

# Exploring the correlation between ionospheric scintillation and GNSS positioning error near the magnetic equator using machine learning techniques

G. Fajardo<sup>1</sup>, E. Pacheco<sup>1</sup>

<sup>1</sup> Jicamarca Radio Observatory, Geophysical Institute of Peru

## Abstract

Ionospheric scintillations are a common phenomenon in the equatorial ionosphere. This phenomenon directly affects the position estimated by GNSS receivers degrading the quality of the radio signals; however, the quantification of the positioning error contributed by the ionosphere over the Peruvian sector has not been studied in detail. In this work, algorithms are being implemented that will allow us to identify and classify amplitude scintillation (S4) levels, we have worked with data from the Huancayo Observatory for the period December 2016-February 2017 obtained from LISN, this data has been plotted to analyze the spatial and temporal occurrence, and to analyze the occurrence of S4 as a function of other space weather variables obtained from OMNI2. The machine learning algorithms were decision tree, Support Vector Machine (SVM), Neuronal Network (NN). Decision tree was implemented as a filtering method, support vector machine for clustering and neuronal network to generate time series in forecasting. This paper shows the initial part of an investigation that aims to correlate qualitatively and quantitatively the occurrence of amplitude scintillations (S4) with errors in the position estimation of GNSS receivers, once the correlation between S4 and position error has been quantified, it may be possible to predict the error by predicting S4.

## Objectives

Processing of S4 data from Huancayo Observatory in the last solar cycle.  
Use of machine learning methods to classify the level of amplitude scintillation (S4).  
Use of machine learning as a new method of statistical analysis.

## Methodology

In this paper, we are using scintillation data and position approximations of a GNSS receiver from the Huancayo Observatory for the local summer December 2016 - February 2017. We analyzed the behavior of space weather variables, scintillation and the position estimation for the same periods of time. The decision tree method has been used as a new method to establish limits for data cleaning, so we used the Decision tree method with different penalty criteria and with space weather variables as inputs (features) and as output (label) 4 levels of scintillation. In this way, we can visually observe which variables are affecting the decision and make a more timely decision in a large amount of data. This is useful because each station has different parameters, references differ establishing thresholds to filter data, which can make them discard information that can be exploited. Spatial projections were used to determine whether there were scintillation patterns in latitude, longitude, elevation or azimuth, that the first method (Decision tree) could not detect, and also to visually observe the dispersion of the position estimates and their actual effect. The Support Vector Machine method was tested as a clustering method for scintillation. Four previous levels were established to classify the intensity of scintillation and these levels of scintillation were set as the label (output of the algorithm).

## Preliminary Results

GNSS position estimates are obtained at Huancayo and we observe the largest errors between 1800 and 2300 local time, as can be seen in Figure 1. The distribution of the S4 index as a function of azimuth and elevation is examined, as it is illustrated in Figure 2. The scintillation level and occurrence have been compared with the position estimations, as observed in Figures 3 and 4. The Machine learning algorithm, decision tree (see Figure 5), has been used as a filter method, each iteration represents a separation of the data, this way given a threshold value, we can observe the amount of data that we are losing, if in some iteration it is observed that a variable separates the data in a considerable percentage for high values of S4, it could suggest a more detailed analysis, in the present work it is observed that for the two first iterations the variables that are taken into account are the elevation, establishing thresholds of 22°, 17° and 31°. The Support Vector Machine method with the data analyzed so far has shown not to be so effective for an intensity classification, the confusion matrix (Figure 6) shows high precision for the low and high level, however the null and medium level does not show a good precision, this suggests not using this classification and joining the null and low level in a label called absence and the medium and high level in a label called presence, this way the algorithm would have a higher precision with this binary problem (presence or absence). The use of neural networks for generating time series (Figure 7) in a short period of time is also being analyzed, so far the results have been good since we obtained 0.153 from Mean Absolute Error (MAE), but it is expected to be compared with the integrated methods at the end of the work.

