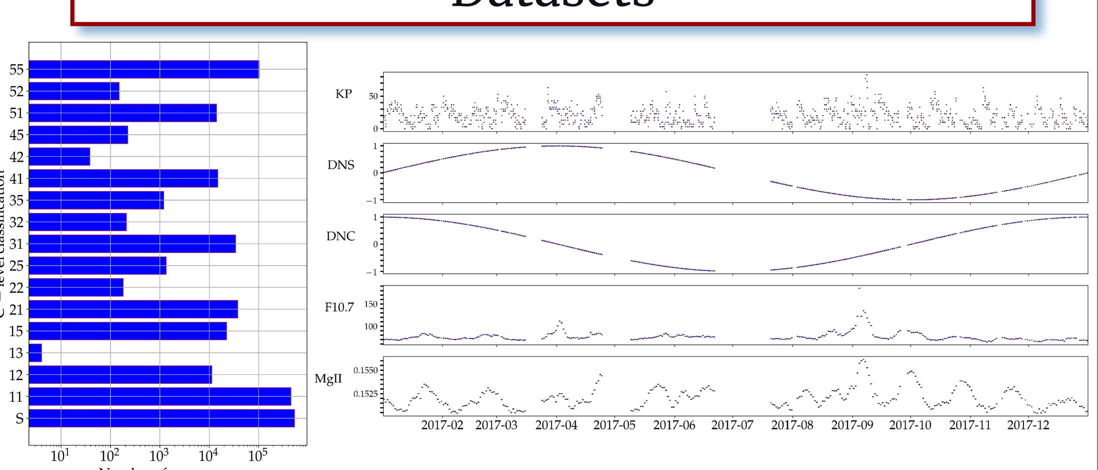


Figure 3. C flag classification(left) and Input parameters time series for some dates.

- Day of year values were converted into 2 sines and cosines to avoid discontinuities as proposed in [4] for solstices and equinoxes model and for climatological model.

- We filtered the Digisonde ionograms used to train the model using ARTIST c-level flag. The c-level flag indicates and qualifies some ARTIST scaled results[3]. 11 means high quality and 55 low quality.

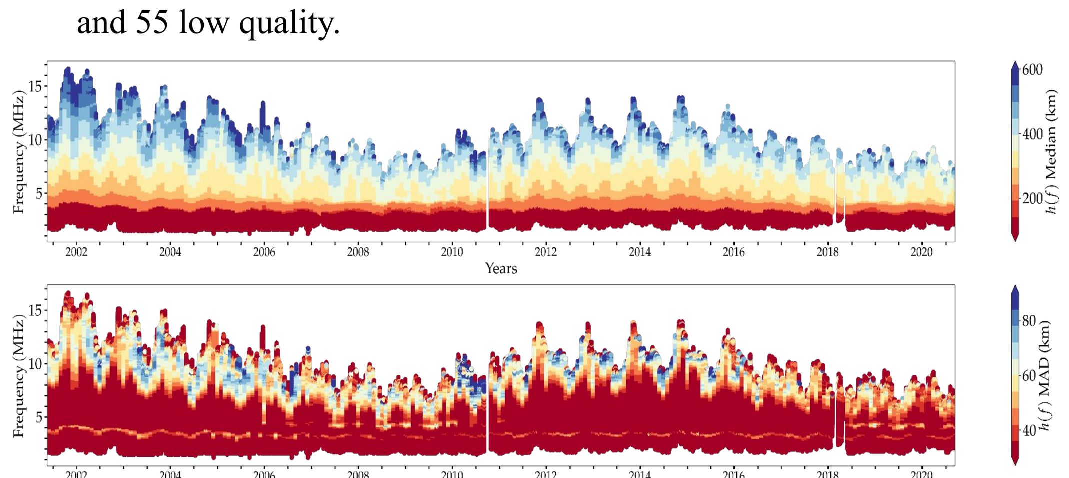


Figure 4. Ionograms medians and standard deviation per month from 8am - 11am ionograms.

