Uncertainty quantification of the DTM2020 thermosphere model

Claude Boniface* and Sean Bruinsma

OMP/GET-CNRS, Space Geodesy Office, 14 Avenue E. Belin, 31401 Toulouse cedex 4, France

Received 11 June 2021 / Accepted 5 September 2021

Abstract—Aims: The semi-empirical Drag Temperature Models (DTM) calculate the Earth’s upper atmosphere’s temperature, density, and composition. They were applied mainly for spacecraft orbit computation. We developed an uncertainty tool that we implemented in the DTM2020 thermosphere model. The model is assessed and compared with the recently HASDM neutral density released publicly in 2020. Methods: The total neutral density dataset covers all high-resolution CHAMP, GRACE, GOCE, and SWARM data spanning almost two solar cycles. We constructed the uncertainty model using statistical binning analysis and least-square fitting techniques, allowing the development of a global sigma error model to function the main variabilities driving the thermosphere state. The model is represented mathematically by a nonlinear manifold approximation in a 6-D space parameter. Results: The results reveal that the altitude parameter presents the most notable error range during quiet and moderate magnetic activity ($K_p \leq 5$). However, the most considerable uncertainty appears during severe or extreme geomagnetic activities. The comparison with density data provided by the SET HASDM database highlights some coherent features on the mechanisms occurring in the thermosphere. Moreover, it confirms the tool’s relevance to provide a qualitative database of neutral densities uncertainties values deduced from the DTM2020 model.

Keywords: Thermosphere / DTM2020 / uncertainty / modeling / neutral density

1 Introduction

The thermosphere operational and research DTM2020 (Drag Temperature Model) models were developed in the framework of the Space Weather Atmosphere Models and Indices project (SWAMI, http://swami-h2020.eu), which was a European Union Horizon 2020 project. Semi-empirical thermosphere specification models are used to compute the atmospheric drag force in the orbit prediction of objects in Low Earth Orbit (LEO), Conjunction Assessment (CA) and subsequent satellite collision avoidance maneuver. Operational decisions are made without taking its utmost contributor to uncertainty, the thermosphere density estimation, properly into account in the orbit determination process (Bruinsma et al., 2021b; Hejduk & Snow, 2018). Until recently, propagation of the neutral density uncertainties to orbital state uncertainties has received little attention, but awareness for this theme is growing (e.g., Vallado & Finkleman, 2014; Emmert et al., 2017; Schiemenz et al., 2019; Lopez-Jimenez et al., 2021). A key hurdle to including the neutral density uncertainty in a space object’s covariance is that thermosphere models do not deliver this information up to now. In this work, we developed a semi-analytical model to characterize the uncertainty of atmospheric modeling for DTM2020 as a function of location (latitude, local time, altitude), season (day of the year), and two indices ($F10.7$ and $K_p$) describing the solar and geomagnetic activities. The model calculates the 1-sigma uncertainty of Observed to Computed density ratio (“O/C”), which combines statistical data binning techniques and the least-squares fitting procedure.

The drag force represents the largest source of uncertainty when extrapolating the orbit of spacecraft and space debris in LEO. Upper atmosphere density varies considerably in space and time due to highly variable solar and geomagnetic activity and seasonal effects (e.g., Guo et al., 2007; Lathuillère et al., 2008; Qian & Solomon, 2011; Emmert, 2015). Consequently, atmospheric drag on satellites varies significantly due to the proportional relation with thermospheric mass density. Other error sources in the drag computation, which are not addressed in this paper, are due to the unknown attitude of the object, approximate knowledge of its mass (Hejduk & Snow, 2018), and its aerodynamic drag coefficient (Mehta et al., 2014). The uncertainty in the thermospheric density modeling is generally 10–15% (Vallado & Finkleman, 2014) for a quiet thermosphere state. However, it can increase rapidly depending on the solar and geomagnetic activity, notably in the case of geomagnetic...
storms (e.g., Liu & Luhr, 2005; Forbes et al., 2005; Sutton et al., 2009; Emmert, 2015).

