



**HAL**  
open science

# Synthetic Dataset of Maneuvering Low Earth Orbit Satellite Trajectories for AI Analysis

Stéfan Baudier, Santiago Velasco-Forero, Franck Jean, Daniel Brooks, Jesus  
Angulo

► **To cite this version:**

Stéfan Baudier, Santiago Velasco-Forero, Franck Jean, Daniel Brooks, Jesus Angulo. Synthetic Dataset of Maneuvering Low Earth Orbit Satellite Trajectories for AI Analysis. Proceedings of SPAICE2024: The First Joint European Space Agency / IAA Conference on AI in and for Space, Sep 2024, Oxfordshire, United Kingdom. hal-04755327

**HAL Id: hal-04755327**

**<https://hal.science/hal-04755327v1>**

Submitted on 27 Oct 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Synthetic Dataset of Maneuvering Low Earth Orbit Satellite Trajectories for AI Analysis

Stéfan Baudier<sup>\*1,2</sup>, Santiago Velasco-Forero<sup>1</sup>, Franck Jean<sup>3</sup>, Daniel Brooks<sup>2</sup>, Jesus Angulo<sup>1</sup>

<sup>1</sup>Mines Paris, PSL University, Centre for Mathematical Morphology (CMM), Fontainebleau, France

<sup>2</sup>Thales LAS France, Advanced Radar Concept (ARC), Limours, France

<sup>3</sup>Thales LAS France, Ultra high frequency Multifonction Radar (UMR), Limours, France

The characterization of satellite behavior is of paramount importance in Space Surveillance Awareness. It involves modeling complex patterns from large operational databases, making AI tools well-suited to handle this use case. Despite existing contributions, no database is dedicated to Pattern-of-Life study in the Low Earth Orbit regime. In this context, we provide a dataset of satellite trajectories, focusing on station-keeping issues. The proposed database contains generated trajectories based on real data. Our experiments on the provided dataset and real trajectories tend to verify the representativity of the data and highlight the complexity of the Pattern-of-Life related tasks.

## 1 Introduction

Over the last decades, the increasing number of satellites has led to a higher complexity for Space Surveillance Awareness (SSA) systems. The Low Earth Orbit (LEO) regime is the playground of many private interests, especially telecom companies like Starlink dominating such orbits. Better reactivity is necessary to address this growth, such as when dealing with orbit determination and maneuver detection tasks. Exploiting SSA data relies on human intervention, which is time-consuming and error-prone. Further, SSA generates a lot of data due to the nature of the system: objects rotating around the Earth and repeatedly passing in front of SSA sensors, as radar systems.

At the same time, AI algorithms prove their ability to treat different use cases dealing with time series, such as anomaly detection [1], classification [2], or prediction [3]. Some AI tools have already been used for maneuver-related tasks [4]. However, these contributions rely on eclectic datasets based on different spacecraft, pre-processing, and data nature, such as mean elements or ephemeris. We propose a new dataset to unify the results of state-of-the-art methods.

A dataset for Geosynchronous Orbit (GEO) satellites already exists [5]. Despite the GEO regime being a key challenge for SSA, it does not share the characteristics of the LEO regime, which is also worthy of interest. Whether GEO satellites have the same orbital period matching with the Earth's rotation rate (24 hours), leading to returning to the same position in the sky after each sidereal day, the satellites in LEO have different orbital periods (from 1h30 to 2h00) leading to less remarkable natural trajectory pattern. Furthermore, LEO satellites are submitted to the same perturbations as GEO satellites but at different scales of intensities, leading, for example, to a high altitude decrease due to atmospheric drag. It drastically changes the maneuvering strategy, called station-keeping, correcting these drifts. Then, the Pattern-of-Life (PoL), the repeatable patterns composed by the natural satellite cinematic and the station-keeping strategy, is very different from the satellites in the GEO regime, raising the need for a new dataset.

We introduce a new dataset composed of synthetic data. The data generation is based on Orekit [6], a well-known core layer for space flight dynamics applications. While this generation aims to generalize the proposed PoL to all LEO satellites, it also focuses on the case of telecom satellites through the Starlink example.

Based on this dataset strongly focused on PoL description, we conduct two related studies with AI tools. The first aims to detect maneuvers in a portion of the satellite trajectories. We test the model on real data to assess the synthetic maneuvers' representativity. The second aims to classify the maneuvers as part of station-keeping or not.

This paper covers i) the problem description focusing on the PoL, ii) the dataset presentation and the underlying assumptions, and iii) the results of the two studies related to maneuver characterization with AI.

