ECAI 2025
I. Lynce et al. (Eds.)

doi:10.3233/FAIA251488

Explainable Artificial Intelligence for Quality Estimation of MARSIS Observations

Benedetta Ferrari[a], Marco Lippi[b,*], Giulio Ganzerli[a], Manuel Iori[a] and Roberto Orosei[c]

[a]Department of Sciences and Methods for Engineering, University of Modena and Reggio Emilia, Italy
[b]Department of Information Engineering, University of Florence, Italy
[c]Institute of Radioastronomy, Italian National Institute of Astrophysics, Italy

Abstract. Planetary remote sensing missions are critical for advancing our understanding of extraterrestrial systems. They operate in highly uncertain environments where reliability and resolution are not always guaranteed, often compromising data analysis and scientific outcomes. In this paper, we consider the challenging task of estimating the quality of the signal acquired by MARSIS, the subsurface sounder aboard ESA's Mars Express mission, which aims to map the presence of liquid water beneath the Martian surface. Quality estimation has a strategic impact on the scheduling of MARSIS observations, since the radar operates with strict constraints that greatly limit the number and size of observation opportunities available per day. Thus, maximizing the quality of scheduled observations becomes a crucial factor in reducing resource utilization and increasing the coverage of the target areas in search of liquid water. To this end, in a previous research we proposed a predict-then-optimize approach, which included a neural network regressor to predict signal quality achievable by future observation opportunities, based on contextual features. In this work, we advance the methodology by applying explainable artificial intelligence techniques that allow domain experts to interpret the results, by enhancing the comprehension of the physical phenomena that have an impact on signal acquisition. Specifically, we applied a SHAP analysis to the neural network predictions and trained an Explainable Boosting Machine (EBM) to provide interpretable models. We then analyzed and compared the results with existing domain knowledge, uncovering promising new avenues for investigation and highlighting limitations in the current dataset construction.

1 Introduction

The Mars Advanced Radar for Subsurface and Ionospheric Sounding (MARSIS) has been operating for 20 years now, since the late deployment of its 40-meter antenna in 2005. It is installed onboard the Mars Express mission, mainly managed by the European Space Agency (ESA) [4]. MARSIS aims to observe the subsurface of Mars to extract information on its layers and their composition. Special attention has been paid to the mapping of liquid water, after its first discovery below the Martian South Pole in 2018 [17]. This discovery gained resonance worldwide because of its implications on the current habitability of Mars. Due to significance in astrobiology, assessing the amount and state of liquid water present on Mars today is considered a key goal in its exploration.

* Corresponding Author. Email: marco.lippi@unifi.it

In subsurface sounding mode, MARSIS transmits 1-MHz-bandwidth pulses at a central frequency of 1.8, 3, 4 or 5 MHz, which penetrate below the surface of Mars, down to a depth of 3.7 km, and are reflected by dielectric discontinuities in the subsurface [21]. These operations, which we call *observations*, are only possible when the spacecraft is less than 800 km above the Martian surface, which corresponds to a time-window lasting 30 minutes per orbit. In addition, due to a strong memory limit, at most a couple of minutes within the time window can be used to collect high resolution data [11]. The data collected are then analyzed and represented as radargrams, showing the reflectivity of the various sub-layers.

Unfortunately, uncertain environmental conditions can badly influence the quality of the data collected by the observations, with a high risk of losing any type of information. Therefore, it is fundamental to select the best possible observations to perform, given their expected quality. This problem was addressed as the Mars Observation Scheduling Problem (MOSP) [8]. Specifically, MOSP aimed to automate and optimize the scheduling of the observations performed by the radar. The decision process was modeled as an optimization problem, involving the selection of the most profitable set of observations from a wide range of future opportunities under specific operational constraints, where the profit is measured as the expected quality of the data collected during an observation. The study was limited to the South Pole of Mars, prioritizing the scientific goal of searching for underground liquid water.

Although the expected quality of future observations is unknown at the time of schedule generation, it depends on several known environmental and technical features. To solve this problem, a two-stage approach has been devised that combines, in sequence, a Machine Learning (ML) algorithm and an optimization procedure, following the so-called predict-then-optimize paradigm (see, e.g., [25]), where the ML algorithm forecasts the potential data quality achievable for each forthcoming observation opportunity, using functions learned from an extensive historical dataset. The results achieved within this setting have been promising and have allowed a thorough evaluation of the scheduling algorithm.

This work moves a step further. We go over the pure predictive aim of the ML algorithm, with the aim of understanding which factors affect the most the quality of a MARSIS observation and, therefore, which make such observation valuable. This is a crucial problem in understanding the physical phenomena that interfere with signal acquisition [16]. To this end, we use techniques that are interpretable-by–design in the context of eXplainable Artificial In-

telligence (XAI) [1]. The main contribution of this research is thus to provide a comprehensive analysis of multiple factors that can have an impact on the acquired signal quality, understanding their relationships and interdependency. The variables considered in our study will cover a wide variety of aspects related to observations, including geometric properties, roughness, solar activity, solar events, and the impact of the ionosphere.

The remainder of this article is organized as follows. Section 2 reviews previous research on artificial intelligence approaches applied to space missions, with a special focus on interpretable approaches. Section 3 describes the MARSIS case study. Section 4 gives details of the methodology used. Section 5 presents the dataset analyzed and the result of the analysis. Finally, Section 6 makes some concluding remarks and highlights several future research directions.

2 Related Works

Artificial intelligence (AI) techniques have seen increased adoption in recent years in the context of space missions [15]. They are used across various mission phases, including mission design and planning, spacecraft health monitoring, and anomaly or fault detection [10, 23]. A particularly active area of research involves the processing and analysis of images collected by satellites orbiting Earth or other planets. This involves image enhancement, classification, and pattern recognition [18]. These techniques are used, for example, to measure land coverage or other outcomes related to sustainable development [2]. Recently, deep neural networks (YOLOv7) have also been applied to detect surface and subsurface features in radargrams obtained from MARSIS observations [20].

