

# Determining a Local Quiet Day with a Machine Learning Model Approach

Castellanos-Velazco C. I. & Castro A.

UNAM - México

Email: [ccastellanos@igeofisica.unam.mx](mailto:ccastellanos@igeofisica.unam.mx)

## Abstract

Quiet Days (QDs) play a crucial role in modeling diurnal variations and removing their contribution to compute regional geomagnetic indices. This project aims to develop a local QD identification process by adapting van de Kamp's criteria for selecting Local Quiet Days (LQDs). However, empirical testing has revealed that days classified as LQDs are not always truly quiet, as they may lack the magnetic signatures associated with diurnal variation. To address this limitation, we propose an automatic, unsupervised machine learning method that combines Long Short-Term Memory (LSTM) networks with Convolutional Neural Networks (CNNs). This approach seeks to enhance the detection of truly LQDs within specified time windows, ensuring more accurate identification and analysis.

## Introduction

For regional studies of space weather and geomagnetic storms (GS), magnetometer records from the region of interest are analyzed. These records provide insights into variations in electric currents within the ionosphere and magnetosphere [1]. However, the data also include magnetic contributions from processes unrelated to GS [2]. Therefore, it is crucial to preprocess the magnetometer data to accurately identify and isolate the magnetic signatures associated with GS.

The Earth's magnetic field (EMF) exhibits variations across different time scales, commonly referred to as regular variations [3], such as the diurnal variation. This variation is driven by the Solar Quiet (SQ) current ionospheric system [4, 5].

Diurnal variation is conventionally modeled using the average of Quiet Days (QDs), which are defined as days with the lowest geomagnetic activity associated with solar activity. On these days, diurnal variation represents the most prominent magnetic contribution. Therefore, averaging QD data for a month is generally considered an adequate approach for estimating the monthly diurnal variation.

Monthly lists of the ten quietest days are generated using the Kp index as a reference [1]. These days are identified by sorting and averaging daily Kp values, selecting those with the smallest variations. However, relying solely on Kp indices can lead to misleading results. Even days classified as quiet may exhibit significant disturbances,

and geomagnetic quiescence can vary regionally. While one region may experience low magnetic activity, another one may show considerable fluctuations, making the day locally non-quiet. Consequently, QDs determined by the Kp index may not truly reflect quiet conditions on a local scale. To ensure accuracy, quietest days should instead be identified based on local magnetic data.

Another issue is that, even relying on local K indices, their calculation requires the prior identification of QDs, creating a circular dependency, as noted by Gjerloev in [6]. Therefore, an alternative to K indices is needed to identify the most probable QDs using local magnetic data.

## Selection of Local Quiet Days

To derive the diurnal variation, Quiet Days (QD) must be identified first. Once selected, the data from the five quietest days are stacked and averaged to compute the diurnal variation, denoted as  $H_{SQ}$ .

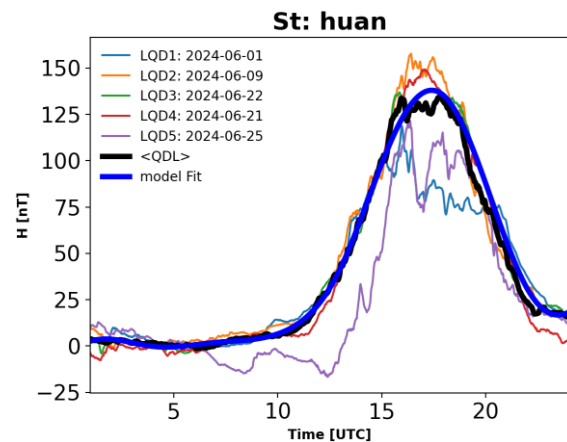


Figure 1: LQD selected for Huancayo magnetic observatory. Black thick curve is the average of the 5 LQD meanwhile blue thick curve is the Fourier fitted model.

As mentioned in the previous section, the conventional Quiet Day (QD) selection method may be unsuitable for regional studies and geomagnetic activity monitoring. To address this, we adapt the criterion proposed by van de Kamp [7] for QD selection using local data. Local Quiet Days (LQD) are identified by determining the maximum daily fluctuation.