<sup>\*</sup>Corresponding author. E-Mail: stefan.baudier@minesparis.psl.eu

## 2 Problem

The PoL of satellites raises great issues in SSA. While the natural trajectory can be well modeled numerically, modeling station-keeping is more challenging because it can be very different from one satellite to another.

The main specificity of the LEO, regarding other orbital regimes, is the preponderance of the drag among the other perturbations. The atmosphere friction depends on characteristics that can differ from one satellite to another, as described in the drag force equation:

$$F_d = \frac{1}{2} \rho u^2 c_d A, \quad (1)$$

where  $\rho$  the mass density of the atmosphere,  $u$  the flow velocity relative to the object,  $c_d$  the drag coefficient and  $A$  the cross-sectional area. These parameters depend on the satellite's altitude, shape, size, and attitude. These parameters evolve with time, leading to high volatility in the drag force value. The same phenomenon occurs in the other perturbations at a smaller intensity. It leads to a personalized calibration of station-keeping maneuvers for each satellite.

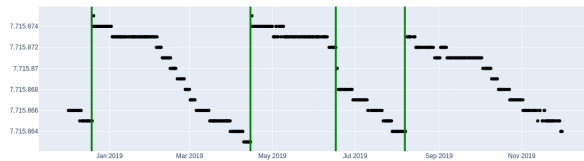
The intensity of the satellites' maneuvers depends on the propulsion system used. Electric propulsors generate low-thrust maneuvers but enable finer orbit control and a longer delta-V budget; meanwhile, using chemical propulsors leads to a smaller delta-V budget but greater maneuver intensity capacity. Further, the station-keeping strategy can be different from one satellite to another. The satellite's tolerance for drifting from its initial orbit relies on the specificity of its mission, for example, being included in a satellite constellation.

We illustrate the difference in station-keeping characteristics with TLE data provided by the USSPACECOM via the Spacetrack website [7]. The trajectories of Jason 3 and Cryosat 2 in Fig. 1 show that the intensity and periodicity of repeatable maneuvers differ from one satellite to another. Moreover, there is a maneuver that does not follow the same pattern as others in the Cryosat 2 trajectory. Indeed, while station-keeping maneuvers represent most satellite maneuvers, spacecraft can perform other maneuvers, for example, changing their mission or avoiding a collision.

## 3 Data description

### 3.1 Data generation

This simulation is based on the Orekit software, which is well-known for its application in space mechanics projects. The algorithm generates scenarios of satellite



(a) Jason 3



(b) Cryosat 2

**Figure 1:** Evolution of the mean semi-major axis for two different satellites. The vertical lines represent the maneuvers.

trajectories, including the main perturbations [8] of LEO satellites: the perturbed gravity field, the drag, the third bodies attraction (Sun and Moon), the solar radiation, and the relativity. These natural perturbations depend on satellite characteristics. We also add non-natural perturbations through maneuvers. We explain the core principles underlying satellite characteristics generation and the station-keeping process.

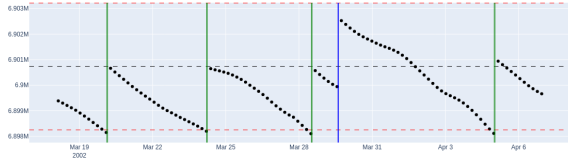
#### 3.1.1 Satellite characteristics

One notes that only three satellite characteristics mainly impact the trajectory: the cross-section, drag, and radiation coefficient. Their values mostly rely on the satellite's shape and attitude. The first can be easily considered constant, but the second evolves. We decide to represent the attitude evolution by only a cross-section evolution. The drag coefficient and radiation coefficient are basically around 2.2. We draw once for all their values for each satellite around this nominal value. The cross-section parameter is generated based on the DISCOSweb database [9]. Minimum and maximum values are drawn based on the available data on LEO satellites. Then, the actual cross-section is initialized and updated randomly based on the previous limits.

We also generate Starlink-like scenarios using Starlink characteristics and orbit values from the DISCOSweb database.

