

Satellite Connectivity Prediction for Fast-Moving Platforms

Chao Yan*, Babak Mafakheri*

* Safran Passenger Innovations GmbH, Munich area, Germany;

emails: {name.lastname}@zii.aero

Abstract—Satellite connectivity is gaining increased attention as the demand for seamless internet access, especially in transportation and remote areas, continues to grow. For fast-moving objects such as aircraft, vehicles, or trains, satellite connectivity is critical due to their mobility and frequent presence in areas without terrestrial coverage. Maintaining reliable connectivity in these cases requires frequent switching between satellite beams, constellations, or orbits. To enhance user experience and address challenges like long switching times, Machine Learning (ML) algorithms can analyze historical connectivity data and predict network quality at specific locations. This allows for proactive measures, such as network switching before connectivity issues arise. In this paper, we analyze a real dataset of communication between a Geostationary Orbit (GEO) satellite and aircraft over multiple flights, using ML to predict signal quality. Our prediction model achieved an F1 score of 0.97 on the test data, demonstrating the accuracy of machine learning in predicting signal quality during flight. By enabling seamless broadband service, including roaming between different satellite constellations and providers, our model addresses the need for real-time predictions of signal quality. This approach can further be adapted to automate satellite and beam-switching mechanisms to improve overall communication efficiency. The model can also be retrained and applied to any moving object with satellite connectivity, using customized datasets, including connected vehicles and trains.

Index Terms—Machine Learning, LEO, GEO, Handover

I. INTRODUCTION

A communication satellite is an artificial satellite designed to provide communication service by transmitting signals through a transponder, creating a link between a transmitter and a receiver located at different points on Earth. These satellites play a critical role in enabling a wide range of services. Satellites move around the Earth due to the gravitational pull along specific paths called orbits. These orbits can vary in altitude and some other factors, depending on the satellite's purpose.

There are various satellite providers nowadays on the market, such as ViaSat, Starlink, SES Sirius, and so on. Starlink satellites mainly run in Low Earth Orbit (LEO), at altitudes from 340 km to 1,200 km above the Earth. Viasat satellites operate in GEO. In this orbit, the satellites match the Earth's rotation and are positioned at around 35,786 km above the equator, in a fixed position to the Earth. SES Sirius is Medium Earth Orbit (MEO) satellite provider, having their satellites operating at an altitude of around 8,000 km. GEO, MEO, and LEO satellites are the three commonly used communication satellite systems.

Satellite communication is widely used in various scenarios, ranging from everyday individual use to business purposes and critical systems for overall infrastructure. The services include telecommunications, global navigation systems, autonomous systems, maritime and aviation, and other fields. Satellites' communication can widely be used for mobile phones, TV, and internet [1][2]. Global navigation systems are utilized for providing global positioning services via Global Positioning System (GPS)[3][4]. Autonomous systems are critical for autonomous vehicles which require high-precision GPS so that tasks such as optimizing routes, avoiding collisions, and recognizing lanes can be performed well [5][6]. Aircraft and ships rely highly on satellites for communication, navigation and so on. Satellite communication provides internet access to crew and passengers during flights at high altitudes targeting on-board internet access, also called In-Flight Connectivity (IFC) [7], [8], and cruises in the middle of the ocean [9][10][11]. On commercial aircraft, good IFC is extremely demanding, since it can enhance passenger satisfaction to some extent which is what the airlines are pursuing [12], [13]. Handover (HO) is another important term in satellite communication which means the whole step of switching an ongoing communication session from one ground station or satellite to another, during which seamless and uninterrupted connectivity should be secured. This is especially critical in aviation since the aircraft is moving at a high speed.

We have already demonstrated the usage of a LEO satellite for IFC in [7] and introduced an AI-based In-Flight Entertainment and Connectivity (IFEC) system in [14]. Moreover, there are several works discussing the LEO, GEO satellite networks, IFC, handover (HO), and application of ML in this domain [15], [16], [8], etc. Authors in [17] proposed a new method for IFC link allocation, which could make dynamic switches between terrestrial cellular and LEO satellite networks in real-time. The paper [15] summarizes the current ML methods for Global Navigation Satellite Systems (GNSS) based positioning, their advantages, disadvantages, and potential challenges. Min J [16] proposes a nonorthogonal multiple access (NOMA) in LEO satellite communication systems using ML techniques. The experiment results prove that the performance is comparable to the optimal one by using standard calculation. On the other hand, [18] proposes solutions based on ML for optimizing handover decisions in non-terrestrial networks (NTN), focusing on decreasing the

