Low-Earth Orbit Prediction Accuracy Review of Modern Empirical Atmospheric Models and Space Weather Data Sources

Pol Mesalles-Ripoll, Roman Rositani, Matthew Duncan
SpaceNav

ABSTRACT

We introduce an end-to-end framework to analyze and evaluate the prediction accuracy of different atmospheric models for satellites in low-Earth orbit. We focus our analysis on the NRLMSISE-00, NRLMSIS 2.0, and JB2008 atmosphere models.

To evaluate the performance of each atmospheric input, we cannot simply compare generated ephemerides with the same initial conditions and different atmosphere models; parameters such as the drag coefficient must also be estimated with the same model before generating new predictions. By processing GNSS tracking data over a timespan of 6 months for two satellites in different orbit regimes, we can generate a large dataset of definitive ephemeris data. Using the last definitive state at the end of each orbit determination arc, we generate a prediction using the solved-for C_D value and the most up-to-date space weather data at the time.

Overlapping predictive and definitive states at the same epochs to determine the prediction errors at 24, 48, and 72 hours, we show that JB2008 slightly outperforms the MSIS models during geomagnetic storms, reducing prediction errors by half in some isolated cases. However, outside of storm periods, the empirical radial–in-track–cross-track uncertainties generated from our sample data with MSIS are smaller than the equivalent results from Jacchia–Bowman: at 400 km, the differences in error are less than 20 %, but at 700 km the errors double. We also show that for this application, the differences between the newer NRLMSIS 2.0 and the classic NRLMSISE-00 are negligible; the lower thermospheric densities result in higher C_D estimates, but the prediction errors are essentially identical for both.

Finally, we analyze the effect of space weather data independently from the atmospheric model. Using the same framework, we compare results with NRLMSISE-00 using space weather data from NOAA's Space Weather Prediction Center and from Space Environment Technologies. We do not see any significant differences between the two, but when we compare the predictive data to the definitive data of either source, prediction errors are on average reduced by 30 % after 48 h of prediction. This shows that if we can decrease the error in the space weather predictions, owner/operator ephemerides will see a major boost in accuracy.

1. INTRODUCTION

As the catalog of space objects in low-Earth orbit (LEO) grows and we witness an increasing number of conjunction events, owner/operators find themselves in need of generating accurate orbit predictions for the purposes of space situational awareness.

Although the 18$^{\text{th}}$ and 19$^{\text{th}}$ Space Defense Squadrons of the U.S. Space Force already generate ballistic trajectories for all trackable objects for operational conjunction assessment, not all operators have access to these ephemerides, and in many cases they might require more frequent predictions that they can use for planning, tasking, and scheduling. Additionally, for all maneuverable satellites, it is of the utmost importance that they share predicted ephemerides with the rest of the space community to communicate their intent to maneuver and allow other operators to plan accordingly.

Many satellite operators already leverage on-board GNSS receivers to obtain definitive solutions and produce their own orbit predictions. In LEO, uncertainty in these predictive ephemerides is highly influenced by drag modeling, but there is a lack of resources detailing which atmospheric models are best suited for the application of generating real-time predicts. By real time, we refer to orbit predictions being generated with the most up-to-date space weather data at a given epoch, which is typically the latest definitive state produced by an orbit determination (OD) process.

Our goal with this paper is to start generating comparative data of the prediction accuracy of different atmospheric models. In particular, we will be comparing the widely used NRLMSISE-00 from the Naval Research Laboratory, their more modern NRLMSIS 2.0 version, as well as the Jacchia–Bowman 2008 (JB2008) atmosphere. For the first two models, we employ the daily short-term space weather predictions from the U.S. Air Force available through NOAA's Space Weather Prediction Center (SWPC), while for the latter we use the data generated by Space Environment Technologies (SET), as we need an expanded set of indices. We also independently analyze the effects of each of these space weather data sources with predictive and definitive data.

For this study, we have selected two satellites that operate in different orbit regimes: the first one is NASA's Global Precipitation Measurement (GPM) satellite, which is in a very low altitude, 400 km, 65° orbit, while the second one—an Earth-Observing satellite—is flying in a much higher, 700 km, 98° orbit and has a slightly higher area-to-mass ratio. Additional details of each satellite can be found in Table 1.

Table 1: Orbital parameters of the satellites in the study.

Satellite	Apogee × Perigee Altitude*	Inclination	Area-to-Mass Ratio†
GPM	405 × 391 km	65°	0.005 m^2/kg
Earth-Observing Satellite	703 × 702 km	98°	0.009 m^2/kg

* Average mean elements relative to the Earth's equatorial radius.
† Average A_D/m, where A_D is the effective drag area and m is the mass.

The accuracy review will be conducted by first processing 6 months of GPS measurements through an orbit determination process to generate definitive ephemerides and then propagating a predictive ephemeris from the definitive states at the end of each arc. After aggregating the results, we end up with multiple sets of predicted states overlapping with the definitive data, allowing us to compare the position differences in the radial–in-track–cross-track (RIC) frame for each model at different prediction lengths.

2. ATMOSPHERIC MODELS

In order to generate ample data for comparisons, three different modern empirical atmospheric models are considered:

(a) NRLMSISE-00,

(b) NRLMSIS 2.0, and

(c) JB2008.

Each of these models has been implemented in SpaceNav's *State Propgation Service* (SPS) and is utilized for both orbit determination and prediction. The following sections describe the particulars of each atmosphere model in more detail. The first section of results, §5.1, will focus on the performance of each atmosphere model.

2.1 The NRLMSISE-00 Model

The inputs to the Naval Research Laboratory (NRL) mass spectrometer and incoherent scatter radar (MSIS) extended to the exosphere (E) from the year 2000, or NRLMSISE-00 for short, are as follows [12]:

$$\rho = f\left(\text{DoY}, t, \text{Altitude}, \text{Latitude}, \text{Longitude}, \overbrace{F_{10.7}, \bar{F}_{10.7}, a_p, A_p}^{\text{Space Weather}}, \text{LST}\right)$$

- DoY is the current day of the year and t are the seconds in the day,
- daily $F_{10.7}$ for the previous day and 81-day $\bar{F}_{10.7}$ average,
- daily A_p (average), as well as 3-hour a_p index for 0, 3, 6, 9, 12–33 (mean), and 36–57 (mean) hours before current time, and finally

- LST, or local solar time, which is interpreted as the apparent local time from the Sun's instantaneous position; it is worth noting that other implementations may use directly the mean local time computed directly as LST = $t/3600 + \text{Longitude}/15$ [hours].

NRLMSISE-00 has become the standard atmosphere for satellite drag modeling over the last few years, in part thanks to its widespread availability. The original source code is available in Fortran [16], but a translation to C is available from Prof. Brodowski [3], and Python bindings exist too.

2.2 The New NRLMSIS 2.0 Model

In [4], the new NRLMSIS 2.0 model from 2020 is introduced as follows:

> "NRLMSIS®2.0 is an empirical atmospheric model that extends from the ground to the exobase and describes the average observed behavior of temperature, eight species densities, and mass density via a parametric analytic formulation."

The new model takes in the same inputs as NRLMSISE-00 but drops the local solar time requirement. Internally, it computes mean local solar time from time of day and longitude (i.e., LST = $t/3600 + \text{Longitude}/15$ [hours]). Fortran source code is available [15]. However, it is worth noting that unlike NRLMSISE-00, NRLMSIS 2.0 is distributed with a license agreement that prohibits usage of the software for any non-academic, commercial purposes without first obtaining written authorization from the Naval Research Laboratory.

Most of the changes focused on altitudes below 200 km. However, the incorporation of new measurements and slight changes to the thermosphere modeling have significant effects on state propagation.

Below we list the main differences or upgrades of the NRLMSIS 2.0 model over its predecessor:

- Species densities are fully coupled to temperature from the mixed region below ~ 70 km altitude to the diffusively separated region above ~ 200 km.
- Atomic oxygen (O) density now extends down to 50 km.
- Geopotential height is now used as the internal vertical coordinate.
- Assimilated extensive new lower and middle atmosphere temperature, O and H data, global average thermospheric mass density from satellite orbits, and validated the model against independent samples of this data.
- Decreased residual biases and uncertainties in the mesosphere and below.
- New model presents a warmer upper troposphere, cooler stratosphere and mesosphere.
- In the thermosphere, the main changes are lower N_2 and O densities.

From a computational performance point of view, the source code from [15] states:

> "For applications in which the horizontal location or time changes with every call (e.g., satellite ephemerides), users may find that NRLMSIS 2.0 is 1/3 to 1/2 as fast as NRLMSISE-00. The reduction in speed is due to the greater number of model parameters in NRLMSIS 2.0, as well its stronger coupling with the lower atmosphere, compared to NRLMSISE-00. However, the model is still very fast: a typical desktop system can process at least 100 000 serial calls per second."