Models and Hyperparameter Tuning

- Two supervised models are presented. Both models use a regression neural network for virtual heights forecasting. Model one uses a regression NN with foF2 data and Model 2 is a binary classification one that unliked other machine learning methods or approaches includes virtual heights and frequencies that are not foF2 in the training data.

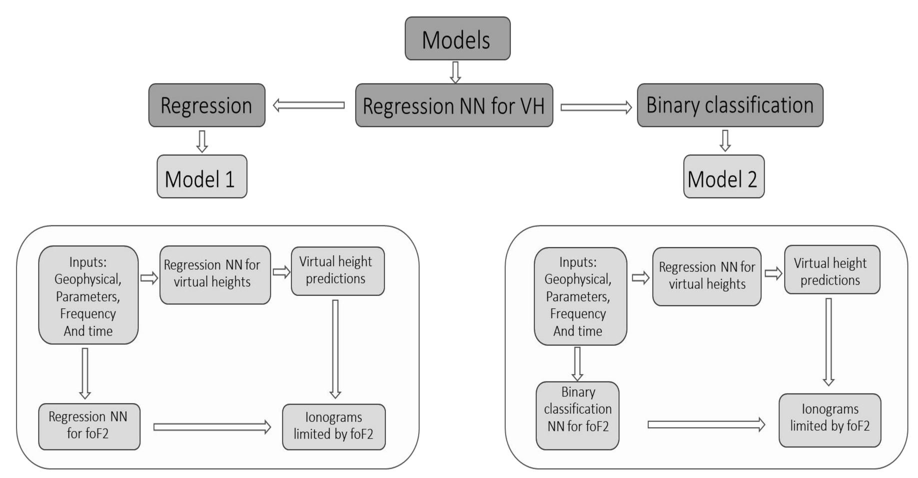


Figure 5. The two proposed models.

- The learning rate and number of nodes by layers were chosen with OPTUNA(an open-source hyperparameter optimization framework[5]).
- Relu, sigmoid and swish activation functions were used.

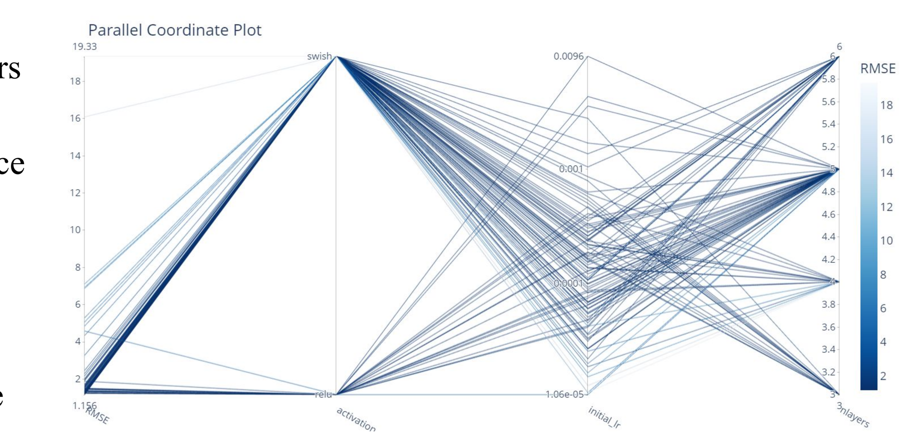


Figure 6. Optuna parallel plot and some results obtained with OPTUNA.

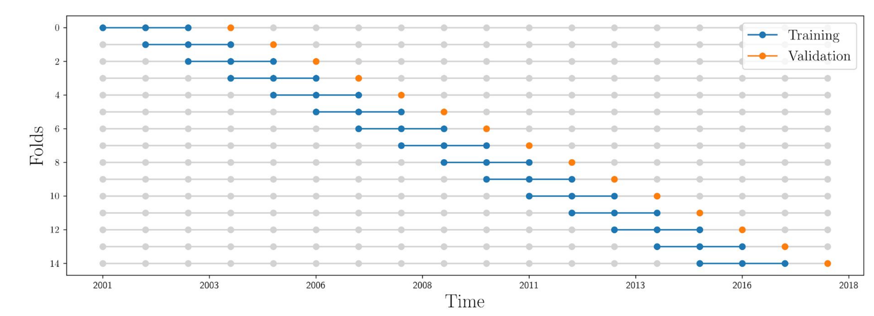


Figure 7. Splitting window technique used in climatological approach

- Different hyper-parameters were obtained for each dataset.
- Validation set was given to OPTUNA to find the best hyper-parameters.