#### 3.1.2 Station-keeping

Following previous studies [10, 11], where station-keeping relied on mean element computation, we opt

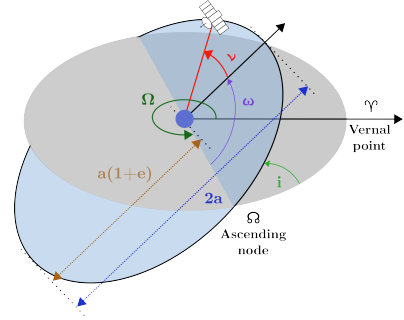


**Figure 2:** Example of the mean semi-major axis correction. The horizontal lines represent the nominal (black) and limits (red) values. The vertical lines represent the station-keeping (green) and random (blue) maneuvers.

for a similar approach, activating maneuvers through a servo system driven by mean element values. Each scenario is defined by a nominal value and a threshold defining the limits the satellite mean parameters can not cross. They are computed based on the orbit drift intensities caused by the perturbations over a variable period of a few days. It aims to generate faithful maneuvers correlated with the orbit. Further, these limits set nominal intensities of maneuvers and aim to simulate the variability in propulsion technology and operational mission described in 2. We implemented a station-keeping strategy correcting the mean semi-major axis and inclination values. In Fig. 2, one observes that when the mean semi-major axis crosses the down limit, a maneuver corrects the mean value to set it to the nominal one. The mean elements are computed using the same theory as TLE, especially based on the SGP4 model fitted on a 3-day ephemeris window.

### 3.1.3 Other maneuvers

Additionally, we introduce two distinct types of maneuvers that deviate from this typical behavior. These anomalous maneuvers constitute the class we aim to detect in 4.3. The first type resembles station-keeping maneuvers in intensity and modifies exclusively the semi-major axis or the inclination. However, they do not adhere to the servo system strategy. On the other hand, the second type exhibits even more significant deviations in intensity compared to the station-keeping distribution. It does not target any specific aspect of the trajectory for alteration. All these maneuvers are composed of impulsive burns, whether Hohmann transfer for semi-major axis change, nodes maneuvers for inclination change, or at random anomaly and orientation for another kind of maneuver. Some scenarios do not include maneuvers; their trajectory only relies on natural forces.



**Figure 3:** Representation of Keplerian parameters:  $a$  is the semi-major axis,  $i$  the inclination,  $e$  the eccentricity,  $\omega$  the argument of perigee,  $\Omega$  the longitude of the ascending node, and  $v$  the true anomaly.

## 3.2 Data format

The dataset gathers 400 synthetic satellite scenarios, with 150 Starlink look-alike trajectories<sup>1</sup>. This diversity of scenarios enables to cover uniformly the LEO domain and to focus on its overcrowded parts. The cinematic data are in the osculating equinoctial representation [12], a variant of Keplerian orbit parameters (Fig. 3). This data representation is robust to singularities, such as when the eccentricity or the inclination equals zero. This kind of property is well-suited for a deep-learning approach. For maneuvering scenarios, another file containing information about the maneuvers is provided.

Considering that the station-keeping process differs for each satellite, we aim to challenge models on scenario change resilience. To achieve this, we partition the dataset into train-test sets, ensuring that each scenario is exclusively assigned to one set in a 70-30 percent manner.

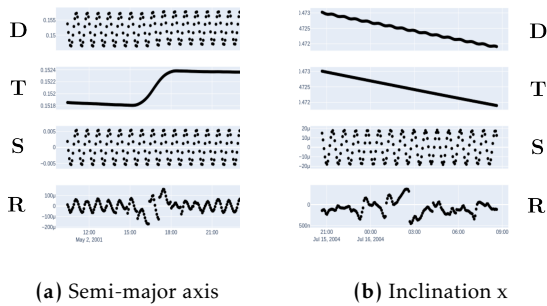
## 4 Results

This section proposes a quick exploration of the data through maneuver detection, allowing to check the data generation's representativity and assess the complexity related to Pattern-of-Life tasks.

### 4.1 Preprocessing and data visualization

We use the Seasonal-Trend decomposition procedure based on Loess (STL) [13], a filter that decomposes time series with periodicity by applying a Loess sequence smoother. As described in Fig. 4, this filter yields three distinct components: a seasonal component capturing periodic evolution, a trend component

<sup>1</sup><https://github.com/StefanBaudier/Synthetic-Dataset-of-Maneuvering-Low-Earth-Orbit-Satellite-Trajectories-for-AI-Analysis>



**Figure 4:** STL decomposition of two equinoctial parameters, during a semi-major axis (a) and an inclination (b) maneuver. The top charts are the raw data (D). The three following are the components of the STL decomposition: Trend (T), Seasonal (S), and Remainders (R).

delineating shifts between successive periods, and a remainder component encompassing unexplained information. Subsequently, we derive various features by computing the differences between consecutive timestamps. One notes that the impact of the maneuvers is clearer on the trend and residual graphs than on the raw ephemeris. It encourages the usage of such methods to treat ephemeris data.