In our testing, propagating an initial state for 1 day with a full set of force models (see Table 4), we have found ephemeris generation with the new NRLMSIS 2.0 atmosphere to be $\sim 8\,\%$ slower than the old NRLMSISE-00 model.

2.3 The Jacchia–Bowman 2008 Model

Jacchia–Bowman 2008, or JB2008 for short, is the latest iteration of atmosphere models based on Jacchia's diffusion equations, following the original work published by Jacchia in 1970 [6], with updates in 1971 [7] and 1977 [8]. Starting with JB2006 [2], Bowman introduced new solar indices to better model the thermospheric density. With JB2008, this is further expanded to also better model the effects of geomagnetic storms [1].

The final form of JB2008 takes in as inputs:

$$\rho = f\left(\text{MJD}, \text{Altitude}, \text{Latitude}, \text{RA}, \overbrace{F_{10}, \bar{F}_{10}, S_{10}, \bar{S}_{10}, M_{10}, \bar{M}_{10}, Y_{10}, \bar{Y}_{10}, dT_c}^{\text{Space Weather}}, \text{Sun Declination and RA}\right)$$

- MJD is the modified Julian date of the current time,
- solar indices F_{10}, S_{10} (EUV), M_{10} (MUV), and Y_{10} (X-ray) for each corresponding wavelength, with $\bar{F}_{10}$, $\bar{S}_{10}$, $\bar{M}_{10}$, and $\bar{Y}_{10}$ being the centered averages,
- the dT_c exospheric temperature change—derived from the Dst index—is the driver of global density changes when modeling the effect of geomagnetic storms, and
- the satellite's and the Sun's right ascensions (RA) are used to compute the local hour angle (LHA), like in all Jacchia models [17].

Fortran source code for JB2008 is publicly available from Space Environment Technologies [13].

JB2008 is the basis for the U.S. Air Force's latest high-accuracy satellite drag model (HASDM), often labeled JBH09, which applies the Dynamic Calibration Atmosphere (DCA) algorithm on top of this reference model to estimate a varying global density field out 3 days from a series of calibration satellites in LEO.

3. SPACE WEATHER DATA SOURCES

As we have seen in the previous section, all atmospheric models require up-to-date space weather inputs. Space weather data comes in many forms, but the most common are:

$F_{10.7}$ The solar radio flux at 10.7 cm is one of the best indicators of solar activity. It correlates well with the sunspot number and solar irradiance records. Daily $F_{10.7}$ data is produced by the Canadian Dominion Radio Astrophysical Observatory and expressed in solar flux units (1 sfu = 1×10^{-22} W m^{-2} Hz^{-1}) [10].

K_p The planetary K-index is an indicator of disturbances in the Earth's magnetic field and is used to characterize the magnitude of geomagnetic storms. NOAA's Space Weather Prediction Center (SWPC) will send geomagnetic alerts when K_p thresholds are exceeded [10].

A_p The K_p index is a logarithmic scale, and thus it is not meaningful to average a set of K-indices. Every 3 hours, the K-value is converted into a linear scale, which results in the a-index. Averaging the 8 daily a-values gives us the A_p index for a certain day.

Daily and historical space weather data can be obtained from different sources:

- $F_{10.7}$ indices can be retrieved from NOAA SWPC [10] or the Canadian Space Weather Forecast Centre [5]. Updated daily.
- K_p indices can be retrieved from NOAA SWPC [10]. Updated every 3 hours. These can be converted into a_p indices with a simple relation if necessary.

For short short-term predictive data, the U.S. Air Force generates 45-day predictions daily for both $F_{10.7}$ and A_p indices, which can then be retrieved through NOAA SWPC [10]. SWPC also distributes a separate 27-day prediction. Note that daily values will not become definitive until, in some cases, months after the day has passed.

In this work, we will discuss two main sources of space weather data:

- *NOAA SWPC* will refer to the combination of daily and historical indices retrieved from SWPC combined with the U.S. Air Force 45-day prediction. With the raw data, an 81-day centered average $\bar{F}_{10.7}$ is computed to use in the atmosphere models.
- *SET* will refer to the space weather files generated by Space Environment Technologies in an hourly cadence. Access to these can be requested to SET through the Unified Data Library. These files are required when using the JB2008 model as they provide the additional S_{10}, M_{10}, and Y_{10} indices, along with the precomputed dT_c. It is worth noting that this file provides 54-day centered averages instead of 81 days.

In the second part of this study (see §5.2), we present results that compare the effects of switching between the NOAA SWPC and SET space weather data sources using a single atmospheric model (NRLMSISE-00). Additionally, we compare the effects of using "definitive" space weather instead of the daily predictions to understand the statistically significant orbit prediction changes caused by errors in space weather predictions.

4. A DESCRIPTION OF THE ANALYSIS FRAMEWORK

In order to evaluate the accuracy of each atmosphere model and space weather source, we have designed a rigorous end-to-end testing framework that can inform us on which model produces more statistically consistent results. In this section, we will cover the details regarding the propagation and filter configuration, as well as detail the procedure for analyzing predictive-definitive overlaps.

4.1 Orbit Determination and Prediction Models

SpaceNav's *Advanced Orbit Determination Service* (AODS) was utilized for processing all tracking data. Both satellites provide GPS NavSOL measurements, i.e., Cartesian position and velocity in a terrestrial frame (ITRF). Table 2 lists the measurement noise values associated with each satellite.

Table 2: Measurement noise values for NavSOL tracking data.

Satellite	Position [m]			Velocity [m/s]		
	Radial	In-Track	Cross-Track	Radial	In-Track	Cross-Track
GPM	5.0	2.75	2.25	0.012	0.008	0.005
Earth-Observing Satellite	5.0	5.0	5.0	0.01	0.01	0.01

For GPM, tracking data is received in 32-hour long files, 3 times per day. For this analysis, we split the data and processed it in 10-hour OD arcs to improve solution quality due to the high variability of drag at GPM's low altitude. For the Earth-observing satellite, since it flies about 300 km higher than GPM, the effects of drag are decreased and we were able to process 24-hour OD arcs from daily high-quality NavSOL measurement files while still preserving accuracy.

The batch least-squares filter, or differential corrector, was used for each OD run. This approach iteratively processes groups of data as opposed to sequentially propagating between measurements and applying updates like the extended or unscented Kalman filters [14]. The batch propagates a reference trajectory numerically, while the covariance is propagated using the state-transition matrix (STM). Further details on the OD setup are shown in Table 3.

Internally, AODS makes use of our in-house propagation software (SPS), which is then also utilized to generate the predictive ephemerides. The latter are always propagated from the last definitive state in each OD arc and employs the latest solved-for C_D for drag and C_R for solar radiation pressure modeling during the entire prediction timespan. Predictive ephemerides are limited to 4 days (96 h) in this study.

Table 3: Orbit determination configuration for each run.

Satellite	GPM	Earth-Observing Satellite
Arc Length	10 h	24 h
Filter Type	Batch Least-Squares	Batch Least-Squares
Solve-Fors*	C_D, C_R	C_D, C_R

* Additional states (other than position and velocity) solved-for during the OD arc, including the drag coefficient (C_D) and the reflectivity coefficient (C_R).

Table 4 details all force models used for propagation during both orbit determination and prediction. These are the same for both satellites. The only differences lie on the individual parameters for each spacecraft: (a) the Earth-observing satellite assumes fixed drag and solar radiation pressure (SRP) effective areas, while (b) the GPM model is significantly more complex. Given the importance of drag modeling for the latter, we make use of a high-fidelity variable area model where we compute the effective drag and SRP areas at each timestep of the propagation from the modeled attitude of the spacecraft, which is derived from an Earth-pointing attitude provider that accounts for a constant pitch offset (an axis aligned with nadir and an axis constrained to point as close as possible to the velocity direction), as well as Sun-tracking solar panels that remain feathered during eclipse and at certain β angles.

Table 4: Special perturbation force models for propagation.

Geopotential	EGM2008, 70×70 (degree and order)
Solid Tides	Enabled
Drag	NRLMSISE-00, NRLMSIS 2.0, JB2008
Third Bodies*	Sun, Moon
Solar Radiation Pressure†	Enabled

* Included third bodies are treated as point-body masses. Their position is determined using the JPL SPICE ephemerides (DE430).

† SRP uses a cannonball model where the radiation pressure is only a function of the satellite-Sun vector, the effective area in that direction, and the reflectivity coefficient.

4.2 Overlap Analysis

Overlaps are simply position and velocity differences between synchronous definitive and predictive data, i.e., prediction errors. Once all OD runs finish, we concatenate all of the definitive data to arrive at one continuous definitive ephemeris spanning the entire analysis time interval. Any duplicate time nodes are removed (only retaining the latest result). As stated, predictive ephemerides begin at the last epoch of each original definitive ephemeris. Because there are a few gaps in the original GPS data for both satellites, each predictive ephemeris roughly covers 4 days of continuous definitive data, though epochs do not always line-up perfectly. To create better, more uniformly formatted overlaps, we define an overlap timespan which covers 4 days (with 5 min buffers) and has a 60 s timestep size. We then find the definitive data that overlaps every predictive file, and we synchronize through Lagrange interpolation of both predictive and defintive datasets. Finally, we take position and velocity differences for each overlapping state at all relative and absolute times.