 Provision of the model uncertainty was also one of the recommendations in COSPAR’s global roadmap 2015–2025 for Space Weather (Schrijver et al., 2015). The uncertainty model for DTM2020 is based on data in the altitude range from 200 to 500 km. Below and above that altitude interval, density data availability is too sparse due to too high and too low drag values, respectively. Neutral density decreases exponentially with height, and satellite drag is too large for satellites to maintain orbit without propulsion below approximately 350 km. In contrast, above approximately 500 km, satellite drag is usually too small to be accurately derived from orbital data. However, owing to the importance of re-entry predictions on the one hand (Visser & Van den Ijssel, 2016; Hejduk & Snow, 2018) and the primary operating altitude of Earth Observation satellites at around 800 km, the uncertainties are provided for the altitude range 100–1000 km, by making presumptions described in the paper.

 Section 2 introduces the datasets used in the model building, and Section 3 briefly describes the DTM2020 thermosphere model. The data analysis and modeling technique are presented in Sections 4 and 5, respectively, whereas the results are discussed in Section 6. A fully independent model validation using SET HASDM data (Tobiska et al., 2021) is specified in Section 7, and the paper is finally summarized in Section 8.

2 Neutral density data

2.1 Model development density datasets

Four density datasets from different satellite missions were used in the study to build the uncertainty model. These are only briefly described in this section.

CHAMP (Challenging Minisatellite Payload) was a small satellite mission dedicated to gravity and atmospheric research led by the GFZ institute (Reigber et al., 1996). The satellite was launched on 15 July 2000 in a near-circular and nearly polar orbit with an inclination of 87.3° at an altitude of nearly 460 km. Densities inferred from accelerometer measurements, computed by TU Delft (Doornbos, 2011) using high-fidelity satellite shape and aerodynamic coefficient models, are available for the entire 10-year mission.

GRACE (Gravity Recovery and Climate Experiment) was a joint NASA and DLR mission to measure temporal variations in Earth’s gravity field (Tapley et al., 2004). The identical pair of spacecraft was launched on 17 March 2002 into a near-circular and polar orbit. Densities were inferred for the entire mission from July 2002, at 500 km altitude, to December 2016, at 350 km due to natural decay.

GOCE (Gravity Field and Steady-State Ocean Circulation Explorer) was an ESA satellite mission to measure the Earth’s gravity field and model the geoid with high accuracy (Drinkwater et al., 2003). It was launched on 17 March 2009 in a 96.5° inclination, near-circular orbit. Its science mission started in November 2009, at a mean altitude of 255 km. The orbit was lowered several times over the mission, and its lowest at 224 km was reached in May 2013. TU Delft derived densities under the ESA contract, using high-fidelity satellite shape and aerodynamic coefficient models, from the thruster data from November 2009 to September 2013 (Bruinsma et al., 2014).

The three aforementioned satellites carried high-resolution and high-precision accelerometers (Touboul et al., 1999), increasingly precise from CHAMP, GRACE to GOCE. The total densities are inferred from the 0.1 or 0.2 Hz measurements of non-gravitational accelerations (even if indirectly in the case of GOCE). The satellite speed is approximately 8 km s⁻¹, which translates to an along-track resolution of approximately 80 or 40 km, respectively, between one and two orders of magnitude smaller than the thermosphere model resolution.

Finally, SWARM is an ESA Earth observation mission consisting of a constellation of three satellites launched in November 2013 (Friis-Christensen et al., 2008). SWARM A and SWARM C are orbiting side-by-side at an altitude of 462 km at an inclination of 87.35°, whereas SWARM B is flying at a higher altitude of 511 km at the same inclination. The neutral density used in this study was deduced from GPS tracking data on SWARM A (Van den Ijssel et al., 2020) instead of the accelerometer due to severe problems with that instrument. At the altitude of about 460 km, neutral density can only be derived accurately for high and medium solar activity and with a spatial resolution of the order of thousands of kilometers with GPS data. In this study, SWARM A densities for F10.7 less than 100 sfu were rejected.

Figure 1 shows the time evolution of neutral density provided by the four satellites. The database 176 spans 2001 to 2019, i.e., covering almost two solar cycles (23 and 24).