Using a monthly time window with a time resolution of 1 minute, the time series is resampled into subsets  $S_i$  of 60 elements (hourly data). The Interquartile Range (IQR) is then calculated as follows:

$$IQR = Q_3(S_i) - Q_1(S_i).$$

Here, the subindex  $i$  represents the  $i$ -th term of the hourly resampled time series. The Interquartile Range (IQR) is used as a measure of the degree of variation  $V$  in the magnetic data. IQR is preferred over standard deviation due to its robustness against extreme values.

Next,  $V$  is resampled into subsets of 24 elements to represent daily variations, and the maximum IQR within each subset is determined as it reflects the peak of daily variation. The days with the lowest  $V$  peaks in each month are designated as the local quietest days (LQD). Using the magnetic field  $H$  data from the LQD, the diurnal variation is then modeled as follows:  $LQD = \frac{1}{N} \sum_{i=1}^5 LQD_i(t) - LQD_H(t_{0:3})$ .

Here, the subindex  $i$  represents the  $i$ -th LQD, which, by convention, ranges from 1 to 5. The resulting average, denoted as  $LQD_{avr}$ , is depicted in Figure 1 as the solid thick black line. As outlined in Equation (2), the offset for each day is normalized by subtracting the average of the night hours (00:00 – 03:00 LT). This step is crucial to mitigate offset variations between days, ensuring they do not distort the resulting average.

As illustrated in Figure 1, we adhered to the criteria established by Chen [8]. In this process, the resulting  $DQL_{avr}$  is resampled to an hourly resolution by selecting the first element of each hour. Subsequently, a Fourier series expansion is fitted to the resampled data:

$$H_{SQ}(UT) = a_0 + \sum_{i=0}^n a_i \cos(2\pi f_n UT + \phi_n),$$

Here,  $a_{0,i}$  represents the fitted coefficients,  $f_n$  denotes the frequency of the  $n$ -th harmonic. In this work, we set  $n=5$ , and  $\phi_n$  is the phase angle of the  $n$ -th harmonic. The fitted model  $H_{FS}$  is then considered as the monthly diurnal variation approach. Figure 1 illustrates the resulting fit (thick blue line) for the magnetic station in Huancayo (HUAN) during June. Notably, the fitted  $H_{FS}$  demonstrates a smoother profile compared to  $DQL_{avr}$ , which, despite the application of a smoothing function, may still exhibit minor fluctuations.

### Method

For training, a time window between 2023 and 2024 was selected, aiming to cover extended periods without significant gaps, using data from the Huancayo (HUAN) magnetometer. The magnetometer data was provided by the Geophysical Institute of Peru (IGP). A moving window of 27 days was applied to the data, which was processed by removing baseline and offset, and then the diurnal variation model was computed following the procedure outlined in the previous section. The moving window is shifted by 10 days, with the process being repeated until the end of the time window. For each result obtained from the moving window, days that deviate from the monthly diurnal baseline magnetic signature,

hereafter referred to as “fake LQDs”, are manually removed.

Data reported by the HUAN magnetometer is sampled every minute, resulting in 1440 data points per day. To prevent minor gaps or temporal errors while preserving the features of daily variation, we resampled the data by computing the median every 30 minutes. This process yields a dataset with 48 data points per day, as shown in Figure 2.

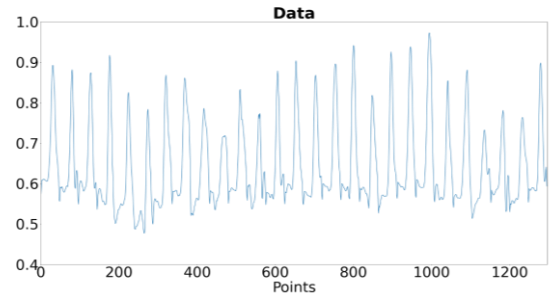


Figure 2: Dataset scaled for training.

### Customize model

The model was customized with CNN, LSTM, and Dense layers. CNNs are effective for extracting features and identifying relationships between days, enabling the generalization of behavior and recovery patterns for estimating quiet days during magnetic storms. As observed with the statistical method, some recovery days of the horizontal magnetic field have been labeled as quiet, even though they do not represent true quiet conditions.

In each layer, we combine regularizer layers as Dropout and Batch Normalization. It is necessary to prevent overfitting.