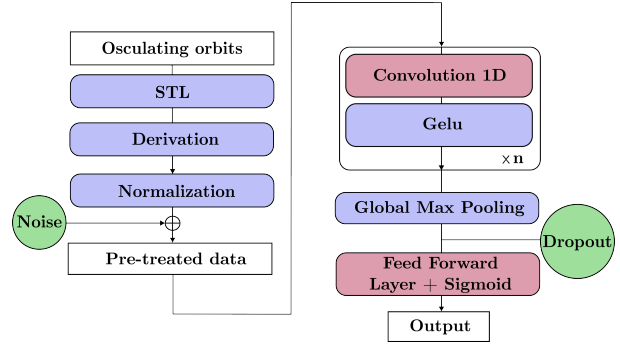
## 4.2 Maneuver detection

The objective is to detect maneuvers within subsequences of 12-hour trajectories using a standard supervised classification paradigm. We train and test the model on the generated dataset and test it on a real dataset. The real data come from the IDS (International Doris System) [14] providing DORIS (Doppler Orbitography and Radiopositioning Integrated by Satellite) data and derived products. This dataset contains 18 real satellites’ orbit positioning at a centimeter precision.

The normal and abnormal classes are balanced for both real and synthetic datasets, i.e., there are as many windows without maneuver as windows containing at least one burst.

The used pipeline is described in Fig. 5. We derive and standardize the data to ensure uniformity, calculating each satellite’s features’ mean and standard deviation. We use a standard AI model, e.g., a convolutional neural network of  $n$  Convolutional-Gelu layers. The kernel size of each layer is set to obtain a perceptive field equal to the maximum Keplerian period in LEO, i.e., 120 minutes. The channel number is increased by 1.2 at each layer.

We test several configurations, changing  $n$  from 1 to 4. The best model contains 2 stacks of convolutional layers. We will use this architecture for the following



**Figure 5:** Pre-treatment and CNN architecture. The blue boxes are non-learnable functions, the red ones denote learnable, and the green circles signify regularization solutions.

experiments.

Test - Dataset	F1-score	RP9	AUPRC
Synthetic	0.98	0.97	0.97
Real	0.92	0.93	0.88

**Table 1:** Results of the maneuver detection with CNN

The F1-score, precision at the recall of 0.9 (RP9), and the Area Under the Precision-Recall Curve (AUPRC) are shown in Tab. 1. Even if the results on real data are slightly lower than those on synthetic ones, the classification performance is still satisfactory. These good results seem to confirm the capacity of an AI model to treat such a task on real data using the generated database as a training base. It tends to prove the good maneuver representativity of the synthetic dataset.

Further, we test our model, returning progressively to the original data balance, i.e., with natural windows representing around 99 percent of the data set. Because this unbalancing only increases the number of natural windows (class 0), it only impacts the precision value. One observes in Tab. 2 that the precision drops significantly for both datasets. We assume the higher decrease for the real dataset is due to unlabeled events impacting the trajectory (probably maneuvers) and minor inaccuracies of synthetic orbit perturbations emphasized by this unbalanced paradigm.

We suppose the architecture needs to model the orbit perturbations better.

## 4.3 Out-of-PoL detection

The objective is to detect maneuvers out of the PoL, represented by the maneuvers out of the station-keeping process described in 3.1.3. We follow the

Test - Dataset	1	2	3	4	5
Synthetic	0.98	0.92	0.86	0.81	0.76
Real	0.91	0.46	0.30	0.23	0.18

**Table 2:** Precision according to the dataset balance uniformly represented from 1 to 5. 1 stands for a balanced dataset, 5 stands for the original data balance.

same paradigm with the same model as in 4.2, a 12-hour supervised classification with the same parameters. The real data do not contain enough labels to distinguish the PoL maneuvers from the others. We test this approach on the synthetic data only.

Test - Dataset	F1-score	RP9	AUPRC
Synthetic	0.51	0.0	0.31

**Table 3:** Results of the out-PoL maneuver classification using CNN

The poor results of this experiment, shown in Tab. 3, highlight that the differences among each satellite station-keeping process make analyzing the Pattern-of-Life with AI difficult. We suppose more information on the past PoL is needed to perform such a task.