2.2 Validation density dataset

Space Environment Technologies (SET) has recently released 20 years of neutral densities through the SET HASDM Database (Tobiska et al., 2021), which is derived from the US Space Force High Accuracy Satellite Drag Model (Storz et al., 2005). However, these densities are not observations but outputs of the HASDM model, which is a Jacchia model corrected in near real-time through data assimilation. The database provides neutral density every 3 h from 2000 to 2019 on 10° latitude × 1 h local time grids, and 25 km stepped altitude from 175 to 825 km. We validated the uncertainty model using the SET HASDM data on the altitudes 250, 325, 400, and 475 km for the entire 20-year time span.

3 The DTM2020 thermosphere model

The DTM2020 model is briefly discussed in this section. A companion paper (Bruinsma & Boniface, 2021) details the improvements of this model and presents a complete assessment and comparison with the CIRA thermosphere models. The Drag Temperature Model (DTM) is a semi-empirical model reproducing temperature, density, and composition of the thermosphere climatology. The total density data of the four satellites were assimilated in the DTM2020 thermosphere model. We developed a 1-sigma uncertainty estimate of the neutral density for operational purposes. That is why the model is constructed based on the operational DTM2020, which uses the classical solar and geomagnetic drivers F10.7 and the planetary index Kp. Although the solar radio flux at
30 cm F30 (Dudok de Wit et al., 2014) and the high-cadence $Hpo$ geomagnetic indices (Matzka et al., 2019) represent variations in upper atmosphere heating accurately than F10.7 and $K_p$, they are not yet certified for operational use. We emphasize that the uncertainty model developed in this study only gives the uncertainty of the DTM2020 model density and not of atmospheric drag, for which more sophisticated tools like state vector covariance of space objects are required (Hilton et al., 2019).

4 Data processing and analysis method

Several works have studied uncertainty in the thermospheric density in orbit prediction, but their quantitative impact on the orbital dynamics has not been thoroughly investigated. Most of them concentrated on the errors in neutral density due to inaccurate space weather inputs and forecasts because these errors make a significant impact on orbit prediction (Anderson et al., 2009; Schiemenz et al., 2019). Therefore, density models must be adjusted, and the uncertainties in the model must be estimated. Several techniques exist for density model calibration (Doornbos et al., 2008; Chen et al., 2019). However, these techniques do not directly assess the uncertainty in the density.

Furthermore, long-term orbit prediction is computationally very time-consuming (Petit & Lemaître, 2016). Thus, the accurate prediction of neutral density and atmospheric drag for orbital uncertainty propagation generally considers simplified thermospheric mass density (Giza et al., 2009). Our goal is not to forecast the orbit of satellites but to provide an upstream estimate of neutral density uncertainties calculated with the DTM2020 thermosphere model for operational use. That is why a simplified statistical approach is preferred and described in the following section.

4.1 Data fusion

The neutral density datasets provided by the satellites GOCE, CHAMP, GRACE, and SWARM are combined. Data fusion by integrating the multiple data sources produces a complete neutral density database in the range of 200–500 km.

Then, the new version of the semi-empirical DTM model (see Sect. 3), using the classical $K_p$ and F10.7 drivers, is used to compute the neutral densities to assess the uncertainty of the model’s prediction. The model and data are not compared by computing density residuals represented as absolute differences (observed minus computed). Indeed, an observed-to-computed density ratio ($O/C$) is preferred because it is a more appropriate quantity to express a model’s ability to reproduce the observations. Using density ratios comprehensively allows interpreting all features of the complex dynamic changes of neutral density as a function of geographical coordinates (latitude, altitude, local time) and solar and geomagnetic drivers (Bruinsma, 2015). Moreover, due to its skewed distribution, the mean density ratio is determined in log space. Equations (1) and (2) are used for the statistical analysis (Bruinsma et al., 2018):

\[
\langle O/C \rangle = e^{\left(\frac{1}{N} \sum_{i=1}^{N} \ln(O/C_i)\right)},
\]

\[
\text{Std} \langle O/C \rangle = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \left(\ln\left(\frac{O_i}{C_i}\right) - \ln\left(\langle O/C \rangle\right)\right)^2}
\]

where $\langle O/C \rangle$ is the mean observed-to-computed density ratio and Std ($O/C$) is the standard deviation expressed in percent.