Results and Comparisons

Solstices and Equinoxes Analysis

- Four activity seasons were evaluated for testing the model: a solstice of a solar minimum, one equinox of solar minimum, one solstice of solar maximum, and one equinox of solar maximum. We present the most representative results and comparisons for the solstice of the solar minimum.

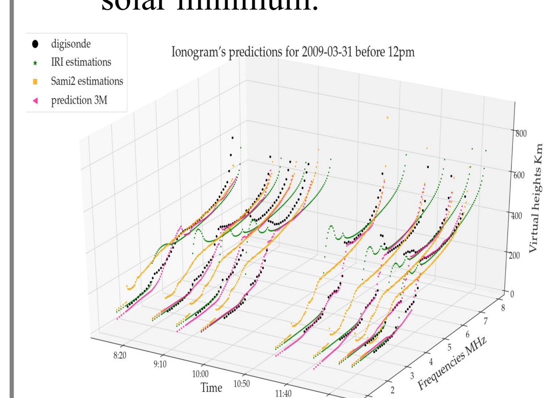


Figure 8. Comparisons of morning ionograms predictions using one month and three months of data to train the model.

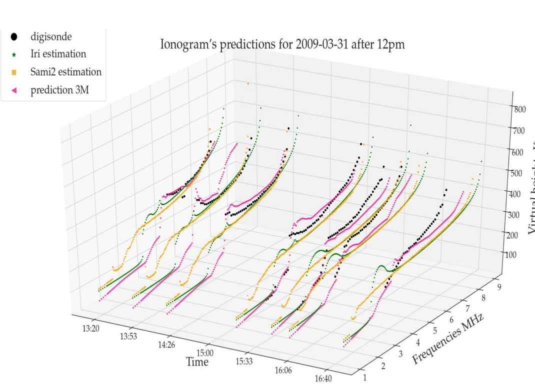


Figure 9. Comparisons of afternoon ionograms predictions to digisonde values, IRI and SAMI2 predictions.

Figure 10. Metrics table to compare performance of the evaluated models to forecast ionograms to IRI and SAMI2 using model 2(uses binary classification NN for foF2) with 1 and 3 months of data.

Evaluation of neural network models to estimate ionograms (RMSE Km)				
Metrics	Model 2 (I month of data)	Model 2 (3 months of data)	IRI estimations	SAMI 2 estimations
SOLSTICE OF SOLAR MINIMUM (DECEMBER 2009)	43.47	51.69	87.23	81.15
EQUINOX OF A SOLAR MINIMUM (MARCH 2009)	25.64	30.37	82.86	70.07
SOLSTICE OF A SOLAR MAXIMUM(June 2014)	**53.04**	40.2	**54.45**	91.68
EQUINOX OF A SOLAR MAXIMUM (MARCH 2013)	32.46	31.15	67.0	49.23

Figure 11. Metrics table to compare performance of the evaluated models to forecast foF2 to IRI and SAMI2 using model 2(uses binary classification NN for foF2).