issue of signaling storms during handovers. The results reveal that achieving optimal HO performance requires considering the distance between the cell center and the user, which is a critical variable in NTN. M Chen [19] proposes a Q-learning-based HO scheme for LEO satellite networks that utilizes both remaining service time and signal quality to optimize the performance of handovers. Experimental results show that this method increases overall signal quality and decreases the number of HOs compared to traditional methods. Feng L[20] presents a scheme by leveraging the timely and accurate orbit and position data of the aircraft to facilitate seamless HO. This scheme not only decreases the packet drop rate during HO but also minimizes data transmission delays. Moreover, S. Mondal [21] proposes a prediction-based solution for HO selection. They derive the cost function and constraints based on dual connectivity variables over the prediction horizon and solve the problem using a 2D Genetic Algorithm (2D-GA) to achieve the best HO solution. The results indicate that network densification combined with the predictive control model improves overall aircraft performance.

Although existing studies propose methods to improve satellite handover, a prediction model specifically tailored to moving platforms can significantly enhance satellite communication handover mechanisms. To the best of our knowledge, no prior research has analyzed historical data to develop a prediction model for satellite connectivity of moving platforms. In this context, our objective is to create a satellite-aircraft link quality prediction model that provides comprehensive analysis into link conditions throughout an entire flight across all potential geographic coordinates. The primary goal of this model is to predict the satellite signal quality in different geographical coordinates which helps enable an automated handover mechanism capable of making network switching decisions based on the aircraft's location and other real-time data such as weather conditions during flight.

The main contribution of this paper is the development of an accurate prediction model capable of forecasting the signal quality of GEO satellites on moving aircraft. This model enables airline operators to take proactive measures before encountering issues related to connection failures, such as beam-switching or network switching. By utilizing the most relevant features for optimal signal quality prediction, the model is categorized into high-altitude and low-altitude scenarios, with low-altitude scenarios incorporating weather data as an influential factor. This adaptability allows the model to be applied to other moving objects on the ground, such as cars and trains.

The remainder of the article is structured as follows: Section II discusses the methodology, including dataset processing, metrics selection, and training. Section III presents the experimental results, and Section IV concludes the work with a discussion of future directions.

II. METHODOLOGY

This section outlines the methods used to analyze the dataset and develop models for predicting the received signal quality

of a Geostationary Orbit (GEO) satellite on an aircraft during flight. We detail the steps involved in data processing, the criteria for selecting performance metrics, and the machine learning experiments conducted. The results of the prediction model are also discussed with a focus on the model's effectiveness in predicting signal quality.

A. Dataset Processing

The dataset used in this study consists of real-world data collected from several flights, capturing detailed logs of aircraft operations. It contains a comprehensive set of parameters, including the status of the aircraft's software and hardware as well as all relevant communication metrics throughout each flight. The logging process begins after the aircraft reaches a certain altitude post-takeoff and records data at one-minute intervals. Thus, it provides a continuous snapshot of communication conditions for nearly the entire duration of each flight.

Within this dataset, we selected a subset for training and experimentation, comprising seven months of real-world data from multiple flights conducted in 2023. The data was extracted from a commercial database logged within the Reliable Affordable and Very Easy (RAVE) [22] system. The initial dataset contains over 10 million entries, with each log containing up to 81 features. These features include numeric values, such as the aircraft's altitude, longitude, and latitude, as well as categorical features, such as TailNumber, AirlineCode, DepartureAirport, and ArrivalAirport. Time-related features, like LogDate, FlightStartTime, and FlightEndTime, are stored in time format. Of particular interest is the downlink carrier feature, Carrier to Noise ratio (CNR), measured in dB, which serves as the target feature for our machine learning model.

First, we analyzed the distribution of CNR values. As shown in Figure 1, most CNR values range between 8 and 10 dB, with a mean of 8.82 dB and a median of 8.8 dB. Since many features remained constant throughout the flight, we filtered out these static features and selected 12 key features for our experiment, including the target feature. To further structure our analysis, we created two distinct datasets. Dataset 1 focuses on the top five airline routes with the highest number of recorded logs, primarily consisting of long-haul flights. Dataset 2 contains the full dataset, covering all recorded flights across various routes. The purpose of using two datasets is to compare the models' performances between specific flight paths (Dataset 1) and general flight paths (Dataset 2).