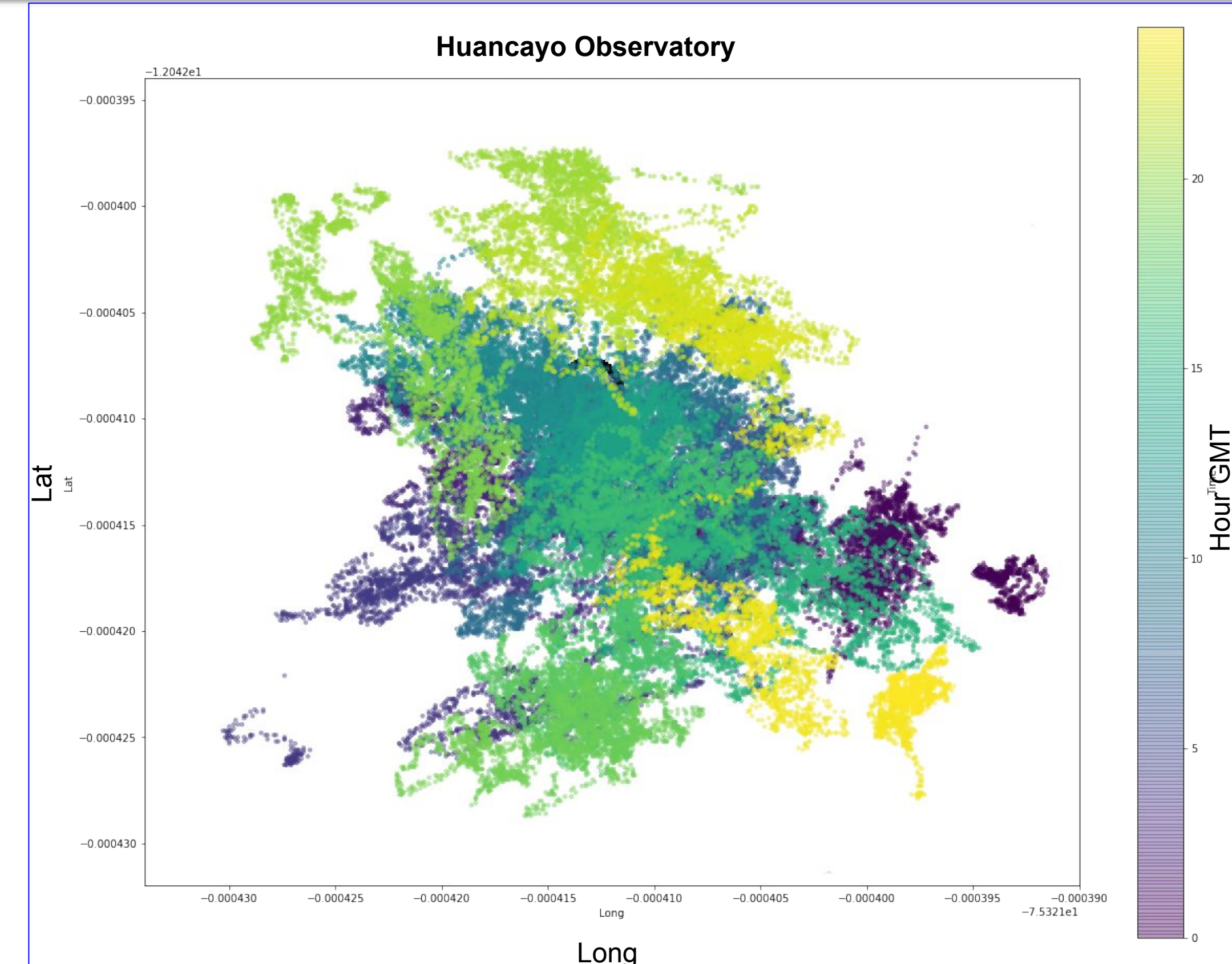


Figure 1. Estimated positions (PSN archive) obtained by GNSS receivers from the Huancayo Observatory, 355 Day of the Year (DOY), year 2016 .