Layer	Output Shape
Input	(None, 27,48)
Conv1D + BN + Dropout + Relu	(None, 27,64)
Conv1D + BN + Dropout + Relu	(None, 27,128)
LSTM	(None, 27,128)
Flatten	(None, 3456)
Dense + Dropout + Relu	(None, 256)
Dense	(None, 1)

Table1: Layers from model

## Results and Discussion

Using June 2024 as an example, days LQD2, LQD3, and LQD4 clearly show a common pattern related to the passage of the Solar Quiet current. However, this is not the case for LQD1 and LQD5, which exhibit significant deviations from the common pattern between 15:00 and 20:00 UT. These deviations may affect the resulting average and the fitted model, as seen in Figure 1. Additionally, the model has some issues. Between 00:00 and 17:00 UT, the fit is generally good, but after 17:00 UT (local noon), the model shows a small deviation. This causes artifacts in the resulting data after the subtraction process.

A manual or semi-manual selection of LQDs could be a better solution, based on both the magnetic pattern and the statistical smoothness of the behavior to determine LQDs. However, this approach can only be applied to specific study cases and is not suitable for long-term analysis or real-time monitoring. Therefore, we proceed by testing a CNN, LSTM, and Dense Layer model to develop an automated process for LQD selection.

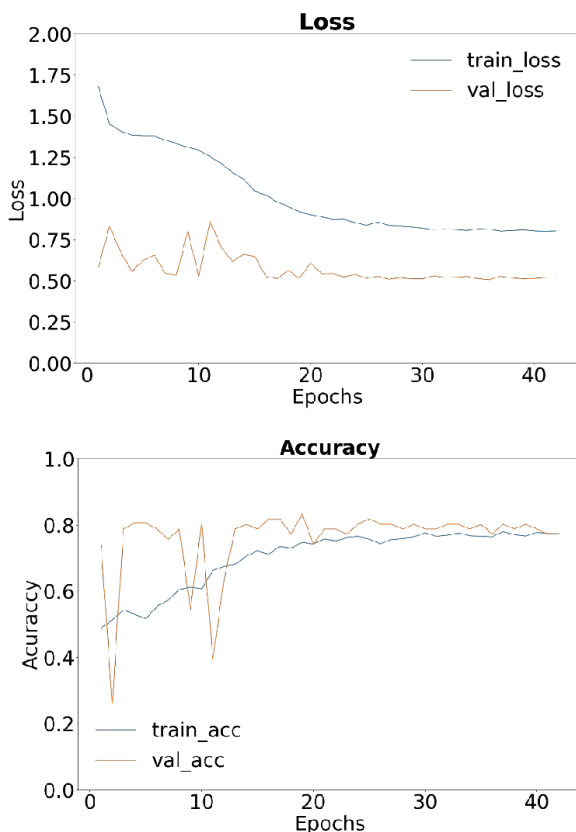


Figure 3: Loss and accuracy during training.

The dataset was divided into 80% for training and 20% for validation. The training process consisted of 75 epochs, achieving an accuracy of 0.81 on the validation dataset, with a loss above 0.5, as shown in Figure 3. In each step, we used a data augmentation technique, changing random values similar to piper and salt noise. Another method consists in adding a low Gaussian

noise. These are preliminary results, and we aim to improve the process either by refining the model construction or by adopting an alternative training approach [7].

The validation dataset consisted of 20 series in total. The confusion matrix shows that 3 out of 10 non-quiet days were misclassified, while 9 quiet days were correctly identified as LQD. This demonstrates that the current model performs better in detecting actual LQDs rather than misclassifying false LQDs.

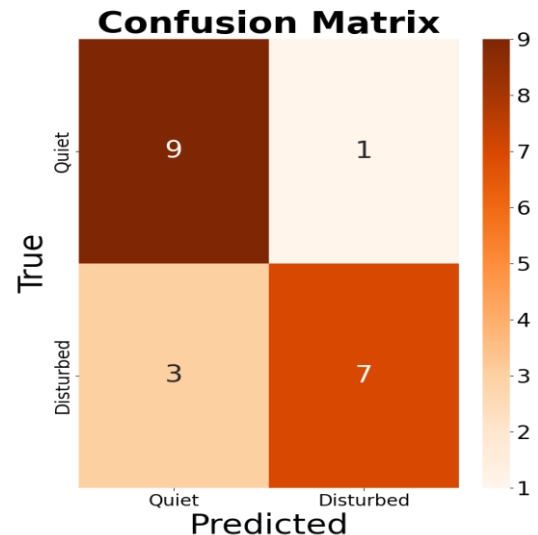


Figure 4: Confusion matrix showing the performance of the ML model used during this work.