Several studies exploit ML algorithms for the prediction of parameters in various contexts. In [5], ML regression techniques were applied to estimate global solar radiation from geostationary satellite data. Specifically, different types of neural networks, support vector regression, and Gaussian processes were tested, considering a complete year of hourly global solar radiation data as the regression target. The input variables included a cloud index, a clear-sky solar radiation model, and multiple reflectivity values derived from Meteosat visible images. In the field of planetary missions, ML algorithms (Gaussian process regression and neural network models) were used to predict seismic energy, and therefore marsquakes, from atmospheric features, such as wind, pressure and temperature [26]. Recently, research in [27] addressed the prediction of solar proton events caused by acceleration processes occurring at the Sun in association with solar flares and coronal mass ejections. This is significant for space mission operations and planning, and for improving monitoring systems of the radiation risk in both deep space and near-Earth environments. To this aim, the authors proposed a random forest algorithm to forecast the energetic proton flux up to one hour ahead by exploiting features derived from the electron flux.

Even if a large variety of applications can be found for AI and ML techniques, few works have been published that employ XAI for space applications. Recently, explainable clustering techniques were applied for the discovery of an equation representing the relationship between solar wind speed, galactic cosmic-ray variations and interplanetary magnetic field intensity observations [24]. The study is applied to the ESA's Laser Interferometer Space Antenna (LISA), which is aimed at the study of gravitational waves and is planned to be launched in 2035.

In spite of the current dearth of applications, XAI appears to have a huge potential in the interpretation of complex phenomena in astrophysics. This research explores that potential, beginning with the case study of MARSIS and the relationship between its activities and the Mars environment.

3 The MARSIS Observation Scheduling Problem

The current MARSIS operative mode limits the acquisition operations to short time window within each orbit around the planet. In addition, due to limited data rate and volume, at most 134 seconds per day can be used to collect high resolution data, which can be stored without on-board processing, thus preserving its full content of information. Then, a down-link window (i.e., a time interval where data can be transmitted to Earth) is needed to empty the memory before a new observation can be acquired. Every observation results in a large amount of data that are analyzed and visualized as radargrams, providing information on the properties of the observed portion of the Martian subsurface. Several factors can influence the quality of data, and hence the effectiveness of the subsequent analysis. Therefore, when planning the acquisition sequence of MARSIS, it is fundamental to consider the expected quality for each observation opportunity, in order to optimize scientific return. The maximization of the expected quality proceeds along with the aim of maximizing the coverage of the planet or specific areas on it, depending on the scientific goals.

MOSP aims to optimize the observation plan of MARSIS, by taking into consideration all these constraints and objectives. Specifically, it studies the coverage of the South Pole of Mars, which is a region of high interest for the mapping of liquid water. The problem falls within a broader class of Satellite Scheduling Problems (SSPs). According to the taxonomy introduced in [7], it is categorized as a regional mapping SSP, as it does not involve scheduling individual observation requests but rather aims to cover a large geographical area. Specifically, in MOSP we have to select a subset of available observation opportunities to perform within a limited time horizon, with the objective of maximizing the total data quality collected from the coverage of the target area. Due to this characteristic, MOSP has previously been formulated as a specific case of the Maximal Covering Location Problem [8].

In the process of selection and scheduling of future observations, an important guideline is the quality with which a specific area can be observed at a certain time (i.e., at a certain position in an orbit). In this context, we define quality as the peak power of the radar signal reflected from the surface and subsurface of Mars, expressed in decibels. The main factors influencing data quality can be grouped into three categories:

- Instrument-related factors: Radar performance may vary due to intrinsic noise, which was not fully characterized during ground testing and is difficult to estimate in-flight. Additionally, the spacecraft altitude as well as the orientation and slope of the two radar antennas may affect measurement accuracy;
- Sun-related factors: Solar activity can significantly affect observation resolution. The extent of the impact may vary depending on the relative position of Mars and the Sun and the phase of the solar cycle;
- Ionosphere-related factors: The Martian ionosphere can disrupt electromagnetic wave transmission, especially since its stability is influenced by solar activity. Sudden solar events, in particular, may cause observation blackouts.

Taking into account all these factors at the same time is a difficult task, which however could greatly increase the effectiveness of the

Table 1: Summary of the features for the ML approaches.

Feature	Description
Orbit number	Id of the orbit of Mars Express
Ephemeris time	Time of the data sampling [s]
Local true solar time	Solar time computed on Mars [s]
Band	Band in which the observation was performed
Altitude	Mars Express probe height above ground level [m]
Latitude	Latitude of the projected position of Mars Express probe [°]
Longitude	Longitude of the projected position of Mars Express probe [°]
X coordinate	Projected x-position of Mars Express probe [m]
Y coordinate	Projected y-position of Mars Express probe [m]
Mars Sun distance	Distance from the center of Mars to center of the Sun [km]
Sun elevation angle	Angle between the Sun rays and the surface [°]
Roughness	Roughness of the point on the surface of Mars [m]
Slope	Slope of the points on the surface of Mars [°]
Solar longitude	Angle between the Mars-Sun line at the ephemeris time and the Mars-Sun line at the vernal equinox [°]
Tangential magnetic field	Magnetic field tangential to the surface [Hz]
Radial magnetic field	Magnetic field perpendicular to the surface [Hz]
Total magnetic field	Sum of the tangential and radial component of the magnetic field [Hz]
Dipole Tilt	Tilt of MARSIS dipole antenna with respect to the nadir [°]
Monopole Tilt	Tilt of MARSIS monopole antenna with respect to the nadir [°]
Quality	Peak power of the electromagnetic signal received by MARSIS during the observation [dB]

acquisition plans. Higher observation quality indeed enables better-resolution radargrams, thus leading to clearer and more informative subsurface analyses. Given the strict resource constraints, it is essential to prioritize high-quality opportunities to maximize early coverage and minimize inefficient use of on-board storage and down-link bandwidth. Although the exact quality of future observations cannot be known in advance, it can be estimated based on observable contextual features, such as local surface morphology and solar elevation angle. These predictive features are detailed in Section 5.1.

