Data Description
To develop AI/ML based models to predict time-varying patterns of the error build up between uploaded and modelled values of both satellite clock and ephemeris parameters of navigation satellites

September 26, 2025

---

1. Introduction

In modern navigation, GNSS satellites are the digital beacons that guide everything from smartphones to aircraft. However, their precision is intrinsically limited by two primary types of errors transmitted from their respective ground control segments: **ephemeris errors** and **clock errors**.

- **Ephemeris errors** refer to the difference between a satellite's predicted (broadcast) orbital position and its true, precise position in space. These errors manifest as discrepancies in the satellite's reported X, Y, and Z coordinates.
- **Clock errors** are the discrepancies between the satellite's on-board atomic clock (which is fundamental for precise timing) and the highly stable, common GNSS time scale maintained by the ground segment.

Both types of errors directly impact the accuracy of a user's computed position and velocity. By statistically analyzing these errors, we can gain a profound understanding of the overall health, reliability, and specific error characteristics of each GNSS constellation. This analysis also helps in identifying potential common error sources that might affect multiple systems, which is crucial for the development and performance assessment of advanced multi-GNSS positioning and integrity monitoring algorithms.

Data Details

It contains a seven-day dataset containing recorded clock and ephemeris errors between uploaded and modeled values from GNSS satellites in both GEO (1 file) and MEO orbit

(2 files). The data is generated at irregular sampling rate. The raw data has already been pre-processed into a readily consumable format: separate Excel files for each orbit (GEO and MEO). This dataset is to be used for training and making appropriate model. The data may contain outliers that require suitable treatment.

Each Excel file is expected to contain time-tagged information for:

- **Time**: The precise timestamp for each error epoch.
- **X_Error**: Position error along the Earth-Centered, Earth-Fixed (ECEF) X-axis, expressed in meters.
- **Y_Error**: Position error along the ECEF Y-axis, expressed in meters.
- **Z_Error**: Position error along the ECEF Z-axis, expressed in meters.
- **Clock_Error**: Satellite clock error, typically expressed in meters.

Clock Error Conversion to Meters: For consistent comparison with positional errors, the satellite clock error (e_{clock}), which is typically in seconds, is converted to an equivalent distance in meters ($e_{clock,m}$) by multiplying it by the speed of light ($c \approx 3 \times 10^8$ m/s):

$$e_{clock,m} = e_{clock} \times c$$

This conversion allows all error components to be analyzed and visualized in the same units (meters).

Evaluation Criteria

The developed model will be used to predict the eighth day errors based on the past seven days data available to the users. The error between the predicted and the ground truth (available with the evaluator) will be compared to the normal disritution. The resemblance to the normal distribution indicates that the developed model has removed the systematic errors and residual errors are random. The Shapiro-Wilk test is a statistical hypothesis test that evaluates whether a sample comes from a normally distributed population. It is generally considered one of the most powerful normality tests. It will be used to evaluate the performance.