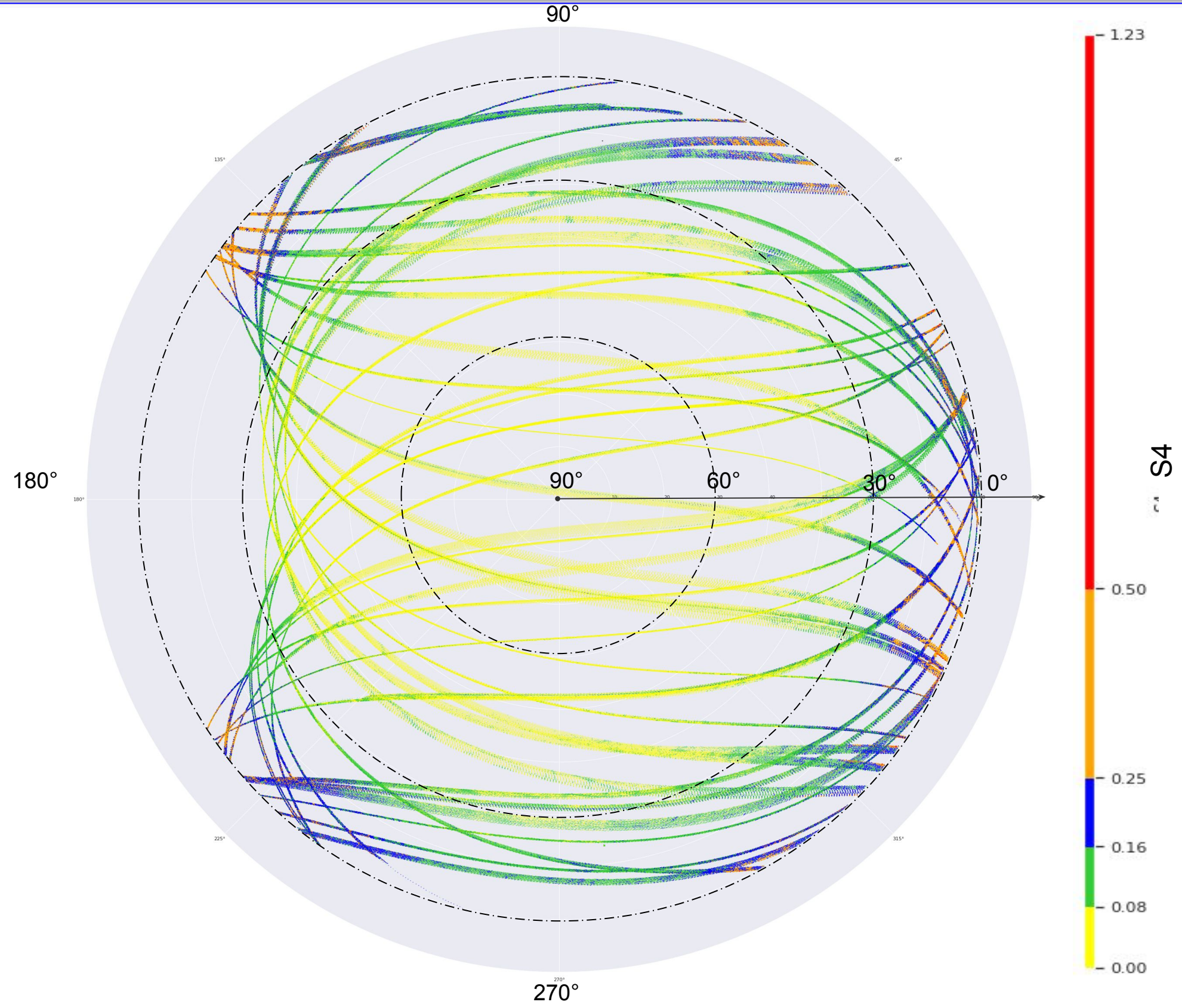


Figure 2. Amplitude scintillation (S4) projection based on Azimuth and Elevation for the local summer December 2016 - February 2017.