After all overlaps are collected, outlier overlaps are removed based on a defined "low" and "high" outlier sigma. Outlier removal first collects the computed predictive errors into intervals of length equal to the orbit period. For each of these intervals, the median of all state errors is computed followed by the absolute deviation of each individual ephemeris from the median. This distance divided by the median of all the distances serves as an analogue for standard deviation that is less sensitive to outliers. All ephemerides falling outside the "high" sigma-level threshold based on this metric are removed. Next, the process is repeated using the standard deviation instead of median absolute deviation and using the "low" sigma threshold [11]. For this analysis, we define "low" and "high" thresholds of 5 and 10 sigma, respectively. Outliers can result from unmodeled maneuvers in the data, a sequence of poor GPS measurements

occurring near the end-of-arc, an unmodeled attitude change, etc.

With outliers removed, an empirical uncertainty profile can be generated. We take the standard deviation of differences at each time, utilizing each overlap as an empirical sample.

4.3 Analysis Time Range

To collect statistically significant results, this study processed tracking data for each satellite over 6 months. We have limited the study to 2023, as this year has been so far the most active of Solar Cycle 25. As we approach solar maximum, the density in the thermosphere increases significantly, and the effects of drag mismodeling are more pronounced. For both satellites we used observations covering the exact same timespan which is detailed in Table 5.

Table 5: Timespan for observations used in the study.

Tracking Data Start	Tracking Data End
February 22, 2023	August 22, 2023

The space weather evolution over the analysis time range (February to August 2023) as described by the "definitive" NOAA SWPC and SET space weather sources is presented in Figure 1a for the solar flux (daily $F_{10.7}$ and $\bar{F}_{10.7}$ centered average) and in Figure 1b for the geomagnetic activity (we plot the daily A_p average, but K_p and Dst follow the exact same pattern). Note that since the period of this analysis covers data all the way until August 2023, not all days have final definitive information. Notice the spikes in the geomagnetic activity history; each corresponds to a past geomagnetic storm which we will see has a significant effect in the prediction results.

5. RESULTS

In this section we present the main prediction accuracy results for this study. We have structured the accuracy review in two main sections: the first one focusing on the atmospheric models in prediction space, and a second one exploring the effects of space weather data sources and prediction errors.

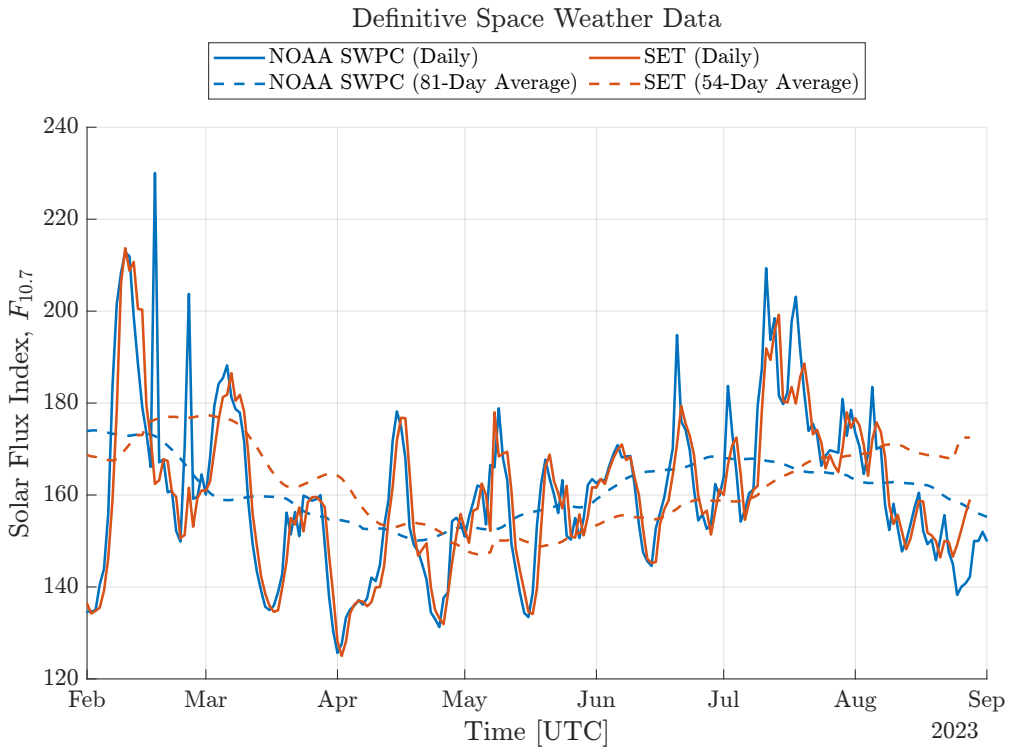
5.1 Atmospheric Models Comparison

5.1.1 Overlap Analysis *Before* Outlier Removal

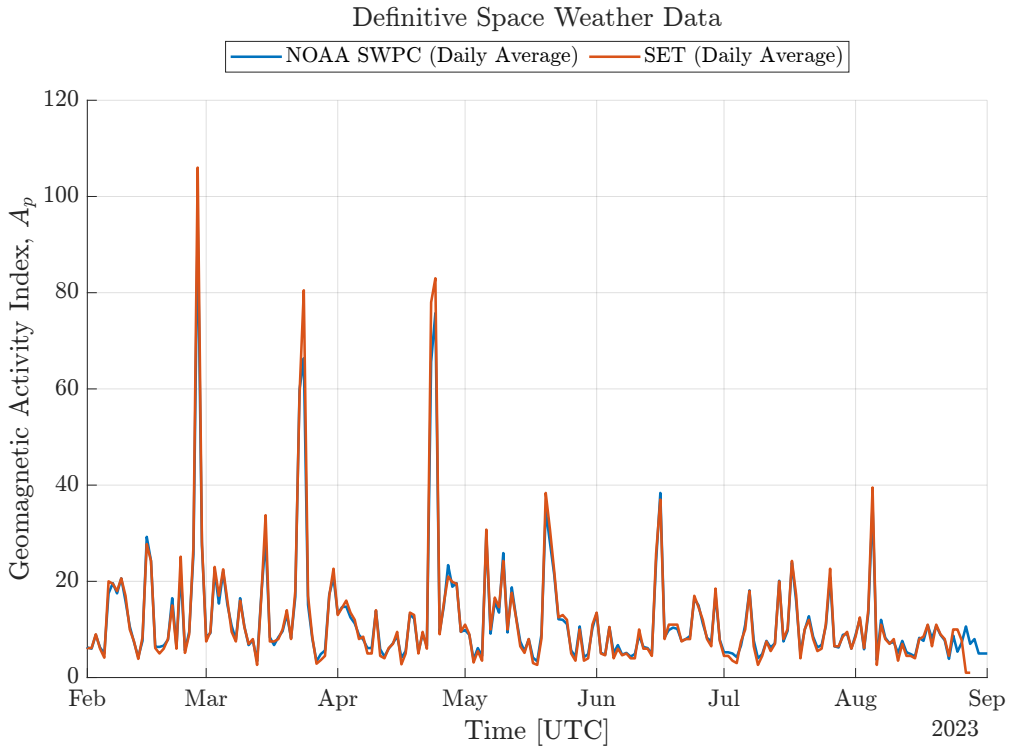
We start by analyzing the predictive-definitive overlaps *without* outlier removal to understand the evolution in time of position differences of the orbit predictions relative to the definitive states (expressed in the RIC frame). To do that, we take individual states from the predicted ephemerides at 24, 48, and 72 hours from the OD epoch and compute the differences with the Lagrange-interpolated definitive state at the same time. Note that we still remove all definitive and predictive overlaps around the times of executed maneuvers to not bias the results.

Figures Figures 2, 3, and 4 plot these RIC differences at discrete prediction times for GPM and each different atmospheric model. In all cases, we see significant spikes in the data—outliers—that all exactly correspond to geomagnetic storms as indicated by high A_p values in Figure 1b. For example, the 72 h prediction line has spikes on March 27, 2023 and April 27, 2023, which perfectly align with the storms on March 24, 2023 and April 24, 2023, respectively. We also see spikes on March 26, 2023 and April 26, 2023 for the 48 h prediction time. This indicates that predictions generated at the time of the storm will suffer from the most inaccuracies. After the storm passes, the prediction errors return to average levels.