Evaluation of neural network models to estimate foF2 (RMSE Mhz)				
Metrics	Model 2 (I month of data)	Model 2 (3 months of data)	IRI estimations	SAMI 2 estimations
SOLSTICE OF SOLAR MINIMUM (DECEMBER 2009)	0.44	0.47	1.12	0.59
EQUINOX OF A SOLAR MINIMUM (MARCH 2009)	0.58	0.51	1	0.75
SOLSTICE OF A SOLAR MAXIMUM(June 2014)	**0.62**	0.82	**0.67**	1.47
EQUINOX OF A SOLAR MAXIMUM (MARCH 2013)	1.81	1.53	1.25	0.70

Climatological Analysis

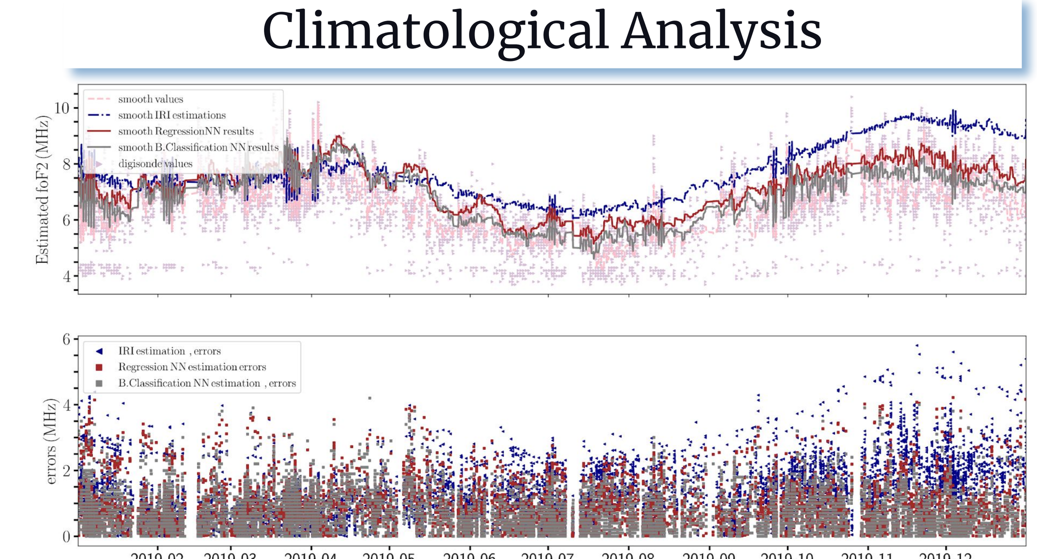


Figure 12. foF2 comparisons of the results for the year 2019

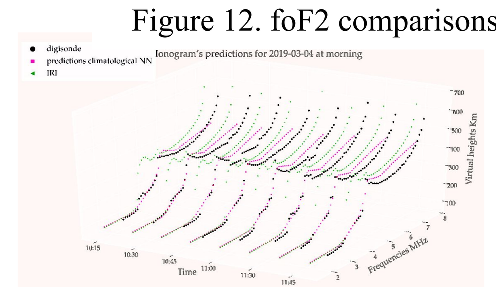


Figure 13. Morning ionogram comparisons

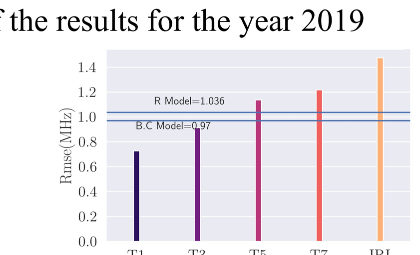


Figure 14. Metrics and comparisons between both foF2 models and persistence.

Rolling Window-Average Analysis

- We selected three consecutive training sets from the year 2019 for analysis. The first study used the months of January to March, the second study covered April to June, and the third study focused on July to September. For each study, we assessed the model's performance using data from the following three months immediately after the training period. In this report,We present the most representative results and comparisons for these studies:

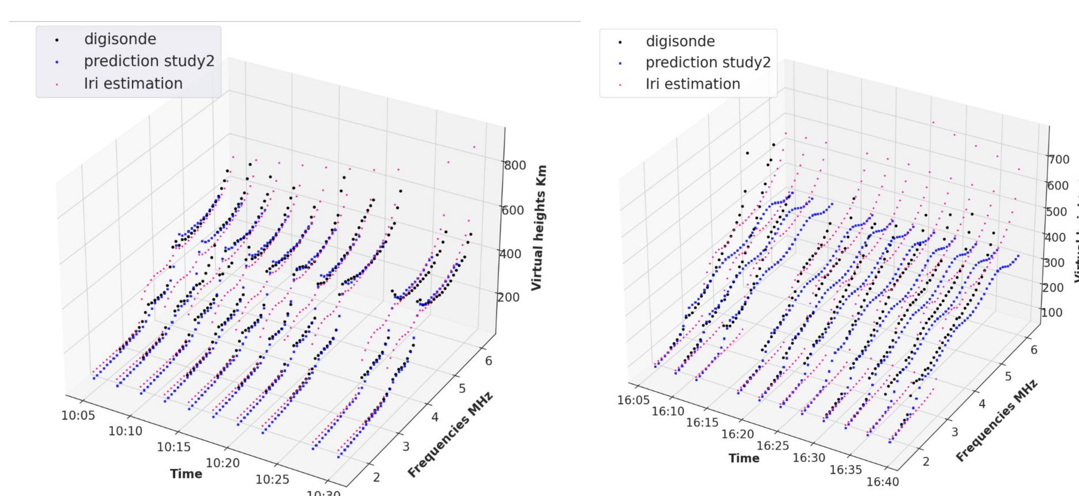


Figure 15. Ionogram predictions for the second study and their comparison to the digisonde values and IRI estimations.

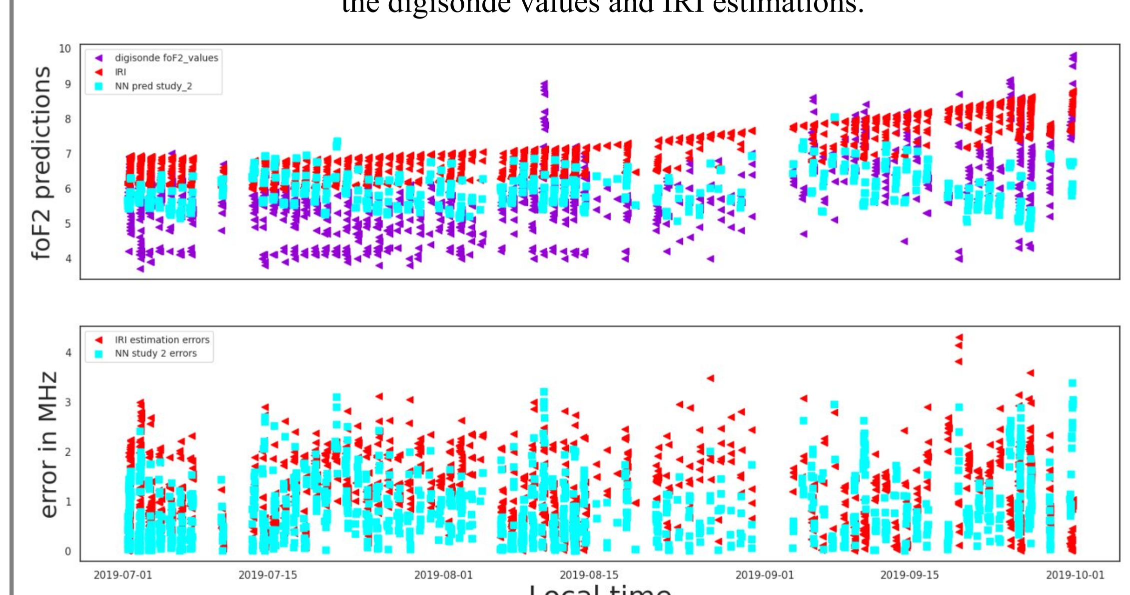


Figure 16. foF2 predictions and errors of our models and their comparisons to digisonde values and IRI.