4 XAI for Quality Estimation of Observations

In our prior research, several ML techniques have been compared to estimate the quality of MARSIS future observations [8]. After extensive testing, a traditional Deep Neural Network (DNN) was found to be the best performing regressor and was integrated into a predict-then-optimize framework. In this approach, predicted quality values are incorporated into the mathematical formulation of MOSP as profit weights in the objective function, guiding the selection of the most valuable observation opportunities.

However, traditional deep networks are notoriously black-box models that completely lack transparency and interpretability, limiting our ability to understand which factors are the most influential in predicting signal quality. Such information would allow us to better analyze the model, thus proving its robustness, and gain an invaluable insight on the physical phenomena that mostly affect the performance of MARSIS. An explainable model would also be easier to trust, due to its ability to motivate its predictions, in stark contrast with opaque, oracle-like, black-box approaches. In this work, we compare two distinct techniques for XAI: namely, Shapley Additive Explanations (SHAP) and Explainable Boosting Machines (EBMs).

SHAP is a model-agnostic, post-hoc explanation tool that takes advantage of concepts from game theory. Each input feature is considered as a player in a game, and it is credited a certain amount for every prediction following fair allocation results from cooperative game theory. By perturbing the inputs in a controlled fashion, SHAP can determine the contribution of a specific value of a certain input feature to the final prediction. Therefore, this method can be applied to explain a posteriori the prediction of any black-box model, such as a DNN [14].

EBMs, on the other hand, are an interpretable-by-design approach that for tabular data have been shown to perform closely to (or even better than) black-box models [3, 13]. An EBM is an implementation of a Generalized Additive Model (GAM), that learns the prediction function as a sum of non-linear functions, one for each feature, independently. Possibly, also pairs of features can be considered, in which case we speak of GA^{2}M (i.e., a GAM with interaction terms). Being p the number of features, the discriminative function is:

$$f(x) = \sum_{i=1}^{p} f_i(x_i) + \sum_{i=1}^{p} \sum_{j=i+1}^{p} f_{ij}(x_i, x_j) \tag{1}$$

In the case of EBMs, individual functions $f_i(x_i)$ or $f_{ij}(x_i, x_j)$ are implemented as boosting trees. The output of the learning process is thus a collection of trees for each individual feature, which can be easily summarized into graphs depicting the one-dimensional (or two-dimensional, in case of interaction pairs) function, that can be easily represented, ensuring interpretability.

5 Experiments

An experimental evaluation was conducted to assess the performance of both XAI techniques discussed in Section 4, namely the SHAP tool applied to a black-box model and EBMs.

5.1 Dataset and Setting

The MARSIS observation dataset that we adopt for our analysis contains information on the observations performed since its deployment, spanning a period of more than 15 years, from December 6, 2005 (orbit 2438), to January 27, 2021 (orbit 21578). To the aim of our research, we selected only the observations that covered the South Pole of Mars, with their features sampled each 2 seconds. Moreover, we only selected data obtained at the frequency of 4.0 MHz, as it is the most widely used and therefore provides the most reliable information. The final dataset is then composed of 2,083,317 entries, each described by 20 features, including the regression target used for the prediction (i.e., the peak power). Such variables are reported in Table 1. The *ephemeris time* is an astronomical convention, which measures the time (in seconds) starting from January 1, 2000, 12:00 UTC. In the MARSIS dataset, it measures the time when the Mars Express probe reached a certain position, and therefore the time in which all the other features were computed.

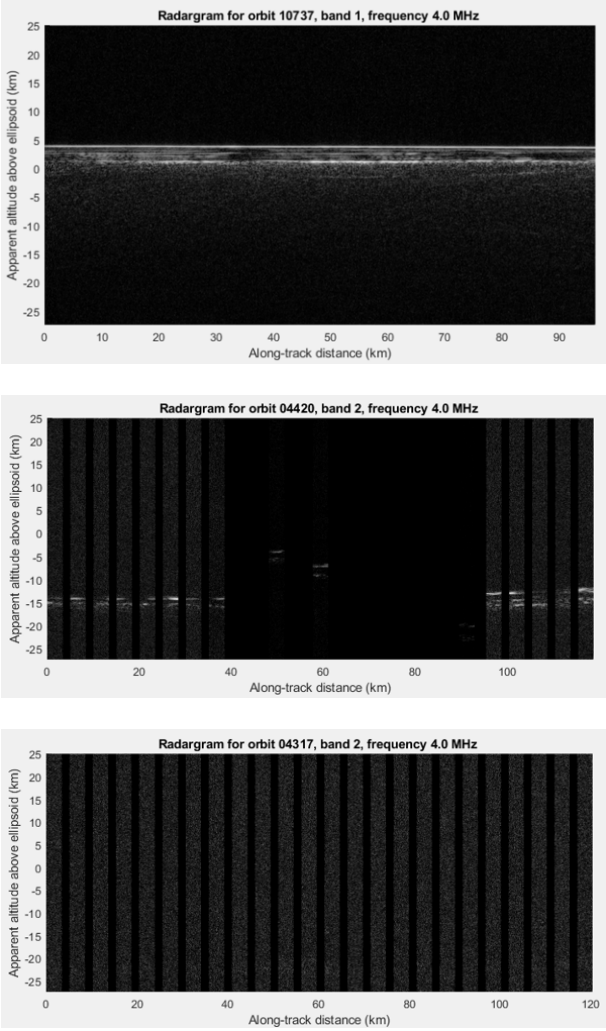


Figure 1: Examples of observation radagrams: high-quality observation (top); observation showing radar malfunction that interrupted the receipt of the signal (middle); observation signal completely canceled as an effect of intense solar activity (bottom).

