Automatic Segmentation of Range-Time-Intensity maps of Equatorial Spread-F

Susan Palacios and Marco Milla

Pontificia Universidad Católica Del Perú

e-mail: spalacios@pucp.edu.pe

2023

Abstract

The Equatorial Spread-F (ESF) is a nighttime ionospheric phenomenon that can disturb the radio signals of global navigation satellite systems (GNSS) or communication systems in the equatorial zone. This phenomenon is related to plasma density irregularities (bubbles) generated at F-region heights in the ionosphere. In Peru, ESF studies have been conducted for many years using the Jicamarca ionospheric radar operating in the JULIA mode. The radar measures the ESF backscattered power registered in Range-Time-Intensity (RTI) maps. These RTI maps show the temporal and spatial (height) occurrence of ESF, allowing us to observe different morphological patterns (bottom-type, bottom-side, radar plumes, and others). In this work, our goal is to automatically segment and classify the ESF patterns in the RTI maps using machine learning and deep learning algorithms. Leveraging the data available in the scientific database Madrigal, different techniques such as Random Forest (RF), eXtreme Gradient Boosting (XGBoost), Neural Networks(NN), and UNET Convolutional Neural Networks are being tested. A comparison study between the techniques reveals the potential of UNET algorithm to segment and classify the ESF patterns. The features used in the segmentation of the RTI maps include geospace physical parameters and SNR texture features that provide spatial information to the learning algorithms.

Data and Methodology

The data set used in this work consists of 221 RTI maps that correspond to the years 2019 and 2020. The methodology involves two stages: building training data, and training a model using RF, XGBoost, NN and UNET methods. Figure 1 illustrates a RTI map.

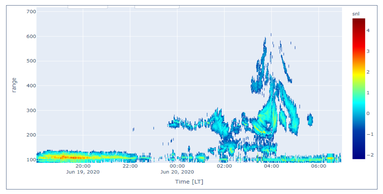


Figure 1: RTI map of echoes measured by JULIA on 19 June 2020.

Building training data

In this stage, we prepared the data sets to train and test the model. First, RTI maps were cleaned to avoid information unrelated to the ESF; therefore, minor artifacts were removed. Second, we manually identified and labeled four patterns (bottom-type, bottom-side, radar plumes, and E-echoes) in each RTI map (Figure 2). These classifications follow the descriptions in [1], [2], and [3]. Third, following the recipe in [4], we decreased the resolution of the RTI maps, where each power value corresponds to an average of a matrix made by a resolution of 15 min and 15 km in height (Figure 3). Then, statistical properties of the backscatter power that describe the texture content of a region were calculated. In total, six statistical properties were determined: mean, standard deviation, smoothness, skewness of the histogram, uniformity, and entropy. In summary, we have considered the following 14 features: SNR, six texture features, range, time, vertical drift, zonal drift, F10.7 solar flux index, Kp geomagnetic activity index, and lunar phase. All these features were merged in a stack of 14 bands (Figure 4).

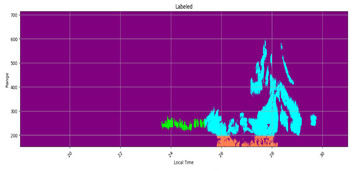


Figure 2: Manually label

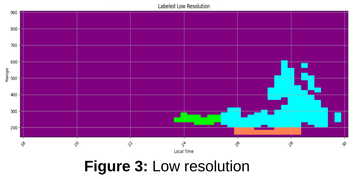


Figure 3: Low resolution

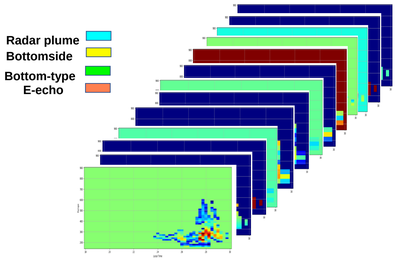


Figure 4: Stack of SNR, statistical textures, range, time, vertical drift, zonal drift, F10.7, Kp, and lunar phase.

Training model

This stage involves training the models with the following methods: RF, XGBoost, NN, and UNET. We used Scikit-Learn and TensorFlow libraries. The 221 RTI maps were divided into 90% of RTI (198) for training and 10% of RTI (23) for testing. The total number of pixels per class are detailed in Table 1. The *K-fold cross-validation* feature was used, and we split the training set into five random subsets named folds. Then, the classification algorithm was trained and evaluated five times, picking a different fold for evaluation every time and training on the other four folds (Figure 5). Grid Search and Randomized Search were used to Fine-Tune the hyperparameters. Table 2 shows the setting parameters for each model.