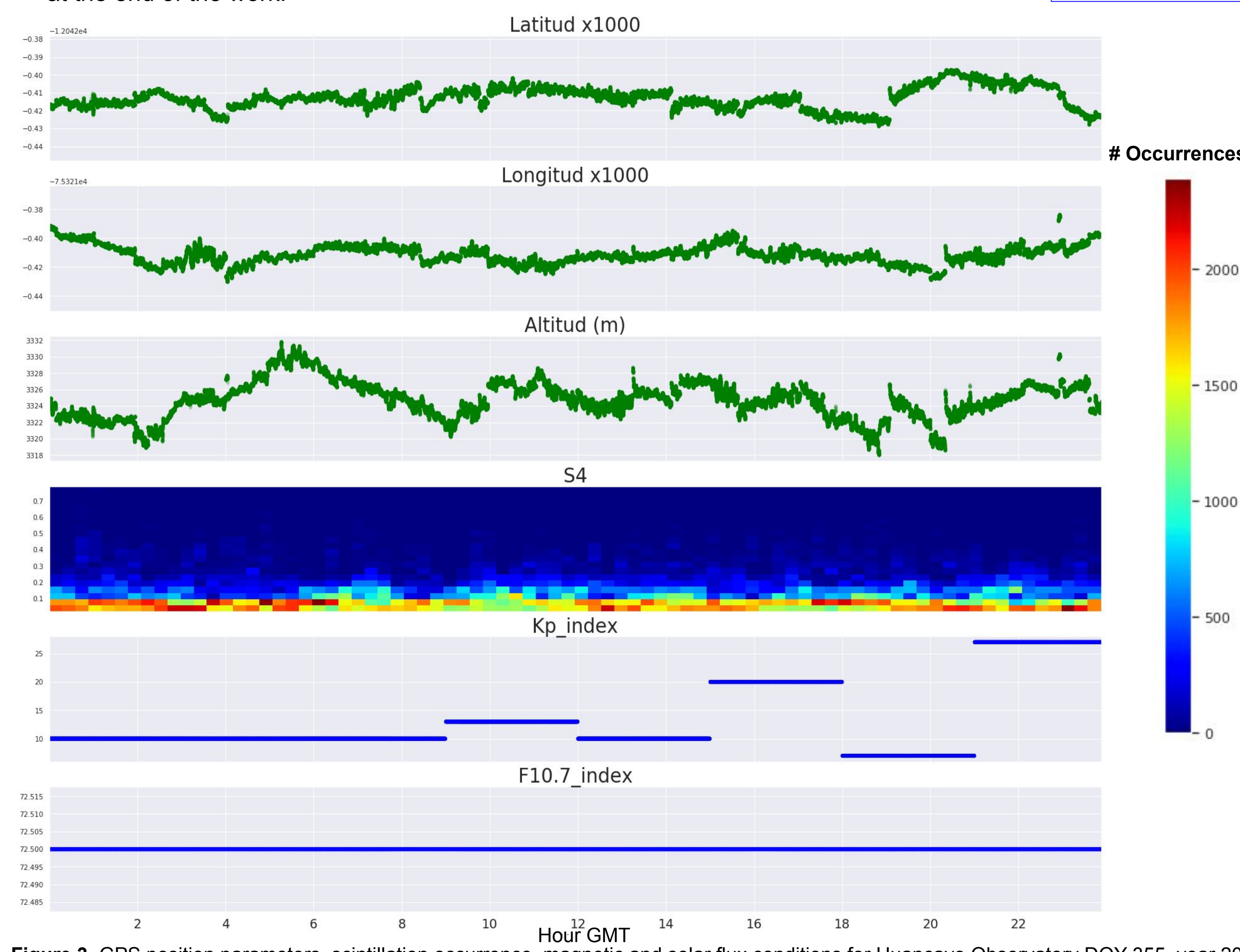


Figure 3. GPS position parameters, scintillation occurrence, magnetic and solar flux conditions for Huancayo Observatory DOY 355, year 2016.