The dataset was pre-processed to be reliably ingested by ML algorithms. First, it was cleaned from spurious patterns, including samples associated with instrumental malfunctions or notable solar activity (see Figure 1). This led to a reduction of 8.5% of the samples and to 1,905,841 observations in the final dataset. Ordinal features such as orbit number and ephemeris time were omitted, and the solar longitude was converted into its sin and cosine components. Finally, standardization was applied for algorithms sensitive to feature scale.

5.2 Results

We performed a 10-fold cross-validation, splitting the dataset into 10 folds according to the chronological ordering of the orbits. We compared the performance of the considered predictors, a DNN and an EBM, in terms of the average Mean Absolute Error (MAE) computed across the 10 folds. The DNN consists in five dense layers containing, respectively, 800, 400, 200, 100, and 1 neuron(s). The model was implemented in the `keras` library. For EBM, we used the default implementation in the `interpretml` library. Not surprisingly, the performance of the DNN is slightly superior to that of the EBM, with a MAE equal to 3.37 against 4.50. Yet, performance in terms of regression error is not the main focus of this work, where our goal is to investigate the explanations behind the physical phenomenon of signal acquisition: this aspect will be covered in the subsequent sections. However, we hereby remark that the achieved performance in terms of regression error is sufficient to successfully implement the predict-and-optimize approach to improve MARSIS scheduling operations [8].

5.2.1 SHAP analysis

As a first attempt to investigate the most important factors in the quality of signal acquisition, we applied the SHAP technique to the trained DNN. Due to the high dimensionality of the dataset, and to the large computational cost of the SHAP analysis, we tested it on a random sample corresponding to 1% of the dataset. The result of the analysis is shown in Figure 2. The figure reports, for each feature in the dataset, a violin shape that represents the impact of the corresponding feature on the regression target, depending on the magnitude of the values (represented by a color scale). Individual violin plots are ordered by the importance of the corresponding feature on the model output (sum of the absolute values of the SHAP values per feature). Each violin displays the distribution and density of SHAP values for the respective feature.

The SHAP analysis reveals that the most influential feature is the Sun Elevation Angle (SEA). This result is not surprising, as SEA significantly affects observation quality, typically with a positive effect when values are low (i.e., during nighttime), and a negative one during daylight hours. This pattern aligns with prior observations (see Figure 3a), which show a Pearson correlation coefficient of -0.623, and with theoretical expectations, as high SEA values increase ionospheric particle activity, which interferes with signal transmission and causes radargram distortion [9].

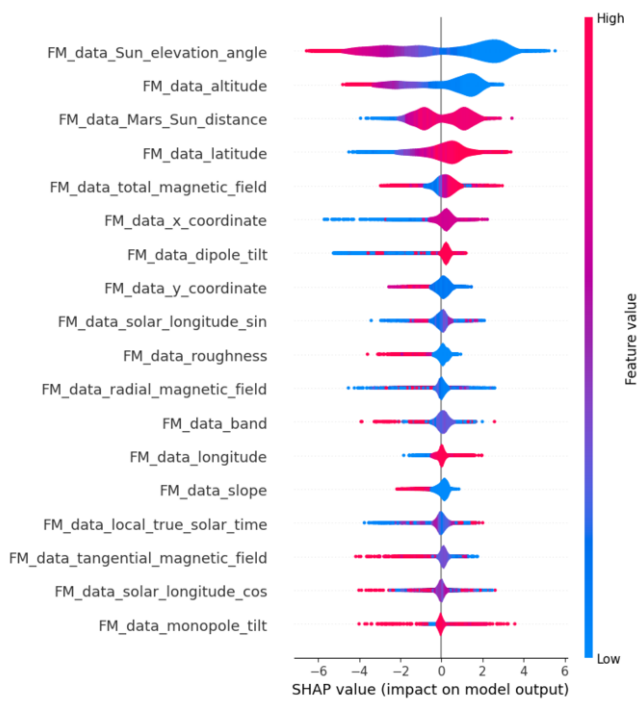
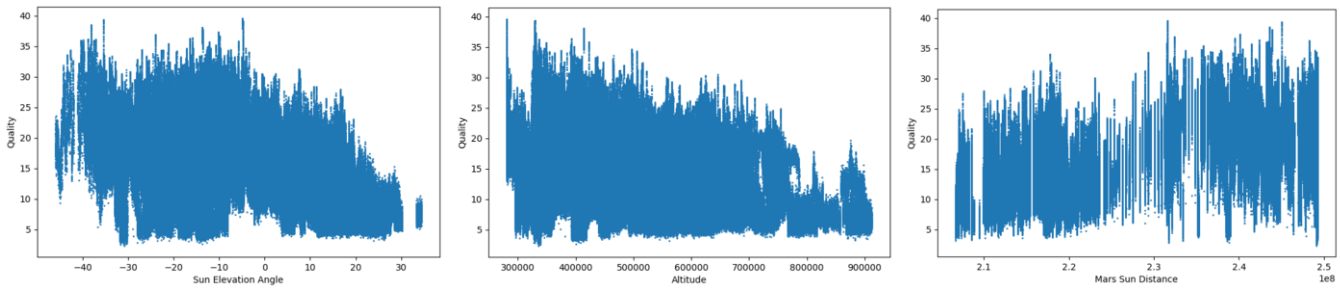


Figure 2: SHAP values distribution for each feature of the dataset. A high feature value corresponds to a positive impact on the observation quality, while a low value corresponds to a negative impact.



(a) Correlation between SEA and quality (b) Correlation between spacecraft altitude and quality (c) Correlation between MSD and quality

Figure 3: Correlation plots of the three most influencing features derived from historical data. The visual representation confirms the results of the SHAP analysis.