The graphs show that most of the time, all models perform roughly the same, with errors typically in the 1000 m range after 24 h of prediction but sometimes reaching more than 10 km in-track and more than 100 m radially during geomagnetic storms. However, it is in the storm periods that we see JB2008 outperforming both MSIS models, with the errors around the March and April 2023 storms sometimes being reduced by half. This makes sense given the focus of the Jacchia–Bowman model on improving density modeling during geomagnetic storms through the Dst-derived dT_c parameter. NRLMSISE-00 and NRLMSIS 2.0 perform nearly identically for GPM's orbit.



(a) Solar flux time history.



(b) Geomagnetic activity time history.

Fig. 1: Definitive space weather data evolution over the study time range.

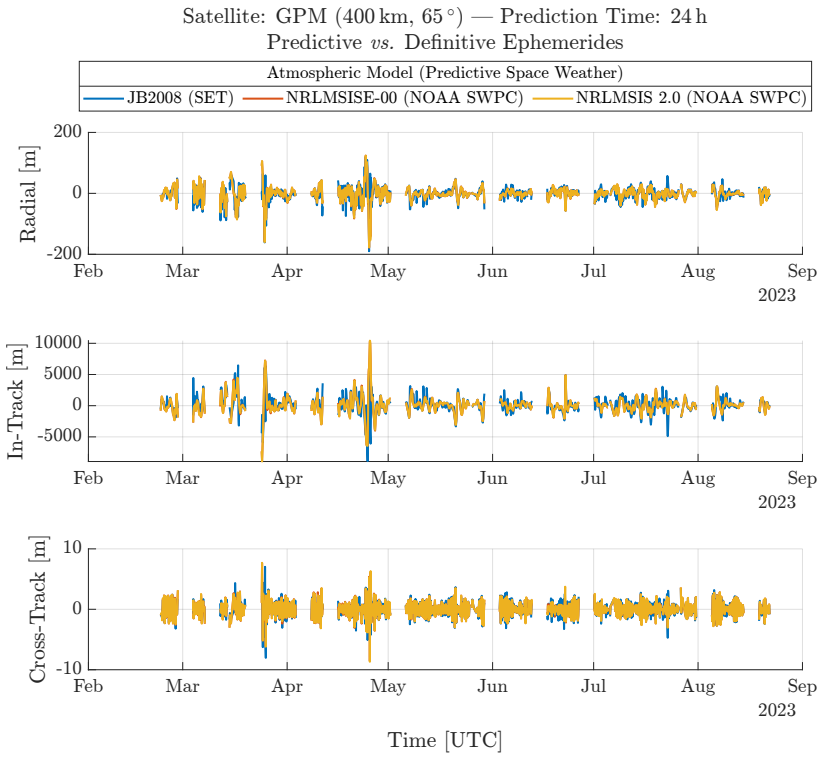


Fig. 2: RIC position differences evolution over time for GPM at 24 h of prediction time after the OD epoch for different atmosphere models using predictive space weather data.

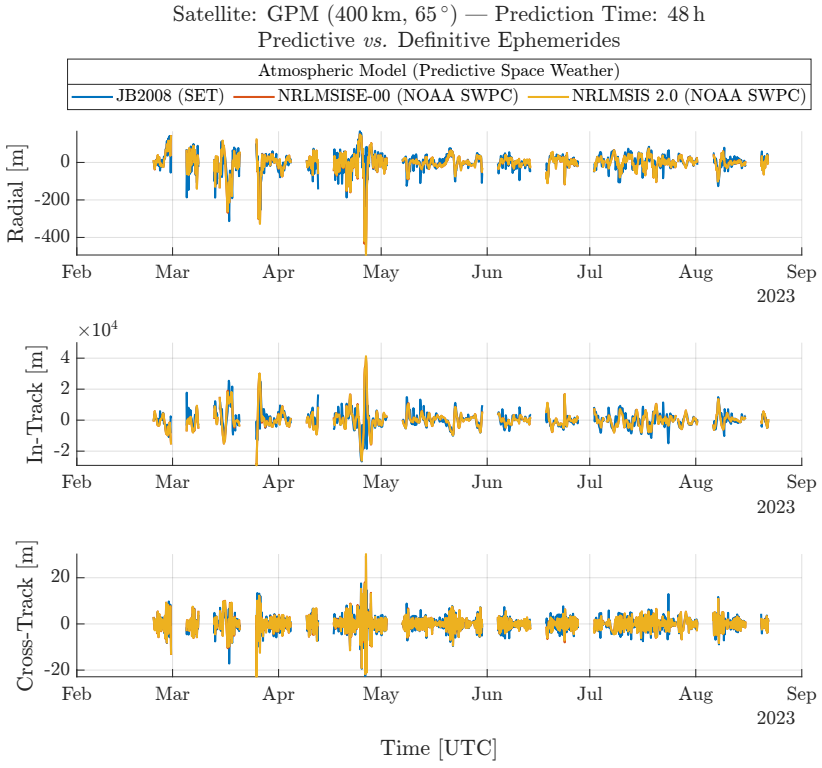


Fig. 3: RIC position differences evolution over time for GPM at 48 h of prediction time after the OD epoch for different atmosphere models using predictive space weather data.

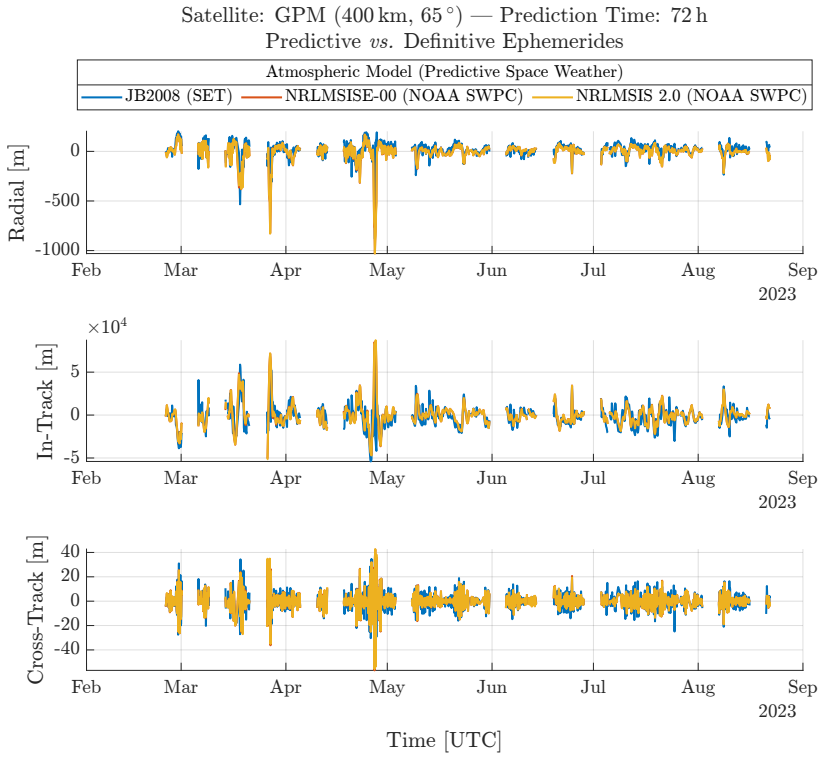


Fig. 4: RIC position differences evolution over time for GPM at 72 h of prediction time after the OD epoch for different atmosphere models using predictive space weather data.

5.1.2 Overlap Analysis *After* Outlier Removal

Next, we complete the overlap process described in §4.2. The total number of overlaps before and after outlier removal are displayed in Table 6 for each case and satellite. Note that this is the only section that will include results from the Earth-observing satellite, as the longer OD arcs resulted in a number of overlaps we deemed too low for any other analysis.

Table 6: Total number of predictive-definitive overlaps.

Satellite	Model	Number of Overlaps	
		Total	Post-Outlier Removal
GPM	NRLMSISE-00	1347	698
	NRLMSIS 2.0	1347	698
	JB2008	1347	689
Earth-Observing Satellite	NRLMSISE-00	161	98
	JB2008	161	96

Figures 5, 6, and 7 show the resulting overlap plots for GPM and each atmospheric model. In these plots, each line represents one predictive-definitive interpolated overlap spanning up to 4 days. Equivalent plots for the higher-altitude Earth-observing satellite are presented in Figures 8 and 9.

From the prediction error overlaps, we can then compute an empirical covariance. This is a realistic covariance since it exactly represents the expected uncertainties at different prediction lengths based on a statistical analysis. Figure 10 plots the empirical error growth for GPM, while Figure 11 includes the counterpart for the Earth-observing satellite. Finally, Table 7 summarizes all the results for all atmosphere models and satellites.