## 5 Discussion

This article describes a novel dataset dedicated to Low Earth Orbit satellite trajectories for AI analysis. This synthetic dataset is focused on Pattern-of-Life modelization and could be used for related analysis. The dataset’s representativity is based on a real database foundation and confirmed through experiments on the presented dataset and real data. Further investigations emphasize the difficulty of dealing with such complex patterns, both for the perturbations and for the station-keeping process modelization. The Deep Matrix Profile introduced in [15] and applied to the provided data shows promising results by exploiting the satellite’s past trajectories to model the Pattern-of-Life. It enables to improve the classification of the natural orbit parts and better distinguish random from station-keeping maneuvers. This study illustrates that AI analysis is well-suited to the complex tasks related to the proposed dataset.

## Acknowledgment

This work was supported by the French Defense Innovation Agency (Cifre-Défense 2021/0003/AID).

This work was granted access to the HPC resources of IDRIS under the allocation 2024-AD011014111R1 made by GENCI.

For our research, we are using information from ESA DISCOS (Database and Information System Characterising Objects in Space), a single-source reference for launch information, object registration details, launch vehicle descriptions, as well as spacecraft information for all trackable, unclassified objects. We acknowledge ESA’s efforts to maintain and operate this database with its APIs.

## References

- Blázquez-García, A., Conde, A., Mori, U. & Lozano, J. A. A Review on Outlier/Anomaly Detection in Time Series Data. *ACM Computing Surveys* **54**, 1–33. ISSN: 0360-0300, 1557-7341. (2022) (Apr. 2022).
- Ismail Fawaz, H., Forestier, G., Weber, J., Idoumghar, L. & Muller, P.-A. Deep Learning for Time Series Classification: A Review. *Data Mining and Knowledge Discovery* **33**, 917–963. ISSN: 1573-756X. (2023) (July 2019).
- Han, Z., Zhao, J., Leung, H., Ma, K. F. & Wang, W. A Review of Deep Learning Models for Time Series Prediction. *IEEE Sensors Journal* **21**, 7833–7848. ISSN: 1530-437X, 1558-1748, 2379-9153. (2024) (Mar. 2021).
- Shen, D. *et al.* *Methods of Machine Learning for Space Object Pattern Classification in 2019 IEEE National Aerospace and Electronics Conference (NAECON)* (IEEE, Dayton, OH, USA, July 2019), 565–572. ISBN: 978-1-72811-416-3. (2022).
- Siew, P. M. *et al.* AI SSA Challenge Problem: Satellite Pattern-of-Life Characterization Dataset and Benchmark Suite (2023).
- Maisonobe, L., Maisonobe, L., Pommier-Maurussane, V. & Pommier, V. Orekit : An Open-source Library for Operational Flight Dynamics Applications.
- Space-Track.Org* <https://www.space-track.org>. (2024).
- Vallado, D. A. *Fundamentals of Astrodynamics and Applications (Vol. 12)* Springer Science & Business Media (Space Technology Library, 2007).
- DISCOSweb* <https://discosweb.esoc.esa.int/>. (2024).
- Schaub, H., Vadali, S. R., Junkins, J. L. & Alfriend, K. T. Spacecraft Formation Flying Control Using Mean Orbit Elements. *The Journal of the Astronautical Sciences* **48**, 69–87. ISSN: 0021-9142, 2195-0571. (2023) (Mar. 2000).
- Garulli, A., Giannitrapani, A., Leomanni, M. & Scortecci, F. Autonomous Station Keeping for LEO Missions with a Hybrid Continuous/Impulsive Electric Propulsion System. *32nd International Electric Propulsion Conference* (2011).
- Broucke, R. A. & Cefola, P. J. On the Equinoctial Orbit Elements. *Celestial Mechanics* **5**, 303–310. ISSN: 0008-8714, 1572-9478. (2024) (May 1972).
- Cleveland, R. B., Cleveland, W. S. & Terpenning, I. STL: A Seasonal-Trend Decomposition Procedure Based on Loess. *International journal of biometeorology* **6**, 3. ISSN: 0282423X. (2023) (1990).

14. Willis, P. *et al.* *The International DORIS Service (IDS): Recent Developments in Preparation for ITRF2013 in IAG 150 Years* (eds Rizos, C. & Willis, P.) (Springer International Publishing, Cham, 2016), 631–640. ISBN: 978-3-319-30895-1.
15. Baudier, S., Velasco-Forero, S., Jean, F., Brooks, D. & Angulo, J. Deep Matrix Profile for Maneuver Classification in Low Earth Orbit Satellite Trajectories.