Table 1: Total number of pixels per class.

Classes	N. Of Pixels
Class 1: Radar Plume	15571
Class 2: Bottomside	9791
Class 3: Bottom-type	3244
Class 4: E-Echos	838

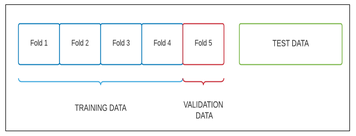


Figure 5: Train, Validation and Test data.

Table 2: Model configurations.

RF	XGBoost	NN	UNET
• Number of trees : 243 • Measure of the split quality: GINI • Max depth of the tree: 14 • Max features for best split: 3.	• Max depth of a tree: 7 • Learning rate: 0.1 • Number of boosting rounds: 100	• Learning rate: 0.0007 • Number of hidden layers: 3 • Number of neurons: 4 • Optimizer: Adam. • Activation: RELU. • Loss function: Categorical Cross entropy. • Number of epoch: 12	• Learning rate: 0.001 • Optimizer: Adam. • Activation: RELU. • Loss function: Categorical Cross entropy. • Number of epoch: 28 • Batch size:16

Discussions and Conclusions

- Four machine learning and deep learning algorithms were tested for ESF segmentation and classification. UNET (83.99%) reached better results than Random Forest (62.19%), XGBoost (52.21%) and Neural Network (67.15%) tested on 23 RTI.
- Even though the algorithms were trained with more radar plume data, the algorithms tend to predict the bottomside instead of the radar plume. On the other hand, the algorithms predict bottomside class instead of bottom-type due to the low quantity of bottom-type training data.
- In this work, we considered features related to the plasma drift velocity, the solar flux, and geomagnetic activity, as well as, texture features based on the region description of the backscatter power to provide context information on the morphological patterns.
- The feature importance using Random Forest shows the range as the most important feature during classification, followed by the time, F10.7 solar flux index, smoothness, and the other features.
- Future work will focus on adding more training data to balance the number of pixels per class.

References

[1] Hysell, D. L., & Burcham, J. D. (1998). JULIA radar studies of equatorial spread F. Journal of Geophysical Research, 103(A12), 29,155–29,167. https://doi.org/10.1029/98JA02655.

[2] Hysell, D. L. (2000). An overview and synthesis of plasma irregularities in equatorial spread F. Journal of Atmospheric and Solar-Terrestrial Physics, 62 (12), 1037-1056. https://doi.org/10.1016/S1364-6826(00)00095-X

[3] Hysell, D. L., & Burcham, J. D. (2002). Long term studies of equatorial spread F using the JULIA radar at Jicamarca. Journal of Atmospheric and Solar-Terrestrial Physics, 64 (12-14), 1531-1543. https://doi.org/10.1016/S1364-6826(02)00091-3

[4] Zhan, W., F. S. Rodrigues, and M. A. Milla. 2018. "On the Genesis of Postmidnight Equatorial Spread F: Results for the American/ Peruvian Sector." Geophysical Research Letters 45(15): 7354-7361, https://doi.org/10.1029/2018GL078822

Acknowledgements

- The authors thank to CEDAR 2023 Science Steering Committee for kindly support the presentation of this work.
- We gratefully thanks the Jicamarca Radio Observatory (JRO). The JRO is a facility of the Instituto Geofísico del Perú operated with support from the NSF Award 1732209 through Cornell University.

Results

The metric evaluation for each model is the accuracy and confusion matrix shown in Figures 6, 7, 8, and 9. The learning models were applied to segment and classify two RTIs (Figure 10 and 11) that were not used for training and testing previously. Comparison between the ground truth label (Figure 13 and 18) and the model predictions are illustrate in figures from 14 to 17 and from 19 to 22.

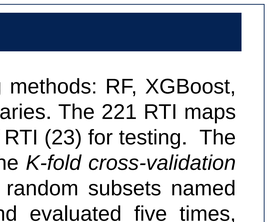


Figure 6: Confusion Matrix and accuracy of RF.

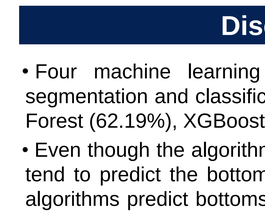


Figure 7: Confusion Matrix and accuracy of XGBoost.