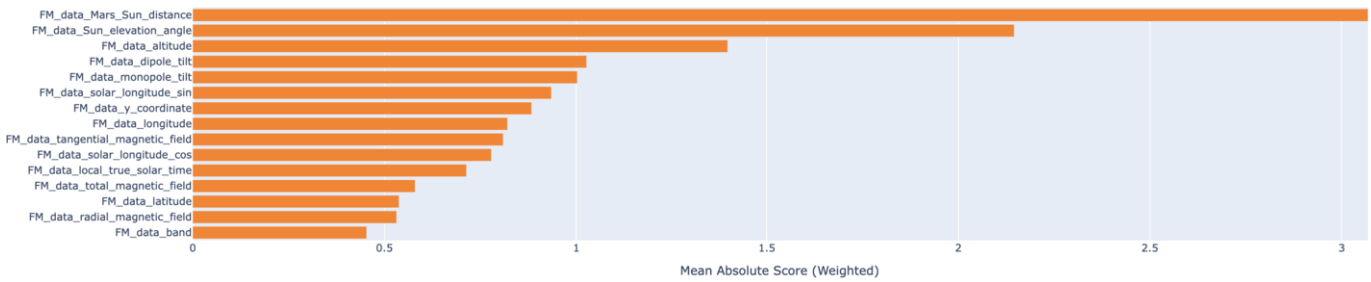


Figure 4: Feature importance at global level, estimated by the EBM regressor (top 15 features are shown).

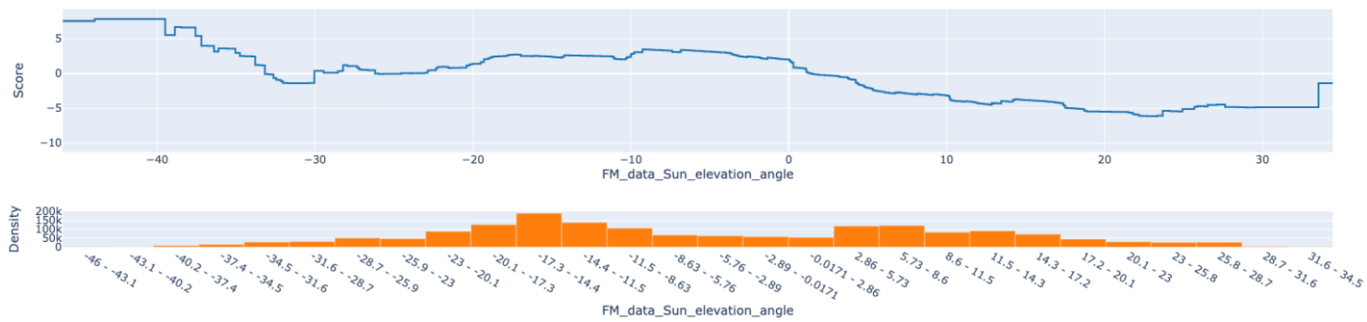


Figure 5: Function learned by the EBM model for the Sun Elevation Angle (SEA) feature (top) and corresponding density histogram (bottom).

Data altitude exhibits a similar trend: higher spacecraft altitudes tend to reduce observation resolution, as confirmed by historical data represented in Figure 3b, where a correlation coefficient of -0.320 confirms a moderate negative correlation. This relationship is also supported by findings in [19]. Another notable result is the importance of the Mars-Sun Distance (MSD), which was unexpectedly ranked among the most impactful features, reporting a correlation value of 0.566. Specifically, observations tend to improve when Mars is farther from the Sun, likely due to reduced exposure to solar particles (Figure 3c). However, a high MSD correlates with poorer observation quality for a significant subset of cases, as also shown by the bimodal distribution of the SHAP values. This apparent contradiction will be explored in more detail later, with EBMs.

As expected, surface roughness and slope also influence observation quality, as irregular surface can cause signal deviation from the nadir direction. Additionally, the 40-meter dipole tilt appears to negatively affect observations when it deviates significantly from the optimal 90-degrees position. Finally, the importance of positional features may be attributed to the geophysical characteristics of specific regions.

5.2.2 *Explainable Boosting Machine*

The results provided by the SHAP analysis were considered interesting by domain experts, although some aspects deserved a deeper investigation. EBMs enable such a detailed analysis by assessing how individual features – or pairs of features – in the dataset influence observation quality. The global ranking of feature importance obtained with EBM is very close to that obtained with DNN (see Figure 4). The two most important features are related to solar activity, i.e., SEA and MSD. Therefore, we investigated in more detail the impact of SEA: Figure 5 displays the score function for SEA, illustrating its impact on observation quality (y-axis) as a function of the values of the variable (x-axis). The orange histogram also shows the distri-

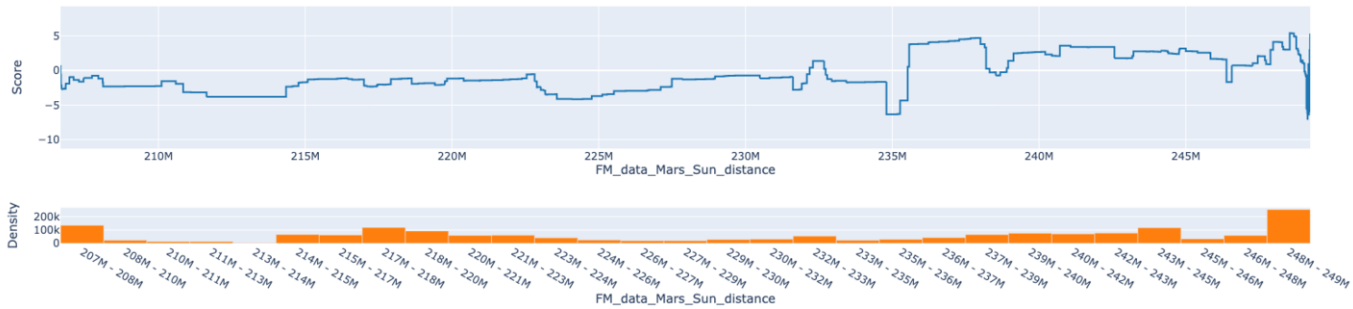


Figure 6: Function learned by the EBM model for the Mars-Sun Distance (MSD) feature (top) and corresponding density histogram (bottom).