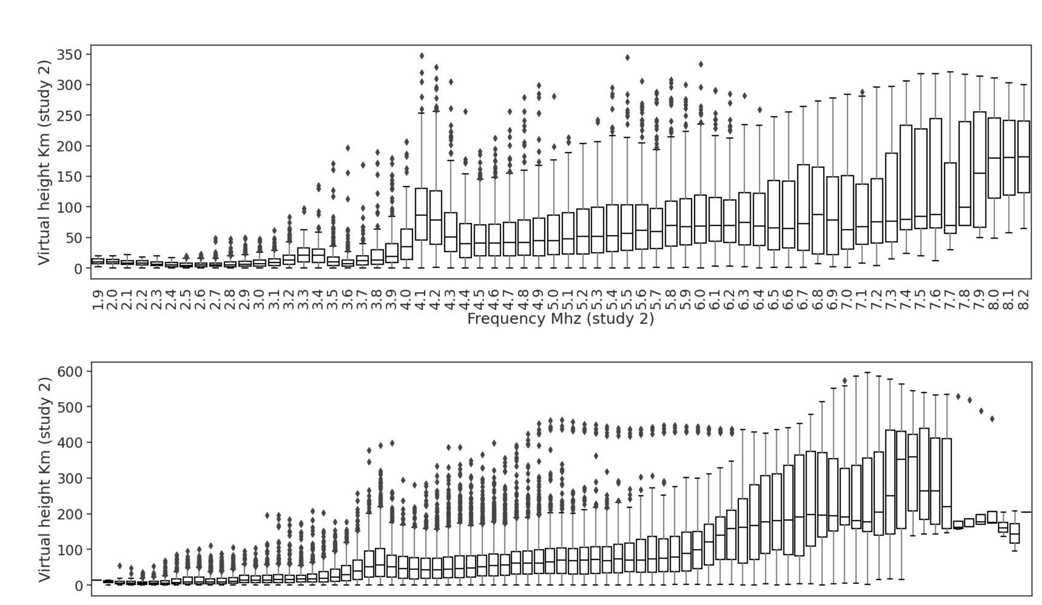


Figure 17. Errors and outliers for morning and afternoon ionograms(study 2 evaluated in the next 3 months of data).

Results and Comparisons

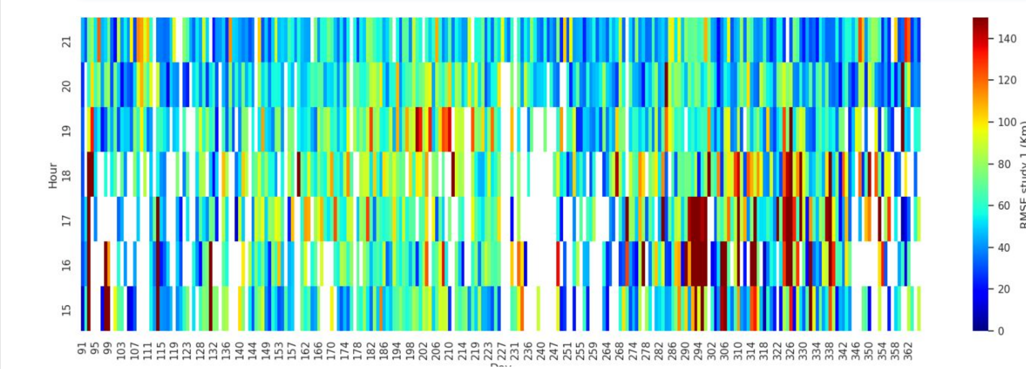


Figure 18. Heatmap to observe how the RMSE evolves over time for ionograms predictions using study 1.