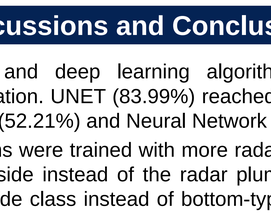


Figure 8: Confusion matrix and accuracy of NN.

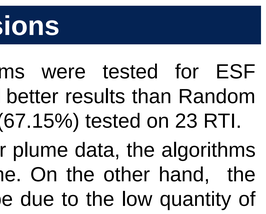


Figure 9: Confusion matrix and accuracy of UNET.

Table 3: Accuracy results for two RTI map measured on 31 January and 31 May 2020..

MODEL	Accuracy % 31 Jan 2020	Accuracy % 31 May 2020
RF	74.17	70.06
XGBoost	88.08	58.68
NN	74.83	62.87
UNET	88.43	90.25

Figure 10: RTI map of echoes measured by JULIA on 31 January 2020.

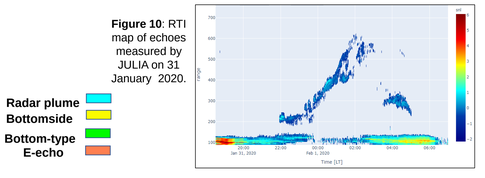


Figure 11: RTI map of echoes measured by JULIA on 31 May 2020.

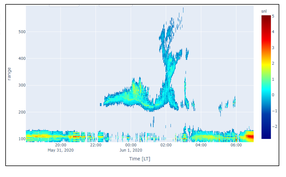
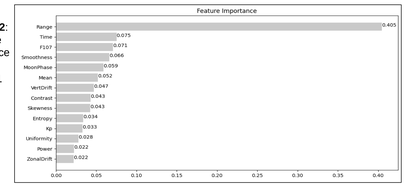


Figure 12: Feature Importance of RF method.



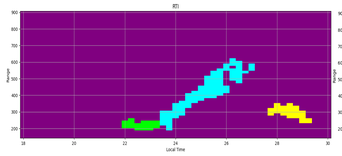


Figure 13: Figure 10 labeled manually.

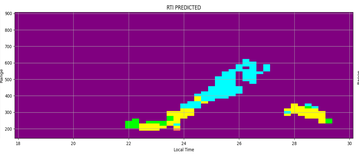


Figure 14: RF prediction of figure 10.

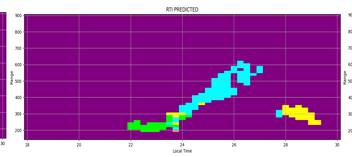


Figure 15: XGBoost prediction of figure10.

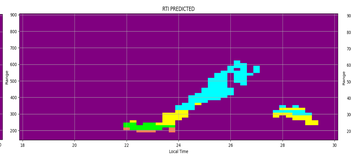


Figure 16: NN prediction of figure 10.

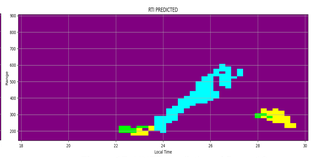


Figure 17: UNET prediction of figure 10.

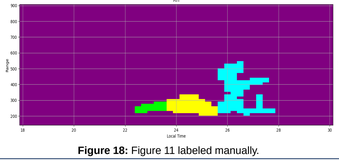


Figure 18: Figure 11 labeled manually.

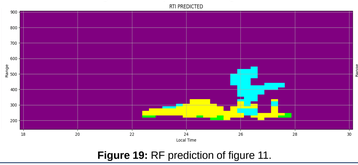


Figure 19: RF prediction of figure 11.

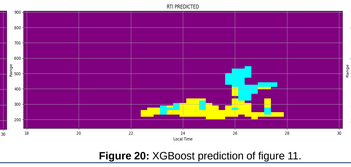


Figure 20: XGBoost prediction of figure 11.

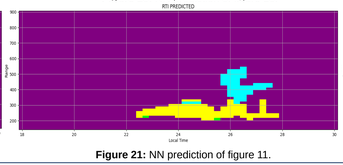


Figure 21: NN prediction of figure 11.

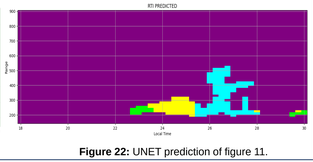


Figure 22: UNET prediction of figure 11.