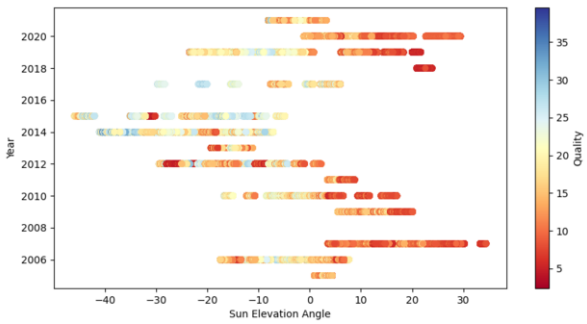


Figure 7: Representation of the SEA and the year of the corresponding observation. Color scale describes observation quality.

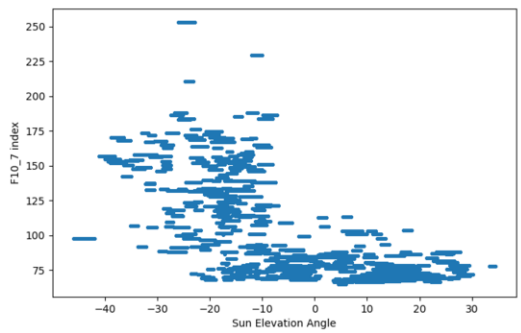


Figure 8: Correlation between SEA and F10_7 index. Several nighttime observations were performed when the solar activity, and hence the F10_7 index, was high.

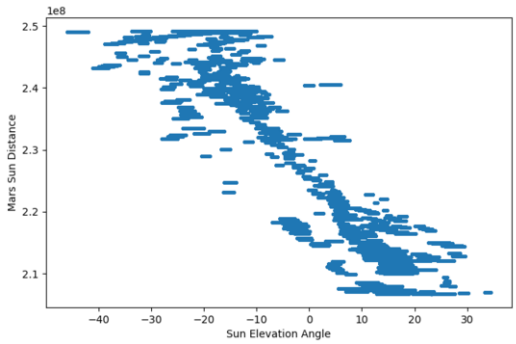


Figure 9: Correlation between SEA and MSD. A clear negative trend is visible.

bution of the variable. Generally, a clear negative trend is observed, which confirms the bad effect of SEA values above zero-degrees, in agreement with the SHAP analysis. Yet, notably, an off-trend behavior is observed in the SEA range between –35 and –20 degrees, where the contribution to signal quality quickly degrades. To investigate this anomaly, we conducted a focused analysis using traditional statistical and data visualization techniques. Figure 7 plots each observation sample according to its SEA value and the year in which it was acquired. The color scale indicates the associated observation quality. A sinusoidal pattern emerges, reflecting the orbital dynamics and seasonal opportunities to observe the Martian South Pole. Interestingly, nighttime observations in the [–35, –20] degree interval occurred primarily between 2012 and 2015, a period that coincides with the peak activity of Solar Cycle 24 [12].

To confirm this idea, we introduced a new variable into the analysis: the F10_7 index. This index represents solar activity levels and is only available after observation, due to the intrinsic difficulty in forecasting it. For this reason it was not included in the initial model. The relation between SEA and F10_7 is shown in Figure 8. The diagram shows a clear increase in the F10_7 index in correspondence with the night-time SEA range: in particular, we notice how all observations with F10_7 values above the third quartile were conducted during the night. Therefore, a significant sample of night-time observations were heavily affected by elevated solar activity, which negatively impacted their quality despite favorable SEA conditions.

The MSD score function (Figure 6), still linked to solar activity, shows a highly irregular shape, although a positive trend is still visible: the solar flux typically decreases with the square of the distance, and thus it is not surprising to observe that larger values of MSD usually correspond to better signal quality. To explore further, we also considered the scatterplot representing SEA and MSD (Figure 9) for which we observed a distinct linear trend. Specifically, when Mars was closer to the Sun (low MSD values), observations were predominantly conducted during daylight (high SEA values). Conversely, when Mars was farther from the Sun (high MSD values), observations were more likely to occur at night (low SEA values). This finding suggests that high-MSD observations benefited from lower SEA values. At the same time, night-time observations frequently coincided with periods of high solar activity (Figure 8). As a result, in many cases, high MSD values were still associated with degraded signal quality, as shown by SHAP values.

5.3 *Discussion*

The implemented XAI techniques produced consistent results. Most of the findings align with theoretical expectations derived from previous studies on the influence of individual features on observation quality. However, differently from previous research, our approach enables a comprehensive analysis of all relevant features simulta-

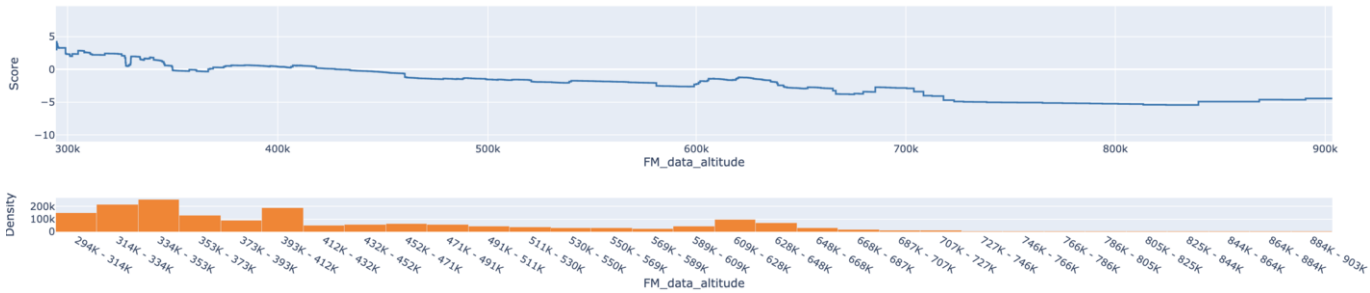


Figure 10: Function learned by the EBM model for the spacecraft altitude feature (top) and corresponding density histogram (bottom).