Several tests were performed to ensure the relevance of the data fusion process technique. Indeed, the neutral density provided by each satellite was analyzed separately. We observed similar 1-sigma uncertainty trends. The histogram of the observed densities to the corresponding DTM2020 model densities (log scale) is illustrated in Figure 2. The neutral density distribution and a normal distribution $N(\mu = 0.02, \sigma = 0.2)$ present a relatively good agreement.

4.2 Binning data

Statistical data binning is applied on the neutral density database to group numbers of more or less continuous values into a smaller number of “bins.” Then, the neutral density is averaged in each interval. This binning process on the density ratios is helpful to assess the quality of the model according
5 The 1-sigma uncertainty semi-analytical model

5.1 1-sigma uncertainty polynomial functions

Density ratios were binned in several physical parameters of DTM2020 (latitude, altitude, day of the year, local time, solar flux, and geomagnetic index). The equations representing the best fit in the least-squares sense were computed for each parameter. The least-square fitting model to data has polynomial function forms. Equation (3) summarizes the set of coupled equations:

\[
\begin{align*}
\sigma_{\text{alt}} &= a_{1,\text{alt}} \times x_{\text{alt}} + a_{0,\text{alt}} \\
\sigma_{\text{alat}} &= a_{2,\text{alat}} \times x_{\text{alat}}^2 + a_{1,\text{alat}} \times x_{\text{alat}} + a_{0,\text{alat}} \\
\sigma_{\text{hl}} &= a_{3,\text{hl}} \times x_{\text{hl}}^3 + a_{2,\text{hl}} \times x_{\text{hl}}^2 + a_{1,\text{hl}} \times x_{\text{hl}} + a_{0,\text{hl}} \\
\sigma_{\text{day}} &= a_{1,\text{day}} \times x_{\text{day}} + a_{0,\text{day}} \\
\sigma_{\text{fp}} &= a_{1,\text{fp}} \times x_{\text{fp}} + a_{0,\text{fp}} \\
\sigma_{\text{kp}} &= a_{3,\text{kp}} \times x_{\text{kp}}^3 + a_{2,\text{kp}} \times x_{\text{kp}}^2 + a_{1,\text{kp}} \times x_{\text{kp}} + a_{0,\text{kp}}
\end{align*}
\]

See the Equation (3) bottom of the page

where \(x_{\text{alt}}, x_{\text{alat}}, x_{\text{hl}}, x_{\text{day}}, x_{\text{fp}}, x_{\text{kp}}\) are altitude, latitude, local time, day of the year, solar flux, and geomagnetic index variables. \(\sigma_{x}\) is the 1-sigma uncertainty for each parameter. Table 1 lists the coefficients \(\{a_{i,x}\}\) of each equation.

To summarize, the altitude parameter presents the largest uncertainty range during quiet and moderate geomagnetic activity \((K_p \leq 5)\). For higher \(K_p\), the most significant uncertainty is mainly due to the geomagnetic activity (i.e., severe or extreme geomagnetic storms).

---

Figure 4 illustrates the 1-sigma uncertainty (in percent) binned in latitude (top) and altitude (bottom) with the corresponding least-squares fits to the data. The 1-sigma error presents a quadratic evolution in latitude ranging from 22% in the tropics to 24–26% to the poles. Moreover, the uncertainty increases strongly linearly with the altitude (from 10% at 230 km to 30% at 500 km).

We applied the exact process to the other parameters (see Fig. 5). The main trends are:

- A quadratic evolution is noted in solar local time (from 20% at 15 h to 30% at midnight).
- A constant error of 23% as a function of the day of the year.
- The 1-sigma error decreases linearly with the solar flux (from 25% at low solar activities to 20% at high solar activities).
- The 1-sigma uncertainty has a cubic dependency with the magnetic index (from 20% to 60% at high \(K_p\)-Index). The uncertainty has the same order of values as the other parameters at low \(K_p\) (between 20 and 30% for low or moderate magnetic activities). A substantial increase in error appears for higher \(K_p\) (typically for severe geomagnetic storms).
Table 1. List of polynomial coefficient values for each parameter.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Altitude</td>
<td>(a_{1,\text{alt}} = 7.92 \times 10^{-2}, a_{0,\text{alt}} = -9.73)</td>
</tr>
<tr>
<td>Latitude</td>
<td>(a_{2,\text{alt}} = 4.72 \times 10^{-3}, a_{1,\text{alt}} = -7.19 \times 10^{-3}, a_{0,\text{alt}} = 22.33)</td>
</tr>
<tr>
<td>Local hour</td>
<td>(a_{3,\text{hl}} = 3.15 \times 10^{-3}, a_{2,\text{hl}} = -5.64 \times 10^{-2}, a_{1,\text{hl}} = -3.95 \times 10^{-1}, a_{0,\text{hl}} = 28.53)</td>
</tr>
<tr>
<td>Day of year</td>
<td>(a_{1,\text{day}} = -3.91 \times 10^{-1}, a_{0,\text{day}} = 23.15)</td>
</tr>
<tr>
<td>Solar flux (F_p)</td>
<td>(a_{1,F_p} = -3.36 \times 10^{-2}, a_{0,F_p} = 26.04)</td>
</tr>
<tr>
<td>Magnetic index (K_p)</td>
<td>(a_{3,K_p} = 6.56 \times 10^{-2}, a_{2,K_p} = -1.99 \times 10^{-1}, a_{1,K_p} = 7.46 \times 10^{-1}, a_{0,K_p} = 22.47)</td>
</tr>
</tbody>
</table>

Fig. 4. (Top) 1-σ uncertainty in percent binned in latitude and the model fitting to data in the least-squares (LS) sense, (bottom) 1-σ uncertainty in percent binned in altitude, and the model fitting to data in the least-squares sense. Data is in blue, and the LS fitting in red color.

5.2 The least-square fit over manifolds

We performed a simplified procedure for fitting a set of least-squares polynomial functions like those obtained in equation (3) to map a manifold structure (Sober et al., 2020; Rubin-Delanchy, 2021). The 1-sigma uncertainty is deduced from the contribution of overall uncertainty described in equation (3). We assume that the 1-sigma uncertainty is defined, in the space of \(N\) dimensions, by the following relations:

\[
\sigma(x_1, x_2, \ldots, x_i, \ldots, x_N) = \sum_{i=1}^{N} \sum_{k=0}^{N_{\text{deg}}} (a_{i,k} \times x_i^k) + A
\]

(4)

\[
\sigma(0, 0, \ldots, 0) = \sum_{i=1}^{N} a_{i,0} = A
\]

(5)

\[
\sigma(x_1, x_2, \ldots, x_i, \ldots, x_N) = \sum_{i=1}^{N} \sum_{k=0}^{N_{\text{deg}}} (a_{i,k} \times x_i^k) + A
\]

(6)

where \(\{a_{i,k}\}\) are the coordinates-wise slopes, \(\{x_i\}\) the variables, and \(A\) is a constant. The coefficient \(k\) is related to the degree of the polynomial, and \(N_{\text{deg}}\) is the maximum degree of the polynomial for the variable \(x_i\). Equation (6) is indeterminate, and the constant \(A\) must be approximated. For that, we postulate that \(A\) is inferred from equations (7) to (8):

\[
\sigma(x_1, x_2, \ldots, x_i, \ldots, x_N) = \sum_{i=1}^{N} \sum_{k=0}^{N_{\text{deg}}} (a_{i,k} \times x_i^k) + \sum_{i \neq j} \sum_{k=1}^{N_{\text{deg}}} (a_{i,k} \times \bar{x}_j^k)
\]

(7)

\[
a_{j,0} = \sum_{i \neq j} \sum_{k=1}^{N_{\text{deg}}} (a_{i,k} \times \bar{x}_j^k) + A, \quad \forall j = 1 \ldots N
\]

(8)

where \(\bar{x}_j\) is the mean value of each distribution of the variable except for \(x_j\) chosen as a reference parameter. Then, the non-linear manifold-like equation (9) is solved step-by-step to estimate the 1-sigma uncertainty:

\[
\sigma(x_1, x_2, \ldots, x_i, \ldots, x_N) = \sum_{i=1}^{N} \sum_{k=1}^{N_{\text{deg}}} (a_{i,k} \times x_i^k) + \sum_{i \neq j} \sum_{k=1}^{N_{\text{deg}}} (a_{i,k} \times \bar{x}_j^k)
\]

(9)
In this work, the 1-sigma relation is simplified in 6-D dimensions with the parameters latitude, altitude, local time, day of the year, solar and geomagnetic index.