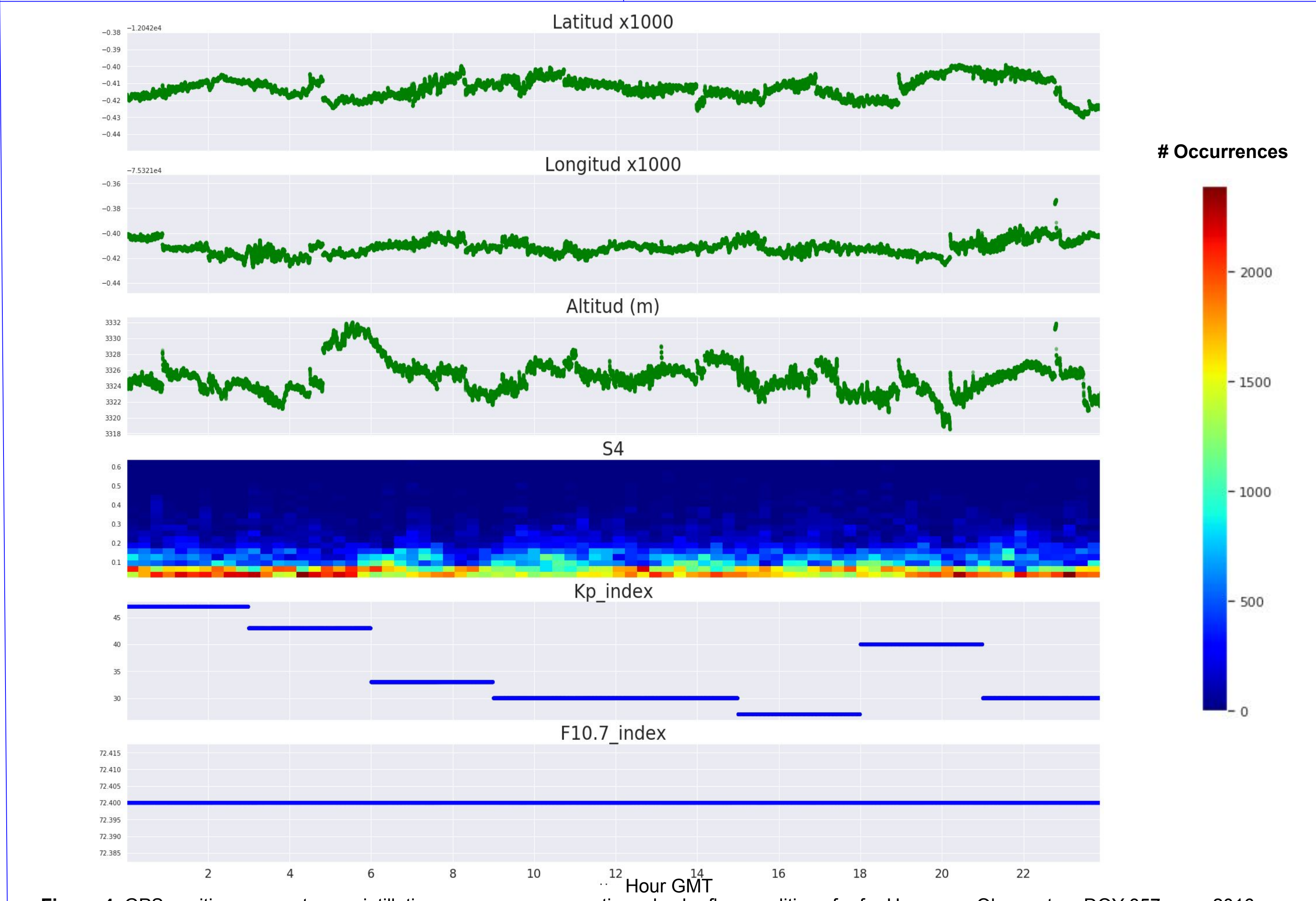


Figure 4. GPS position parameters, scintillation occurrence, magnetic and solar flux conditions for Huancayo Observatory DOY 357, year 2016.