B. ML Experiment Conduction & Metrics Selection

We initially conducted the experiment using the original numeric CNR values, building regression models to predict these values, with Mean Squared Error (MSE) and Mean Absolute Error (MAE) as our validation metrics. MSE and MAE are commonly used in regression tasks. However, the results for both MAE and MSE were around 0.2, which is high and not ideal. Our ultimate goal is to predict CNR values to trigger a network switching, including HO when necessary. This means that as long as the CNR reaches a specific threshold or range,

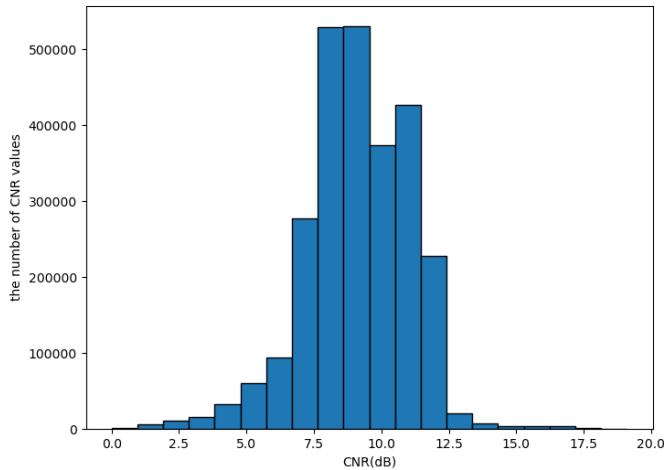


Fig. 1: Distribution of CNR Values

a switch should be performed. In other words, precise CNR values are not required directly for this purpose; instead, we can predict the range of CNR values to determine when a switch should be activated and this could increase the error-tolerant rate in our use case. Therefore, we transformed the continuous CNR values into categorical ranges, where predicting the correct category would be sufficient for identifying when to activate HO. Dividing the CNR values into categories presented another challenge. Different modulation types and coding rates have different CNR requirements. In our case, we convert the CNR values into 4 categories: Good 15-20 dB; Medium 10-15 dB; Weak 6-10 dB; and Bad <6 dB, which is based on our use case. Our use case then turns into a multi-class classification problem with imbalanced dataset labels. In such scenarios, accuracy or precision alone may not provide a comprehensive assessment of model performance, as they do not account for the imbalance in class distribution. Instead, the F1 score is more appropriate, as it balances precision and recall, offering a better evaluation of the model's effectiveness across all classes.

We then randomly selected a sample flight and plotted the CNR values along the flight path. Fig.2 shows the CNR values of the selected flight from Singapore to London, with data recorded every minute. We noticed that the CNR values fluctuated during the flight, with certain segments showing low performance, falling into the Medium or Bad categories, or even missing entirely. Several factors could contribute to these fluctuations, including weather conditions at low altitudes, restricted areas, or other contributing factors.

In our experiment results, we observed that the prediction model's performance was worse at low altitudes compared to higher altitudes. We believe the lack of weather data in our dataset could be a vital issue leading to this. To address and verify the impact of weather data on model performance, we divided our model into two sub-models: one for low-altitude data, which incorporates weather information, and another for high-altitude data, which excludes weather data.

We sought external sources and identified some open-source weather datasets for our second sub-model. Historical weather records are typically available on an hourly basis, whereas our dataset is recorded at one-minute intervals. After evaluating options, we selected a historical weather data provider [23]. To integrate the weather data with our primary dataset, we utilized geo-coordinates and timestamps as common reference points. For each instance in our dataset, we identified corresponding weather conditions by matching its recorded coordinates and timestamp to the nearest available weather data point (precision to about 10 km). This approach ensured alignment between the two datasets and allowed us to incorporate environmental factors into our analysis. For model development, we utilized the open-source H2O AutoML framework[24] in our local environment. H2O AutoML simplifies the machine learning model-building process and provides a user-friendly interface that enables data scientists to construct reliable models efficiently. It includes built-in algorithms such as H2O Gradient Boosting Machines, XGBoost Gradient Boosting Machines, Generalized Linear Models, Distributed Random Forests, Deep Learning, and Stacked Ensemble models (generated by combining other base models).

III. EXPERIMENTS RESULT

We conducted multiple experiments using two datasets under varied conditions, as shown in Table I. This table presents datasets filtered by different altitude thresholds, unique satellite identifiers, and the presence or absence of weather information.