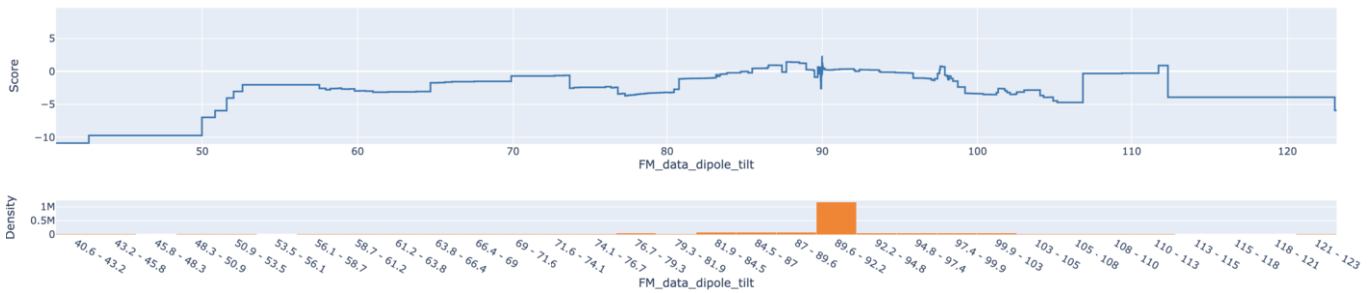


Figure 11: Function learned by the EBM model for the dipole tilt feature (top) and corresponding density histogram (bottom).

neously, allowing us to assess their relative importance in a unified framework. This evaluation is particularly valuable for identifying key parameters to be monitored when designing observation plans.

Overall, observation quality strongly depends on three families of variables: (i) instrumental activity; (ii) solar activity; (iii) ionosphere.

Regarding instrumental activity, Figures 10 and 11 show the impact of spacecraft altitude and dipole tilt angle, respectively: the former clearly shows a clear decrease in signal quality as the spacecraft altitude grows, whereas the latter exhibits a distortion for values different from the optimal tilt values, known to be equal to 90 degrees.

Regarding solar activity, the specific analyses conducted on SEA and MSD revealed a potential bias in the dataset, which is influenced by orbital dynamics and also human choice of when performing the observations. For these reasons, models cannot be trained on samples that fully cover the data distribution, and this may introduce errors both in the estimated quality, and in the interpretation.

Regarding the third factor, namely the impact of the ionosphere, the current dataset does not contain variables that directly measure it. Yet, its importance is underlined by other related variables, such as solar activity, which is known to have a direct effect on the polarization of ionosphere, which can induce a degradation in signal transmission and acquisition. In the future, to achieve a more accurate characterization of the factors affecting observation quality, it is essential to expand the feature set. In particular, introducing additional variables that explicitly represent solar activity (e.g., solar flux indices, solar wind parameters) and ionospheric conditions could help mitigate these effects and improve model generalization.

The results of SHAP and EBM would allow experts to confidently rank and evaluate the weight of each environmental parameter when manually establishing an observation plan. In addition, we are integrating DNN into the decision process, providing an effective tool that predicts the quality of future observations and automatically generates optimized plans. Although the use of XAI does not directly improve the prediction performance over DNNs, explainable methods would allow AI models to gain the trust of their users, and therefore enable their operational deployment.

6 Conclusions

This study demonstrates the value of integrating eXplainable AI (XAI) techniques to analyze the factors influencing observation quality in planetary missions. By exploiting SHAP analysis and EBMs, we provided insights that are both consistent with theoretical expectations and capable of revealing previously unnoticed patterns. Key features such as the Sun Elevation Angle (SEA), spacecraft altitude, magnetic field, and surface properties were confirmed to play significant roles in observation performance. Additionally, the Mars-Sun Distance (MSD) has shown an interesting behavior that needs further, more detailed investigation.

Targeted analyses also revealed potential biases in the dataset, particularly due to orbital geometry and solar activity, that may affect interpretation. These findings underline the need to incorporate additional variables related to solar and ionospheric behavior to further refine model accuracy and physical interpretability.

To conclude, this work represents a first advancement in the application of XAI to planetary science, supporting both informed operational decision-making and scientific exploration. The comprehensive feature evaluation offers practical guidance for future observation planning, overcoming the drawbacks of black-box regressors previously trained for predicting observation quality. Similar approaches could allow mission designers to prioritize the monitoring and control of the most impactful parameters. Moreover, expanding this kind of analysis to additional variables will also provide new and independent insights regarding the factors that influence radio wave propagation across the Martian ionosphere. Finally, the same kind of methodology could be applied to future planetary missions, such as BepiColombo for the observation of Mercury [6] and Juice for the icy moons of Jupiter [22].

Acknowledgements

We kindly acknowledge financial support by the Italian Ministry of University and Research, under project PRIN2022PNRR - M4C2INV1.1, NextGenerationEU - Award 1409/2022 - Project CALIPSO, grant n. P2022XF72W; by the Italian Space Agency (ASI) through contract 2024-40-HH.0; by the CAI4DSA action (Collaborative Explainable neuro-symbolic AI for Decision Support Assistant), of the FAIR national project on artificial intelligence, PE 1 PNRR (https://fondazione-fair.it/).

References

[1] A. B. Arrieta, N. Díaz-Rodríguez, J. Del Ser, et al. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information fusion*, 58:82–115, 2020.

[2] M. Burke, A. Driscoll, D. B. Lobell, and S. Ermon. Using satellite imagery to understand and promote sustainable development. *Science*, 371(6535):eabe8628, 2021.

[3] R. Caruana, Y. Lou, J. Gehrke, P. Koch, M. Sturm, and N. Elhadad. Intelligible models for healthcare: Predicting pneumonia risk and hospital 30-day readmission. In *Proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining*, pages 1721–1730, 2015.