## Conclusions

During this project, we employed a machine learning technique to model the diurnal variation without relying on the QD or Kp index. Initially, we adapted the Van de Camp method [6] to identify the local quiet days (LQD) as an alternative approach for modeling the diurnal variation in the Huancayo region, Peru. Subsequently, we applied a Fourier series expansion model to fit the smooth daily variations, functioning as a low-pass band filter.

We acknowledge the existence of statistically quiet days that do not follow the observed diurnal magnetic variation, which are selected as LQDs, potentially affecting the resulting model. To address this, we implemented an LSTM model to automatically remove days that can be considered “fake LQDs”, ensuring a more accurate diurnal variation model for real-time regional geomagnetic index computation.

After training, the results show an accuracy of 0.81 during the validation test, with a loss above 0.5, indicating an acceptable performance of the model. The validation of day classification shows 90% accuracy in predicting LQDs, while 30% of the days were misclassified as non-disturbed.

For the application of this model in real-time analysis, it is crucial to minimize false positives in estimating LQDs. While these results are promising, there remains room for improvement, which we aim to achieve in future work.

### Acknowledgements

The authors gratefully acknowledge the financial support provided by the Jicamarca International Research Experience for this project. We also extend our appreciation to the Geophysical Institute of Peru for supplying the geomagnetic data from their network of magnetometers. Lastly, we thank the organizers of the 20th IAGA Workshop for the opportunity to share the results of this project.

### References

- [1] J. Bartels, "III - THE GEOMAGNETIC MEASURES FOR THE TIME-VARIATIONS OF SOLAR CORPUSCULAR RADIATION, DESCRIBED FOR USE IN CORRELATION STUDIES IN OTHER GEOPHYSICAL FIELDS," in *Geomagnetism*, vol. 4, International Council of Scientific Unions Comité Spécial de l'Année Géophysique Internationale (CSAGI), Ed., in *Annals of The International Geophysical Year*, vol. 4. , Pergamon, 2013, pp. 227–236. doi: <https://doi.org/10.1016/B978-1-4832-1304-0.50007-5>.
- [2] C. Amory-Mazaudier, O. S. Bolaji, and V. Dumbia, "On the historical origins of the CEJ, DP2, and Ddyn current systems and their roles in the predictions of ionospheric responses to geomagnetic storms at equatorial latitudes," *Journal of Geophysical Research (Space Physics)*, vol. 122, no. 7, pp. 7827–7833, Jul. 2017, doi: 10.1002/2017JA024132.
- [3] W. Younas, C. Amory-Mazaudier, M. Khan, and R. Fleury, "Ionospheric and Magnetic Signatures of a Space Weather Event on 25-29 August 2018: CME and HSSWs," *Journal of Geophysical Research (Space Physics)*, vol. 125, no. 8, p. e27981, Aug. 2020, doi: 10.1029/2020JA027981.
- [4] D. J. Knecht and . Shuman B.M, *HANDBOOK OF GEOPHYSICS AND THE SPACE ENVIROMENT*, chapter 4, THE GEOMAGNETIC FIELD. ed by A. S. Jursa, Boston: Air Force Geophysics Laborator, 1985.
- [5] W. Baumjohann and R. A. Treumann, *Basic Space Plasma Physics*. Imperial College Press, 57 Shelton Street Covent Garden London WC2H 9HE, 1999.
- [6] J. W. Gjerloev, "The SuperMAG data processing technique," *Journal of Geophysical Research (Space Physics)*, vol. 117, no. A9, p. A09213, Sep. 2012, doi: 10.1029/2012JA017683.
- [7] M. van de Kamp, "Harmonic quiet-day curves as magnetometer baselines for ionospheric current analyses," *Geoscientific Instrumentation, Methods and Data Systems*, vol. 2, no. 2, pp. 289–304, Nov. 2013, doi: 10.5194/gi-2-289-2013.
- [8] S. S. Chen, C. M. Denardini, L. C. A. Resende, R. A. J. Chagas, J. Moro, and G. A. S. Picanço, "Development of an Empirical Model for Estimating the Quiet Day Curve (QDC) Over the Brazilian Sector," *Radio Science*, vol. 55, no. 12, p. e2020RS007105, Dec. 2020, doi: 10.1029/2020RS007105.