First, Indexes 1 to 4 compare various altitude thresholds within Dataset 1. We observed that the prediction model becomes more reliable at higher altitudes, achieving its best result with an F1 score of 0.97349 when filtering for altitudes above 6000 m. Given these findings, we assumed that incorporating weather data might improve prediction accuracy, especially at lower altitudes where performance declines.

To test this, we used Index pairs 5 and 6 for Dataset 1, as well as 7 and 8 for Dataset 2, to examine the impact of weather information on predictions at lower altitudes. As expected, we saw a slight increase in F1 scores for both datasets when weather data was included, indicating that weather conditions positively influence prediction accuracy in these scenarios.

Meanwhile, the F1 score of the model using Dataset 1 (Index 1) is higher than that using Dataset 2 (Index 9), and they have the same altitude threshold (altitude >6000 m). The same trend happens also in Indexes 5 and 7, with the identical altitude threshold, the F1 score of Index 5 is higher than Index 7. This is likely because Dataset 1 contains more similar data (focusing on long-haul flights), whereas Dataset 2 is more generalized.

Lastly, Indexes 10 to 14 correspond to models trained on data from individual satellites. Among these, CNR predictions based on Dataset 2 from Satellite 15F3 outperform those from the other satellites. However, further research is required to determine the underlying reasons for this performance.