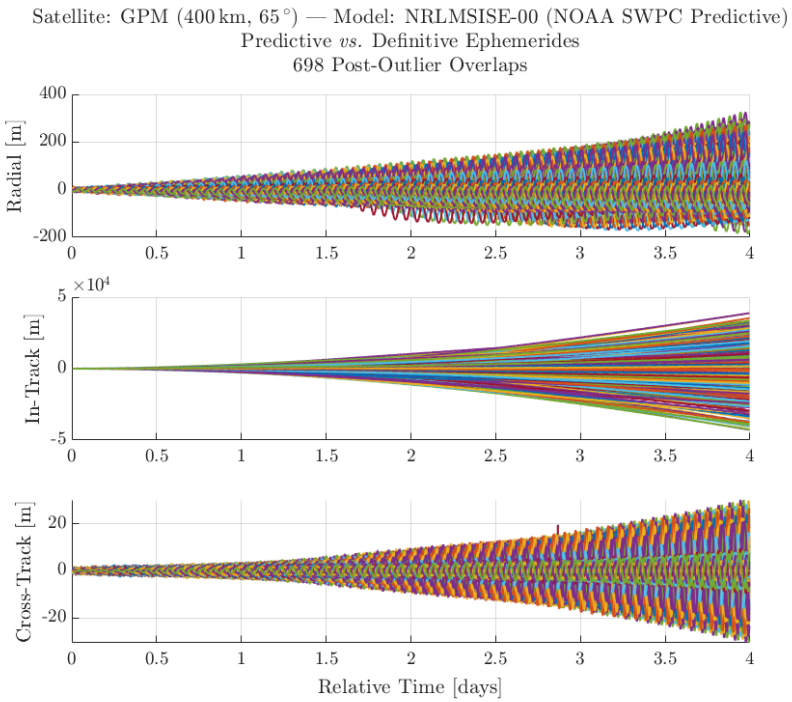


Fig. 5: Overlaps of RIC position differences between predictive and definitive ephemerides for GPM using the NRLMSISE-00 atmosphere model.

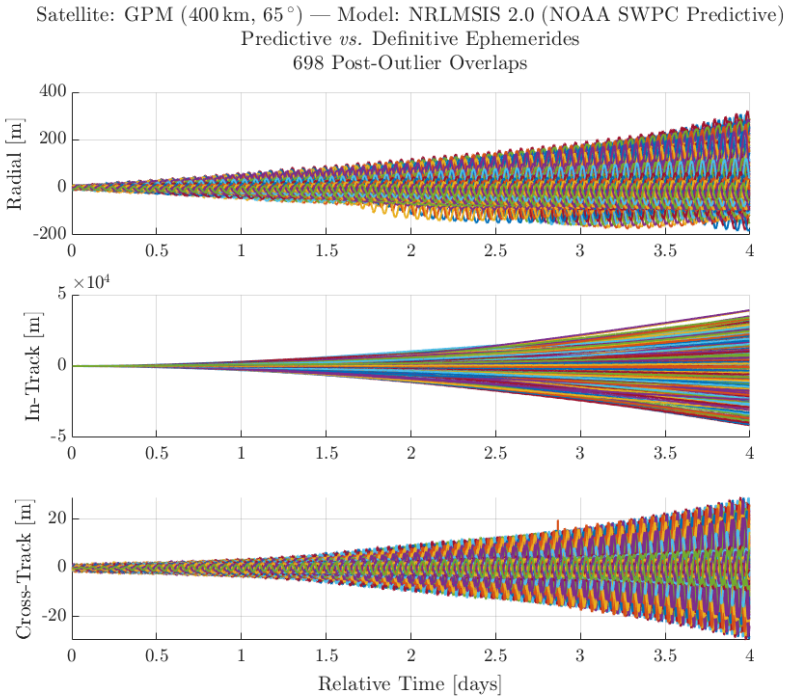


Fig. 6: Overlaps of RIC position differences between predictive and definitive ephemerides for GPM using the NRLMSIS 2.0 atmosphere model.

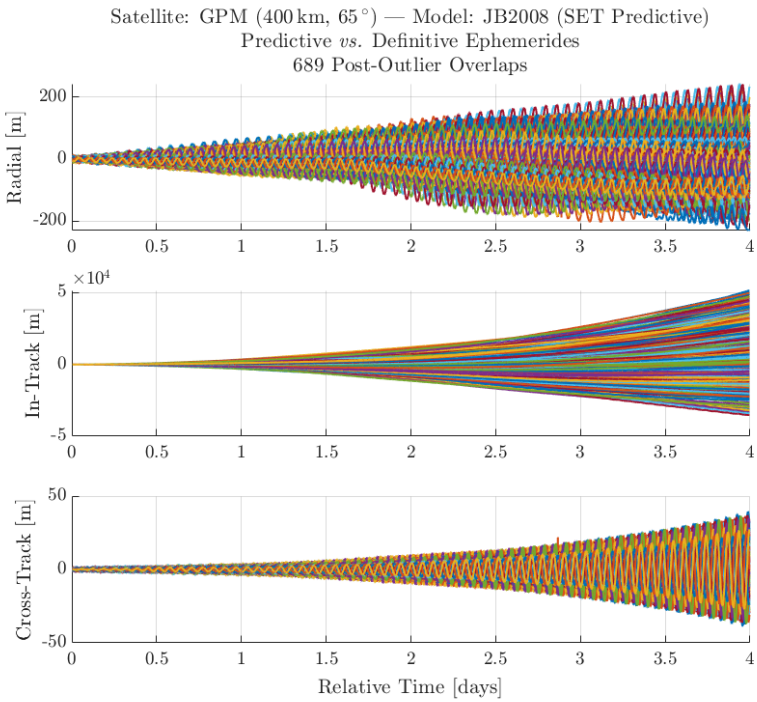


Fig. 7: Overlaps of RIC position differences between predictive and definitive ephemerides for GPM using the JB2008 atmosphere model.

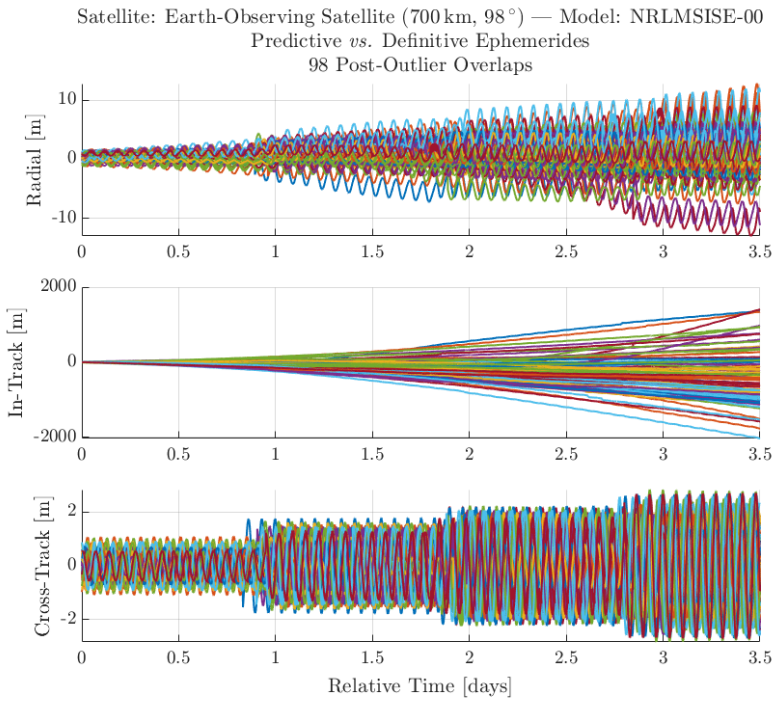


Fig. 8: Overlaps of RIC position differences between predictive and definitive ephemerides for an Earth-observing satellite using the NRLMSISE-00 atmosphere model.

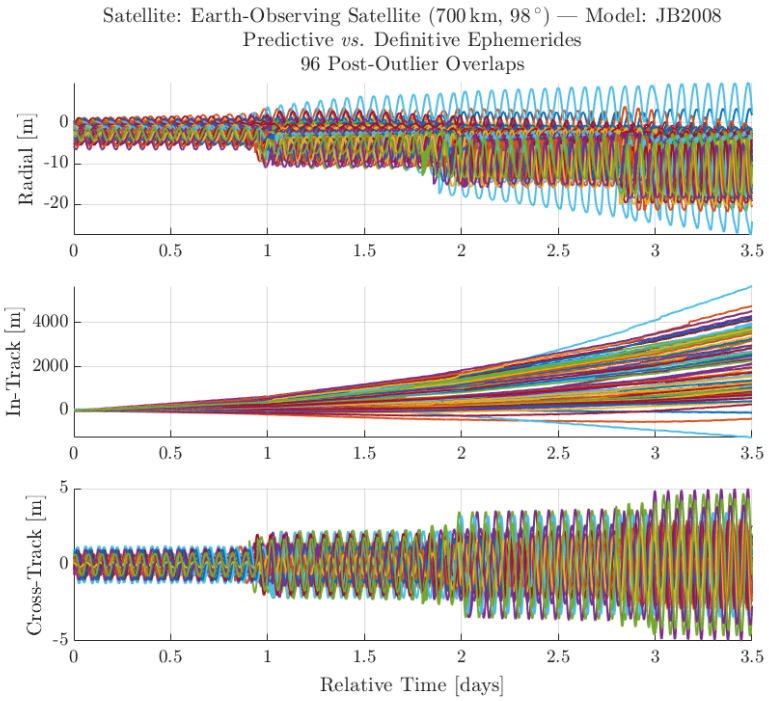


Fig. 9: Overlaps of RIC position differences between predictive and definitive ephemerides for an Earth-observing satellite using the JB2008 atmosphere model.