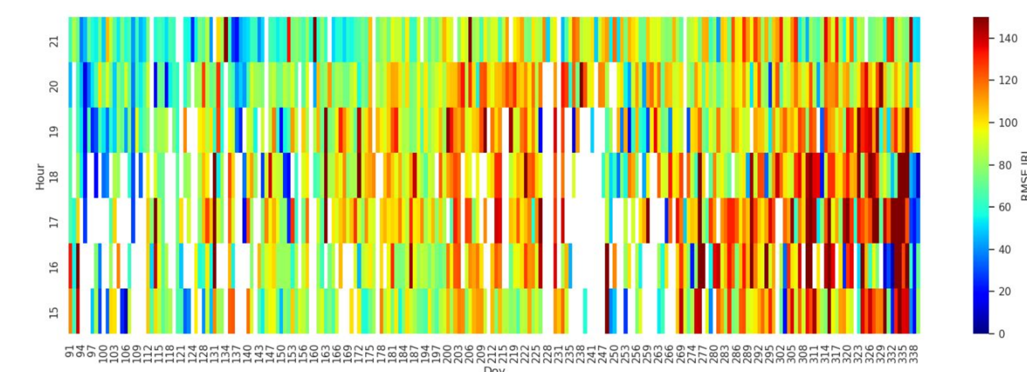


Figure 19. Heatmap to observe how the RMSE evolves over time for ionograms predictions using study 1.

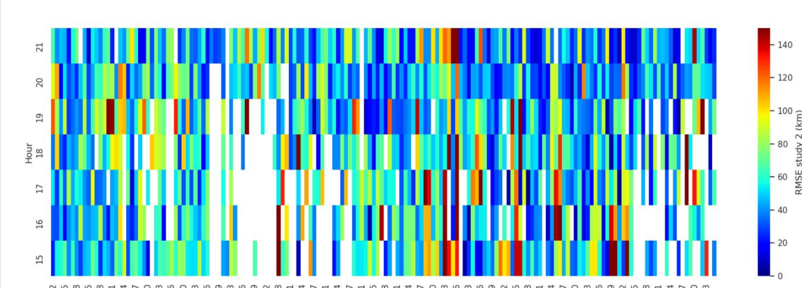


Figure 20. Heatmap to observe how the RMSE evolves over time for ionograms predictions using study 2.

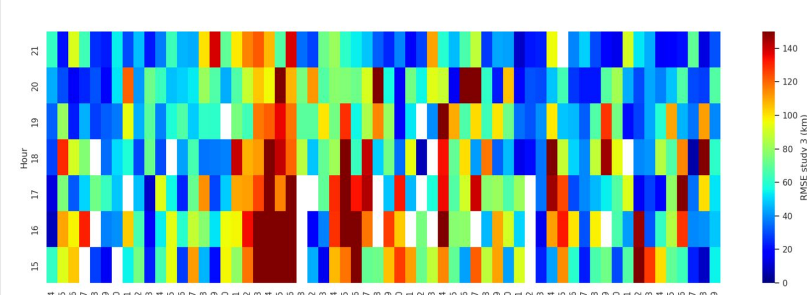


Figure 21. Heatmap to observe how the RMSE evolves over time for ionograms predictions using study 3.

Conclusions and Future Work

Our preliminary results show:

- Neural networks models trained with 3 months of data can capture geophysical parameters and virtual heights variations to show virtual heights results better than IRI and SAMI2 estimations according to solstices and equinoxes analysis.
- foF2 neural network model trained with ionograms is slightly better than using a regression neural network for foF2 during the solar minimum and for the climatological model analysis.
- Morning estimated ionograms seem to be better than afternoon ionograms.
- A short training with three months of data seems to be enough to outperform IRI ionograms for more than 3 months of predictions.
- After making tests on small datasets, we can observe through the good estimations that using deep learning or Machine learning approaches with non-complex models can have potential applications to make ionosonde parameters forecasting using ionosondes with few data or recently installed ionosondes.
- Future work will be oriented toward electron densities forecasting.

References

1. De la Jara, César. 2019. Ionospheric Echoes Detection in Digital Ionograms Using Convolutional Neural Networks. Lima : s.n., 2019.
2. Sai 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
3. https://www.digisonde.com/pdf/Digisonde4DManual_LDI-web.pdf
4. Poole, A. W. V., & Mckinnell, L. A. (2000). On the predictability of foF2 using neural networks. Radio Science, 35(1), 225–234. https://doi.org/10.1029/1999RS900105.
5. Optuna: A Next-generation Hyperparameter Optimization Framework. Takuya, Akiba, y otros. 2019. 2019. Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. DOI: 10.1145/3292500.3330701.