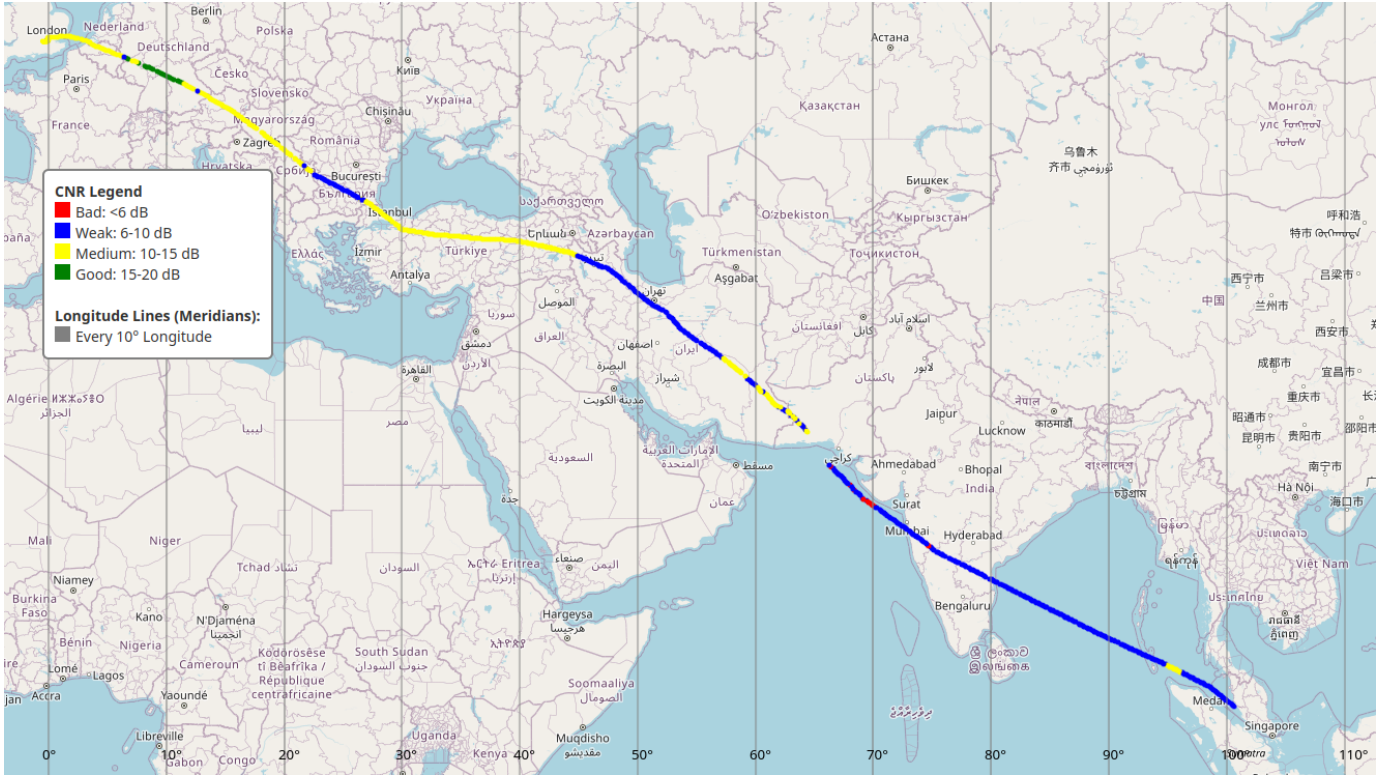


Fig. 2: CNR values of some flight from Singapore to London

TABLE I: F1 Scores of Various ML Models

Index	Dataset	Dataset Size	F1
1	Dataset 1 with altitude >6000m	811,233	0.97349
2	Dataset 1 with altitude >5000m	817,817	0.97065
3	Dataset 1 with altitude >4000m	824,659	0.97172
4	Dataset 1 with altitude >3000m	832,284	0.97133
5	Dataset 1 with altitude <3000m	161,388	0.94569
6	Dataset 1 with altitude <3000m + 4 weather data	161,388	0.96054
7	Dataset 2 with altitude <3000m	230,685	0.90216
8	Dataset 2 with altitude <3000m + 4 weather data	230,685	0.91685
9	Dataset 2 with altitude >6000m	2,617,623	0.96802
10	Dataset 2 with altitude >6000m for I5F1	804,578	0.95916
11	Dataset 2 with altitude >6000m for I5F2	563,815	0.95421
12	Dataset 2 with altitude >6000m for I5F3	434,963	0.96804
13	Dataset 2 with altitude >6000m for I5F4	226,659	0.95700
14	Dataset 2 with altitude >6000m for GX5	37,353	0.914340

IV. CONCLUSION

In this paper, we developed and evaluated machine learning models to predict CNR values during flights, with the ultimate goal of improving satellite handover mechanisms. Our results demonstrate that at higher altitudes, the model's performance is more reliable, achieving an F1 score of 0.97349 with selected flights. This indicates the model's high effectiveness in predicting signal quality, which could be potentially used in customized flight paths. Furthermore, we found that incorporating weather data at lower altitudes slightly improved the

F1 score, confirming that external environmental factors like weather play a role in signal quality fluctuations.

By comparing models trained on different datasets, we observed that Dataset 1, which contains more homogenous data from long-haul flights, outperforms the more generalized Dataset 2, particularly when using the same altitude thresholds. Additionally, models trained with data from a single satellite revealed that predictions using Satellite I5F3 data outperformed those from other satellites, indicating that some satellites might provide a more stable link quality for aircraft.

Overall, this work proves the viability of using machine learning for predicting link quality and highlights the potential for these models to support the automation of satellite handover processes. This approach could significantly reduce signal disruptions, improve in-flight connectivity reliability, and enhance passenger experience and operational efficiency by predicting signal quality in advance and securing the best link. Future work will focus on automating handover decisions based on these predictions, integrating both satellite and weather data for real-time optimization.

ACKNOWLEDGMENT

This work received support from the German Federal Ministry for Economic Affairs and Energy (BMWi) through the German Aeronautical Research Program LuFo - the INTACT project (grant agreement No. 20D2128F).