Table 7: Statistical prediction errors from predictive-definitive overlaps using different atmosphere models, all with predictive space weather data (NOAA SWPC for MSIS models, SET for JB2008).

Satellite	Model	Prediction	Mean Error, μ [m]			Std. Dev., 1σ [m]		
			Radial	In-Track	Cross-Track	Radial	In-Track	Cross-Track
GPM	NRLMSISE-00	24 h	0.92	−36.86	−0.02	14.65	859.89	1.12
		48 h	−1.04	−17.73	0.02	30.44	3542.65	2.43
		72 h	−3.22	664.81	−0.06	46.66	8196.73	4.80
	NRLMSIS 2.0	24 h	0.81	−33.74	-0.02	14.63	848.08	1.12
		48 h	−1.26	7.80	0.01	30.44	3513.67	2.44
		72 h	−3.60	732.58	−0.05	46.33	8155.85	4.82
	JB2008	24 h	0.29	−66.14	−0.01	16.22	1037.46	1.18
		48 h	−4.87	151.26	−0.01	34.48	4060.35	2.76
		72 h	−26.83	2875.25	−0.06	49.18	9474.61	6.00
Earth-Obs. Sat.	NRLMSISE-00	24 h	1.14	−83.60	0.06	1.17	82.60	0.56
		48 h	1.17	−213.76	0.15	2.26	251.96	0.88
		72 h	1.12	−347.99	0.12	3.11	520.57	1.24
	JB2008	24 h	−2.67	236.00	−0.07	2.75	193.07	0.58
		48 h	−5.11	731.48	0.06	4.07	560.05	0.95
		72 h	−7.36	1595.61	0.05	4.88	1061.93	1.49

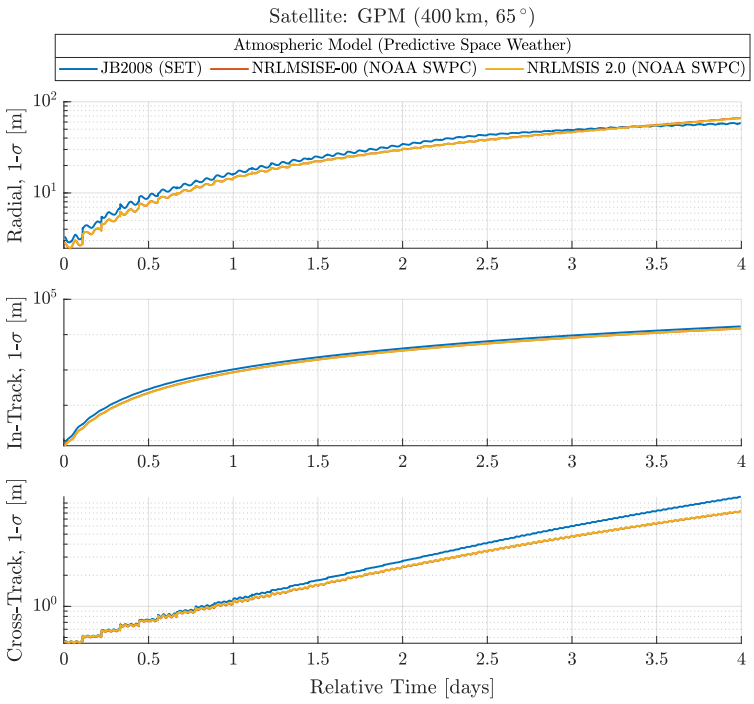


Fig. 10: Empirical covariance error growth *vs.* prediction time for GPM with different atmospheric models.

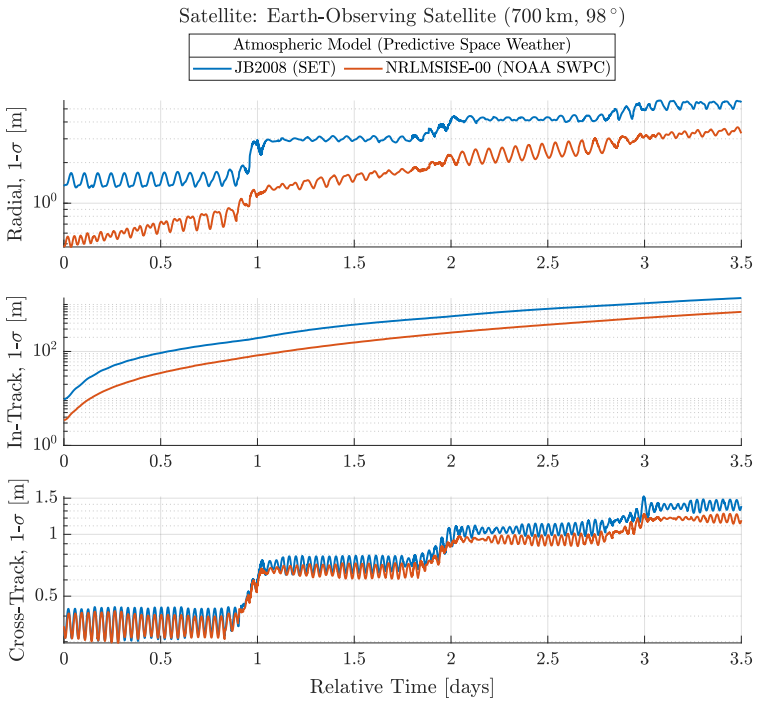


Fig. 11: Empirical covariance error growth *vs.* prediction time for an Earth-observing satellite with different atmospheric models.

Overall, we can see that *after* outlier removal, both MSIS models slightly outperform Jacchia–Bowman 2008 in terms of prediction accuracy despite requiring significantly less space weather inputs. The differences are fairly small for GPM, where prediction errors are already quite high due to the low 400 km altitude orbit being highly perturbed by drag. Differences between either MSIS and JB2008 in the in-track standard deviation prediction errors are around 20 % at most after 24 h of prediction, and slightly lower than that afterwards. However, at the 700 km altitude of the Earth-observing satellite, the prediction error differences are more pronounced; after 48 h of prediction, the empirical in-track uncertainty is more than doubled when comparing JB2008 to NRLMSISE-00.

Furthermore, we do not see any significant difference in accuracy between the NRLMSISE-00 and NRLMSIS 2.0 models. Despite the thermospheric density modeling changes in the newer version and updates for lower-altitude computations, these are not reflected in our results here. Differences in density are compensated with differences in C_D (see §5.1.3), but we do not see improvements in the prediction accuracy. It could be that since the focus of NRLMSIS 2.0 was on altitudes below 200 km, other satellites and, in particular, reentry studies could see improvements from the new model, but for satellite operations in LEO the difference appears to be negligible.

5.1.3 Drag Coefficient Analysis

Finally, since each OD solution returns a new solved-for drag coefficient, we can analyze the consistency of the C_D results as well as its changes throughout the analysis time interval. We plot this in Figure 12 for GPM and in Figure 13 for the Earth-observing satellite, with each graph containing data for all atmosphere model runs.

Again, we can see that NRLMSISE-00 and NRLMSIS 2.0 generate very similar results. However, in the C_D time series we can see some differences between the models. NRLMSIS 2.0 consistently produces higher drag coefficient estimates when compared to the older NRLMSISE-00. This is expected for a satellite flying at GPM's altitude since—all else being equal—the lower N_2 and O densities in the thermosphere need to be compensated with an increase in C_D.

At the same time, the JB2008 drag estimates vary wildly, particularly for GPM. This is very surprising to see since the rest of the results are all in line with the MSIS models, but the solved-for C_D appears to be significantly more variable in this model. While the estimate from MSIS is always around $C_D \approx 2$, some of the JB2008 estimates peak at $C_D > 10$, which is much higher than the typical nominal or physical drag coefficients for satellites. Overall, JB2008 appears to be predicting much lower neutral mass density values than MSIS, with the density being half of MSIS in some cases. This is true for both GPM and the higher-altitude Earth-observing satellite.

5.2 Space Weather Data Sources Comparison

In this last section, we focus on the effects of the space weather data only. We perform a variation of the previous analysis over the same timespan for GPM using the NRLMSISE-00 atmosphere and the NOAA SWPC and SET space weather files. Since the SET indices include both $F_{10.7}$ and A_p, they can be used with any atmospheric model, not only JB2008. However, recall that $\bar{F}_{10.7}$ from SET is a centered-average over 54 days, not the standard 81 days like in the NOAA SWPC data. We also compare the results of using predictive and definitive space weather from both data sources, i.e., using the known space weather data as of August 2023 instead of the daily or hourly predictions for each day.