[4] A. Chicarro, P. Martin, and R. Trautner. The Mars Express mission: an overview. In A. Wilson and A. Chicarro, editors, *Mars Express: the Scientific Payload*, volume 1240 of *ESA Special Publication*, pages 3–13, 2004.

[5] L. Cornejo-Bueno, C. Casanova-Mateo, J. Sanz-Justo, and S. Salcedo-Sanz. Machine learning regressors for solar radiation estimation from satellite data. *Solar Energy*, 183:768–775, 2019.

[6] G. Cremonese, F. Capaccioni, M. Capria, and al. Simbio-sys: Scientific cameras and spectrometer for the BepiColombo mission. *Space Sci Rev*, 216(75), 2020. URL https://doi.org/10.1007/s11214-020-00704-8.

[7] B. Ferrari, J.-F. Cordeau, M. Delorme, M. Iori, and R. Orosei. Satellite scheduling problems: A survey of applications in Earth and outer space observation. *Computers & Operations Research*, page 106875, 2024.

[8] B. Ferrari, M. Delorme, M. Iori, M. Lippi, and R. Orosei. Mars observation scheduling problem: Optimizing the search for underground water. Technical report, Università di Modena e Reggio Emilia, 2024.

[9] D. A. Gurnett, D. L. Kirchner, R. L. Huff, D. D. Morgan, A. M. Persoon, T. F. Averkamp, F. Duru, E. Nielsen, A. Safaeinili, J. J. Plaut, and G. Picardi. Radar Soundings of the Ionosphere of Mars. *Science*, 310 (5756):1929–1933, Dec. 2005. doi: 10.1126/science.1121868.

[10] S. K. Ibrahim, A. Ahmed, M. A. E. Zeidan, and I. E. Ziedan. Machine learning techniques for satellite fault diagnosis. *Ain Shams Engineering Journal*, 11(1):45–56, 2020.

[11] R. Jordan, G. Picardi, J. Plaut, et al. The Mars Express MARSIS sounder instrument. *Planetary and Space Science*, 57(14-15):1975–1986, 2009.

[12] K. Kaplan. The characteristic properties of solar activity in solar cycle 24. *Kinematics and Physics of Celestial Bodies*, 40(2):105–115, 2024.

[13] Y. Lou, R. Caruana, and J. Gehrke. Intelligible models for classification and regression. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 150–158, 2012.

[14] S. M. Lundberg and S.-I. Lee. A unified approach to interpreting model predictions. *Advances in neural information processing systems*, 30, 2017.

[15] P. A. Oche, G. A. Ewa, and N. Ibekwe. Applications and challenges of artificial intelligence in space missions. *IEEE Access*, 12:44481–44509, 2021.

[16] R. Orosei, R. Jordan, D. Morgan, et al. Mars advanced radar for subsurface and ionospheric sounding (marsis) after nine years of operation: A summary. *Planetary and Space Science*, 112:98–114, 2015.

[17] R. Orosei, S. E. Lauro, E. Pettinelli, et al. Radar evidence of subglacial liquid water on Mars. *Science*, 361(6401):490–493, 2018.

[18] H. Ouchra, A. Belangour, and A. Erraissi. Machine learning for satellite image classification: A comprehensive review. In *2022 International Conference on Data Analytics for Business and Industry (ICDABI)*, pages 1–5. IEEE, 2022.

[19] G. Picardi, D. Biccari, R. Seu, L. Marinangeli, W. T. K. Johnson, R. L. Jordan, J. Plaut, A. Safaenili, D. A. Gurnett, G. G. Ori, R. Orosei, D. Calabrese, and E. Zampolini. Performance and surface scattering models for the Mars Advanced Radar for Subsurface and Ionosphere Sounding (MARSIS). *Planetary and Space Science*, 52:149–156, Jan. 2004. doi: 10.1016/j.pss.2003.08.020.

[20] R.-E. Pierau, M. Cartacci, A. Cicchetti, B. Ferrari, M. Iori, A. Meehan, R. Orosei, H. Rezatofighi, and P. J. Stuckey. Detection of subsurface layer instances on radargrams from Mars through hyper-spectral image construction. In *Congresso Nazionale di Scienze Planetarie*, number 20, February 2025.

[21] J. J. Plaut, G. Picardi, A. Safaeinili, et al. Subsurface radar sounding of the south polar layered deposits of mars. *Science*, 316(5821):92–95, 2007.

[22] F. Poulet, G. Piccioni, and Y. L. et al. Moons and Jupiter imaging spectrometer (MAJIS) on Jupiter Icy Moons Explorer (JUICE). *Space Sci Rev*, 220(27), 2024. URL https://doi.org/10.1007/s11214-024-01057-2.

[23] A. Russo and G. Lax. Using artificial intelligence for space challenges: A survey. *Applied Sciences*, 12(10):5106, 2022.

[24] F. Sabbatini, C. Grimani, and R. Calegari. Bridging machine learning and diagnostics of the ESA LISA space mission with equation discovery via explainable artificial intelligence. *Advances in Space Research*, 74 (1):505–517, 2024.

[25] U. Sadana, A. Chenreddy, E. Delage, A. Forel, E. Frejinger, and T. Vidal. A survey of contextual optimization methods for decision-making under uncertainty. *European Journal of Operational Research*, 320(2): 271–289, 2025.

[26] A. E. Stott, R. F. Garcia, A. Chédozeau, A. Spiga, N. Murdoch, B. Pinot, D. Mimoun, C. Charalambous, A. Horleston, S. D. King, et al. Machine learning and marsquakes: a tool to predict atmospheric-seismic noise for the NASA InSight mission. *Geophysical Journal International*, 233 (2):978–998, 2023.

[27] M. Stumpo, M. Laurenza, S. Benella, and M. F. Marcucci. Predicting the energetic proton flux with a machine learning regression algorithm. *The Astrophysical Journal*, 975(1):8, 2024.