REFERENCES

- [1] H. A.-D. F. Kokez, "On Terrestrial and Satellite Communications for Telecommunication Future," in *2020 2nd Annual International Conference on Information and Sciences (AiCIS)*. IEEE, 2020, pp. 58–67.
- [2] O. Kodheli and et al, "Satellite Communications in the New Space Era: A Survey and Future Challenges," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 1, pp. 70–109, 2021.
- [3] A. Kumar and et al, "Chapter 1 - introduction to gps/gnss technology," in *GPS and GNSS Technology in Geosciences*, G. p. Petropoulos and P. K. Srivastava, Eds. Elsevier, 2021, pp. 3–20. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/B9780128186176000019>
- [4] G. Singh, "Satellite Communication Constellations as Sources of Alternate PNT (Position, Navigation, and Timing)," *IET Conference Proceedings*, vol. 2022, pp. 176–185, 2023. [Online]. Available: <https://digital-library.theiet.org/doi/abs/10.1049/icp.2023.1383>
- [5] Z. Wang and et al, "A Joint and Dynamic Routing Approach to Connected Vehicles via LEO Constellation Satellite Networks," *Wireless Networks*, pp. 1–13, 2021.
- [6] K. Çelikbilek and et al, "Survey on Optimization Methods for LEO-Satellite-Based Networks with Applications in Future Autonomous Transportation," *Sensors*, vol. 22, no. 4, 2022. [Online]. Available: <https://www.mdpi.com/1424-8220/22/4/1421>
- [7] B. Mafakheri, C. Yan, K. Narayanaswamy, I. Trang, T. Betz, K. Pientka, and L. Goratti, "Edge Intelligence in 5G and Beyond Aeronautical Network with LEO Satellite Backhaul," in *2023 Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit)*. IEEE, 2023, pp. 579–584.
- [8] A. Papa and et al, "Enabling 6G Applications in the Sky: Aeronautical Federation Framework," *IEEE Network*, 2023.
- [9] V. Monzon Baeza and et al, "Enhanced Communications on Satellite-Based IoT Systems to Support Maritime Transportation Services," *Sensors*, vol. 22, no. 17, 2022. [Online]. Available: <https://www.mdpi.com/1424-8220/22/17/6450>
- [10] I. Kovalev, A. Andronov, I. Kartsan, and M. Karaseva, "Technical and Economic Feasibility of the On-Board Location System Application for the Aircraft and Satellite Communication System," in *IOP Conference Series: Materials Science and Engineering*, vol. 862, no. 4. IOP Publishing, 2020, p. 042055.
- [11] L. D. Earley, "Communication in Challenging Environments: Application of LEO/MEO Satellite Constellation to Emerging Aviation Networks," in *2021 Integrated Communications Navigation and Surveillance Conference (ICNS)*, 2021, pp. 1–8.
- [12] M.-J. Jin and J. K. Kim, "Customer Adoption Factors for In-Flight Entertainment and Connectivity," *Research in Transportation Business & Management*, vol. 43, p. 100759, 2022. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2210539521001425>
- [13] [Online]. Available: <https://www.seamlessalliance.com/wp-content/uploads/seamless-ntn.pdf>
- [14] B. Mafakheri, L. Schmidt, A. Prakash, R. Richter, M. Harasic, M. Heindlmaier, P. Hecker, and L. Goratti, "Ai-powered edge-cloud continuum for in-flight entertainment and connectivity," *IEEE Aerospace and Electronic Systems Magazine*, vol. 38, no. 12, pp. 14–27, 2023.
- [15] A. Mohanty and G. Gao, "A Survey of Machine Learning Techniques for Improving Global Navigation Satellite Systems," *EURASIP Journal on Advances in Signal Processing*, vol. 2024, no. 1, p. 73, 2024.
- [16] M. J. Kang, J. H. Lee, and S. H. Chae, "Machine Learning-Based NOMA in LEO Satellite Communication Systems," in *2024 Fifteenth International Conference on Ubiquitous and Future Networks (ICUFN)*, 2024, pp. 448–450.
- [17] K. Biswas and et al, "Enhanced In-Flight Connectivity for Urban Air Mobility via LEO Satellite Networks," 2024. [Online]. Available: <https://arxiv.org/abs/2406.07298>
- [18] M. K. Dahouda, S. Jin, and I. Joe, "Machine Learning-Based Solutions for Handover Decisions in Non-Terrestrial Networks," *Electronics*, vol. 12, no. 8, 2023. [Online]. Available: <https://www.mdpi.com/2079-9292/12/8/1759>
- [19] M. Chen and et al, "Reinforcement Learning Based Signal Quality Aware Handover Scheme for LEO Satellite Communication Networks," in *Human Centered Computing: 5th International Conference, HCC 2019, Čačak, Serbia, August 5–7, 2019, Revised Selected Papers 5*. Springer, 2019, pp. 44–55.
- [20] F. Liu, H. Wu, and X. Xu, "Application of Mobile IP in the Space-Ground Network Based on GEO Satellites," in *Communications and Networking: 11th EAI International Conference, ChinaCom 2016 Chongqing, China, September 24–26, 2016, Proceedings, Part II 11*. Springer, 2018, pp. 421–430.
- [21] S. Mondal and et al, "Handover Prediction for Aircraft Dual Connectivity Using Model Predictive Control," *IEEE Access*, vol. 9, pp. 44 463–44 475, 2021.
- [22] A Reliable, Affordable, and Very Easy IFEC System. Access on October, 2024. [Online]. Available: <https://www.safran-group.com/companies/safran-passenger-innovations>
- [23] [Online]. Available: <https://open-meteo.com/en/docs>
- [24] [Online]. Available: <https://docs.h2o.ai/h2o/latest-stable/h2o-docs/automl.html>