Figure 14 displays the evolution of the RIC differences over the interval of analysis at 24 h of prediction after the OD epoch. In contrast with the results from §5.1.1, here we do not see any significant differences between each data source. This makes sense, as NRLMSISE-00 cannot take advantage of the SET-provided dT_c parameter that JB2008 uses for geomagnetic storm modeling.

After outlier removal, we follow the same procedure from §5.1.2 and compute the empirical covariance error growth profile for each space weather data source. The results are plotted in Figure 15 and detailed in Table 8. Note that the graph here uses a linear scale to more clearly showcase the differences between each source.

As expected, predictions using NOAA SWPC and SET predictive space weather behave very similarly. Despite the latter using a shorter 54-day centered average for the solar flux, that does not reflect in significant prediction error differences. However, when we compare either predictive source to the definitive data, the differences are more prominent—of the order of 30 %. This is the percentage of orbit prediction accuracy errors introduced by errors in the space weather predictions.

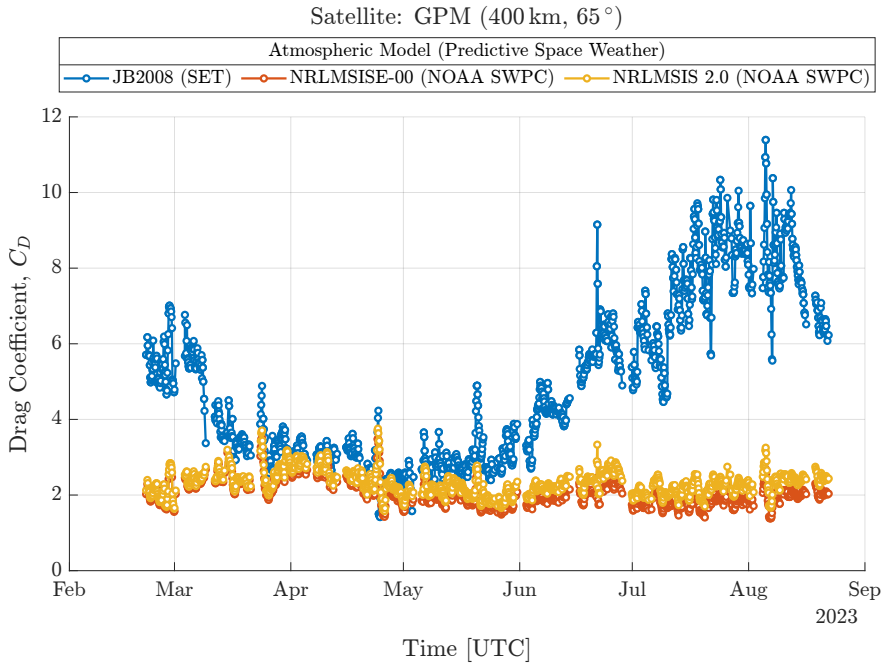


Fig. 12: Time series of the solved-for drag coefficient (C_D) history for GPM with different atmosphere models.

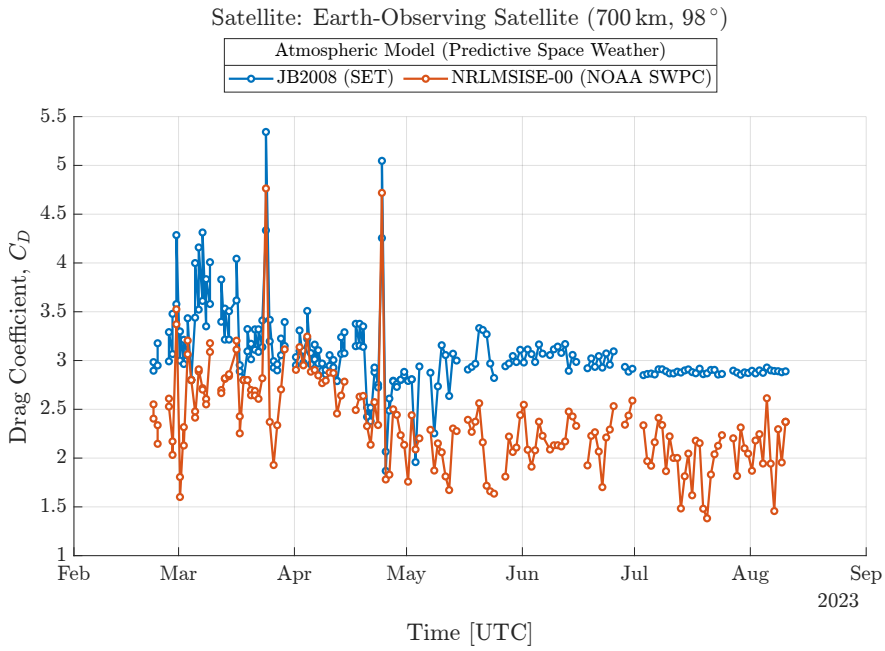


Fig. 13: Time series of the solved-for drag coefficient (C_D) history for an Earth-observing satellite with different atmosphere models.

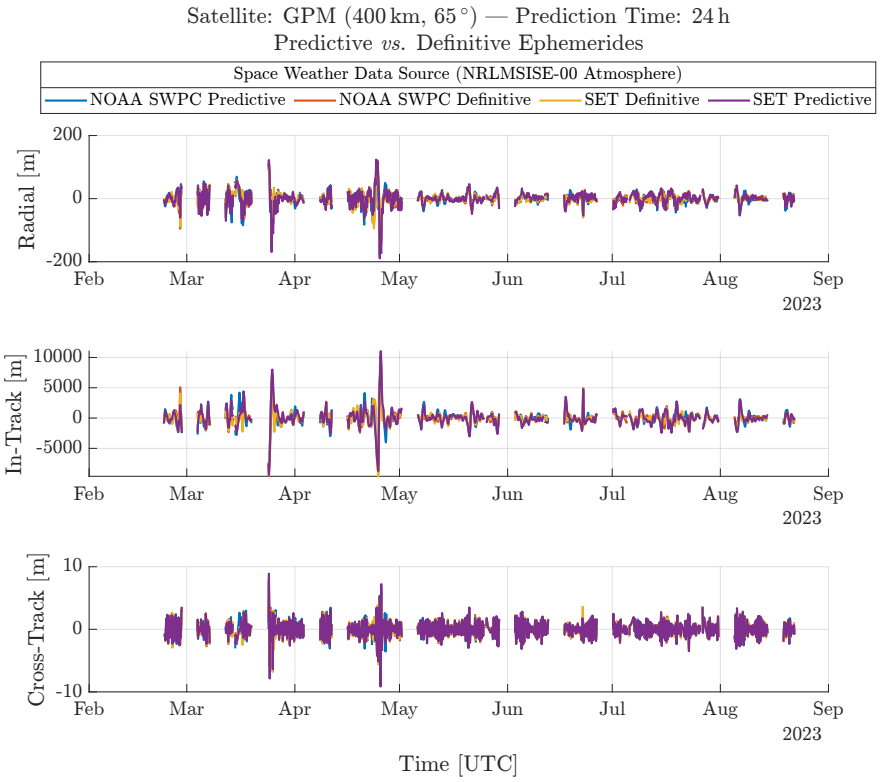


Fig. 14: RIC position differences evolution over time for GPM at 24 h of prediction time after the OD epoch using the NRLMSISE-00 atmosphere and NOAA SWPC and SET predictive and definitive space weather data.

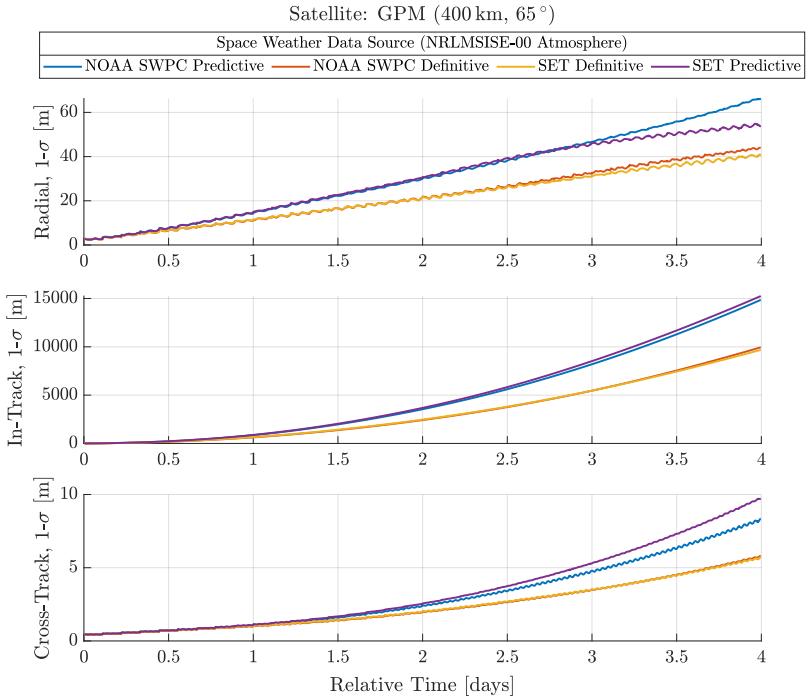


Fig. 15: Empirical covariance error growth *vs.* prediction time for GPM with NOAA SWPC and SET predictive and definitive space weather data.