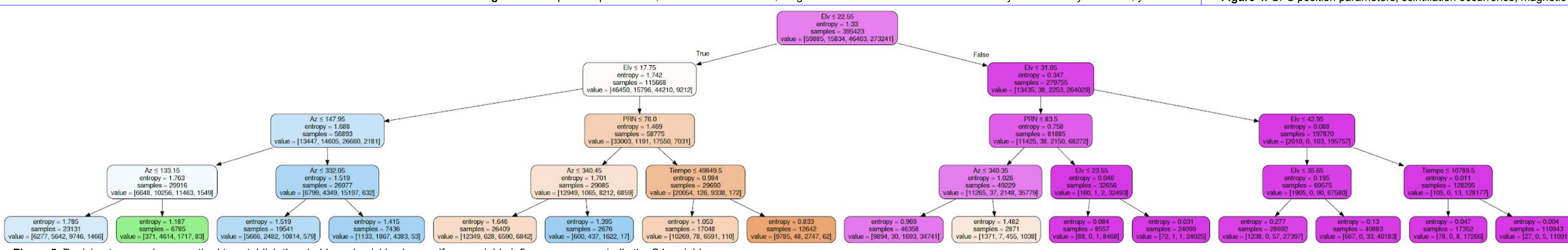


Figure 5. Decision tree used as a method to establish thresholds and quickly observe if any variable influences unexpectedly the S4 variable.

Normalized confusion matrix

	Nulo	Bajo	medio	alto	
Nulo	0.45	0.29	0	0.26	0.8
Bajo	0.03	0.96	0	0.0084	
medio	0.1	0.0009	0.0098	0.89	
alto	0.3	0.044	0.0015	0.66	
		Bajo	medio	alto	0.0

Figure 6. Normalized Confusion Matrix obtained after training the data from Summer 2016-2017 with Support Vector Machine (SVM) with 4 labels for S4.

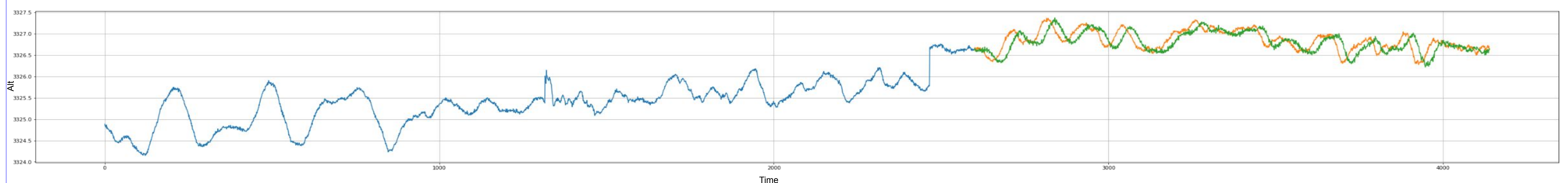


Figure 7. Results of a time series obtained using a neural network. In blue training data, in orange test data, in green predicted data.

## Future Work

- Currently in the work we are analyzing the correlation between S4 and Positioning Error, once this is done we can join the binary prediction of SVM to extrapolate an error given the values of space weather variables.
- We want to analyze multivariate methods such as Long-Short Term Memory (LSTM) for error forecasting and compare it with the method described above.

## Conclusions

- The development of Machine and Deep learning methods, the improvement in the computational capacity and the great amount of accumulated data, especially from the last solar cycles, invite us to analyze these data with techniques not used before. This work presents tests for algorithms, obtaining favorable results, used in a new dataset obtained from LISN and OMNI2 data, which was created and will be shared for use. In addition, using these techniques we are exploring the correlation and quantification of ionospheric error in the approximations of position of GNSS receivers, which at the time of the study, still does not present a determinant correlation, however the data is still under analysis.

## References

- N. Linty, A. Farasin, A. Favenza, and F. Dovis. Detection of gnss ionospheric scintillations based on machine learning decision tree. IEEE Transactions on Aerospace and Electronic Systems, 55(1):303-317, Feb 2019.
- Ryan M. McGranaghan, Anthony J. Mannucci, Brian Wilson, Chris A Mattmann, and Richard Chadwick. New capabilities for prediction of high-latitude ionospheric scintillation: Anovel approach with machine learning. Space Weather, 16(11):1817-1846, 2018.
- C E Valladares and J L Chau. The Low-Latitude Ionosphere Sensor Network: Initial results. Radio Science, 47(4), 2012.

## Acknowledgments

We gratefully acknowledge the thesis program of the Geophysical Institute of Peru and the Jicamarca Radio Observatory staff for their support. We also thank Cesar Valladares, Cesar De La Jara for the LISN data.

## For further information

Contact: [george.fajardo@iro.igp.gob.pe](mailto:george.fajardo@iro.igp.gob.pe)