Table 8: Statistical prediction errors from predictive-definitive overlaps for GPM using the NRLMSISE-00 atmosphere model and different space weather data sources from NOAA SWPC and SET.

Space Weather	Prediction	Mean Error, μ [m]			Std. Dev., 1σ [m]		
		Radial	In-Track	Cross-Track	Radial	In-Track	Cross-Track
NOAA SWPC Predictive	24 h	0.92	−36.86	−0.02	14.65	859.89	1.12
	48 h	−1.04	−17.73	0.02	30.44	3542.65	2.43
	72 h	−3.22	664.81	−0.06	46.66	8196.73	4.80
NOAA SWPC Definitive	24 h	−0.74	62.36	−0.02	11.40	628.24	1.03
	48 h	−1.15	255.94	−0.01	21.95	2418.26	2.01
	72 h	−1.30	595.14	0.03	32.68	5460.61	3.51
SET Predictive	24 h	−0.41	33.23	−0.00	15.04	881.68	1.14
	48 h	−4.19	343.69	−0.07	30.77	3682.71	2.57
	72 h	−18.29	2331.42	0.15	46.27	8522.80	5.30
SET Definitive	24 h	−0.60	54.96	−0.01	11.66	652.35	1.07
	48 h	−0.88	219.24	−0.03	21.81	2468.13	2.05
	72 h	−1.18	529.50	0.07	31.47	5459.38	3.53

6. CONCLUSIONS

In this work, we have presented an approach to evaluate the prediction accuracy of different atmospheric inputs when used in an operational setting for satellite propagation. In this framework, we run *both* orbit determination and prediction with each set of models so that predicted ephemerides can be generated with the correct drag coefficient estimates. Results in §5.1.3 show why this is the right approach, as we see how the solved-for C_D varies significantly at a given epoch based on the chosen model without necessarily accruing more or less errors in prediction.

Overall, the comparison of atmospheric models for GPM showed that in some cases, particularly during geomagnetic storms, the Jacchia–Bowman atmosphere can slightly reduce prediction errors compared to the MSIS models, with a limited number of cases showing a two-fold improvement in accuracy. However, on average, the differences between these models are fairly small, and when we aggregate all overlaps after outlier removal to compute an empirical covariance, we see that both MSIS models result in slightly smaller uncertainties: ~ 3500 m in-track uncertainty after 48 h compared to ~ 4000 m. Note that in all cases, the prediction errors for GPM are rather large due to the low 400 km altitude that is highly influenced by drag perturbations. The differences between JB2008 and MSIS are more obvious for the higher-altitude Earth-observing satellite data, where, for example, the in-track empirical standard deviation after 48 h of prediction is doubled with JB2008.

In our analysis, we did not see any significant differences between the more modern NRLMSIS 2.0 model and the classic NRLMSISE-00. It is possible that other use cases like reentry prediction could see improvements from the newer model given the focus on the lower atmosphere, but even at GPM's altitude, the main differences were observed in terms of slightly higher C_D estimates to compensate for the lower thermospheric density values. It is when analyzing the drag coefficient that we see that the solved-for C_D with the Jacchia–Bowman model is always significantly larger than the one from either MSIS model, indicating that JB2008 is estimating much lower neutral mass densities; in some extreme cases, the density values can be half of what MSIS predicts. This is the case for both GPM's 400 km orbit and the 700 km case.

In the last part of the study, we showed that using either the NOAA SWPC or SET predictive space weather sources produce fairly similar results with the NRLMSISE-00 atmosphere and GPM. This means that the SET data is only advantageous when used along with the JB2008 model which employs the extended set of solar flux and geomagnetic indices. However, the difference between using predictive or definitive space weather data for predictions is striking: radial and in-track uncertainties are reduced by more than 30 % after just 2 days of prediction or more. This indicates that a considerable percentage of the orbit prediction errors are introduced by space weather prediction errors. Future work by NOAA's Space Weather Prediction Center on physics-based atmosphere models, like WAM-IPE [9], has the potential to reduce our reliance on solar and geomagnetic indices and, in the process, improve prediction accuracy.

7. ACKNOWLEDGMENTS

The authors would like to personally thank W. Kent Tobiska from Space Environment Technologies for providing access to SET's real-time expanded JBHSGI solar and geomagnetic indices, as well as answering several questions on the implementation details of the Jacchia–Bowman 2008 atmosphere model.

8. REFERENCES

[1] Bruce Bowman, W Kent Tobiska, Frank Marcos, Cheryl Huang, Chin Lin, and William Burke. A New Empirical Thermospheric Density Model JB2008 Using New Solar and Geomagnetic Indices. In *AIAA/AAS astrodynamics specialist conference and exhibit*, page 6438, 2008.

[2] Bruce R. Bowman, W. Kent Tobiska, Frank A. Marcos, and Cesar Valladares. The JB2006 Empirical Thermospheric Density Model. *Journal of Atmospheric and Solar-Terrestrial Physics*, 70(5):774–793, March 2008.

[3] Dominik Brodowski. C Source Code for the NRLMSISE-00 Empirical Atmosphere Model. `https://www.brodo.de/space/nrlmsise`.

[4] J. T. Emmert, D. P. Drob, J. M. Picone, D. E. Siskind, M. Jones Jr., M. G. Mlynczak, P. F. Bernath, X. Chu, E. Doornbos, B. Funke, L. P. Goncharenko, M. E. Hervig, M. J. Schwartz, P. E. Sheese, F. Vargas, B. P. Williams, and T. Yuan. NRLMSIS 2.0: A Whole-Atmosphere Empirical Model of Temperature and Neutral Species Densities. *Earth and Space Science*, 8(3):e2020EA001321, 2021. e2020EA001321 2020EA001321.

[5] Government of Canada. Space Weather Canada. `https://www.spaceweather.gc.ca/index-en.php`.

[6] L. G. Jacchia. New Static Models of the Thermosphere and Exosphere with Empirical Temperature Profiles. *SAO Special Report*, 313, May 1970.

[7] L. G. Jacchia. Revised Static Models of the Thermosphere and Exosphere with Empirical Temperature Profiles. *SAO Special Report*, 332, May 1971.

[8] L. G. Jacchia. Thermospheric Temperature, Density, and Composition: New Models. *SAO Special Report*, 375, March 1977.

[9] National Oceanic and Atmospheric Administration. NOAA Whole Atmosphere Model-Ionosphere Plasmasphere Electrodynamics (WAM-IPE) Forecast System (WFS). `https://registry.opendata.aws/noaa-nws-wam-ipe`.

[10] National Oceanic and Atmospheric Administration. Space Weather Prediction Center. `https://www.swpc.noaa.gov`.

[11] Jamie Pawloski, Matthew Duncan, Siamak Hesar, Travis Lechtenberg, and Joshua Wysack. Realistic Covariance Generation for the GPM Spacecraft. In *2018 SpaceOps Conference*, page 2699, 2018.

[12] J. M. Picone, A. E. Hedin, D. P. Drob, and A. C. Aikin. NRLMSISE-00 Empirical Model of the Atmosphere: Statistical Comparisons and Scientific Issues. *Journal of Geophysical Research: Space Physics*, 107(A12):SIA 15–1–SIA 15–16, 2002.

[13] Space Environment Technologies. JB2008. `https://sol.spacenvironment.net/JB2008`.

[14] Byron D. Tapley, Bob E. Schutz, and George H. Born. *Statistical Orbit Determination*. Academic Press, Burlington, 2004.

[15] U.S. Naval Research Laboratory Geospace Science & Technology Branch. NRLMSIS 2.0 Publicly Accessible Data Downloads. `https://map.nrl.navy.mil/map/pub/nrl/NRLMSIS/NRLMSIS2.0`.

[16] U.S. Naval Research Laboratory Geospace Science & Technology Branch. NRLMSISE-00 Publicly Accessible Data Downloads. `https://map.nrl.navy.mil/map/pub/nrl/NRLMSIS/NRLMSISE-00`.

[17] David A. Vallado and Wayne D. McClain. *Fundamentals of Astrodynamics and Applications*. Number 21 in Space Technology Library. Microcosm Press, Hawthorne, CA, fourth edition, 2013.