The following equation approximates the 1-sigma uncertainty by using the set of equations (3) and the process (4) → (9) for \( j = 0 \):

\[
\sigma(x_{\text{alti}}, x_{\text{alat}}, x_{\text{hl}}, x_{\text{day}}, x_{\text{Fp}}, x_{\text{Kp}}) = \left[ a_{1, \text{alti}} \times x_{\text{alti}} + a_{2, \text{alti}} \times x_{\text{alti}}^2 + a_{3, \text{hl}} \times x_{\text{hl}} + a_{4, \text{hl}} \times x_{\text{hl}}^2 + a_{5, \text{day}} \times x_{\text{day}} + a_{1, \text{Kp}} \times x_{\text{Kp}} + a_{2, \text{Kp}} \times x_{\text{Kp}}^2 + a_{3, \text{Fp}} \times x_{\text{Fp}}^2 \times a_{4, \text{Fp}} \times x_{\text{Fp}} \times a_{5, \text{hl}} \times x_{\text{hl}} + a_{6, \text{hl}} \times x_{\text{hl}}^2 + a_{7, \text{day}} \times x_{\text{day}} - a_{8, \text{Kp}} \times x_{\text{Kp}} + a_{9, \text{Fp}} \times x_{\text{Fp}}^2 - a_{10, \text{Kp}} \times x_{\text{Kp}} \times x_{\text{Kp}} \right].
\]

The coefficients of the polynomial functions \( \{a_{1, \text{alti}}\}, \{a_{1, \text{alti}}\}, \{a_{1, \text{hl}}\}, \{a_{1, \text{day}}\}, \{a_{1, \text{Fp}}\}, \{a_{1, \text{Kp}}\} \) are related respectively to altitude, latitude, local time, day of the year, solar flux, and geomagnetic index. \( x^- \) denotes the mean value of each physical parameter. The altitude is the reference variable in the equation. The averaged values are deduced from the neutral density distribution: \( x_{\text{alti}} = 415 \text{ km}, \ x_{\text{alat}} = 0.15^\circ, \ x_{\text{hl}} = 12 \text{ h}, \ x_{\text{day}} = 180 \text{ days}, \ x_{\text{Fp}} = 109, \ x_{\text{Kp}} = 1.90 \).

The following remarks can be made here. First of all, a direct least-square approach on the overall neutral density dataset (i.e., without using binning process) may be possible. But due to the very large number of sparsely distributed data (not sparse and so much more for the HASDM dataset), the problem is too time-consuming to solve. The main idea was to use the least-squares mechanisms on a local rather than global level. This simplified approach can capture the local trends of the data to reconstruct better broader sets of functions like the one obtained in equation (3). Secondly, the first modeling performed to reduce the problem linearly was done using hyperplane (Reister, 1997) instead of manifold techniques. Indeed, density ratios were binned using piecewise linear functions deduced in the least-square sense. We observed that the two approaches lead to practically the same results (not presented in the paper). We adopted the manifold technique, allowing both a fitter model charting representation and a more efficient code. Finally, we highlight that the model calculates a mean 1-sigma uncertainty \( \sigma(x_{\text{alti}}, x_{\text{alat}}, x_{\text{hl}}, x_{\text{day}}, x_{\text{Fp}}, x_{\text{Kp}}) \) of 22%, in excellent agreement with the standard deviation deduced from the complete neutral density distribution (see Fig. 3).

5.3 1-sigma error model

Equation (10) is well justified in the range 200–500 km. These presumptions were made to extend the altitude range to 100–1000 km due to the unavailability of data:

- at a lower altitude (< 200 km), the worst case was chosen by fixing the evolution of 1-sigma to their values at 200 km,
- at a higher altitude (> 500 km), an optimistic case was assumed by keeping the evolution of 1-sigma for an altitude of 500 km.

Therefore, the global model describing the 1-uncertainty in the range [100 km–1000 km] is:

\[
\begin{align*}
\sigma(x_{\text{alti}} &= 200, x_{\text{alat}}, x_{\text{hl}}, x_{\text{day}}, x_{\text{Fp}}, x_{\text{Kp}}) \text{ if } x_{\text{alti}} < 200 \\
\sigma(x_{\text{alti}}, x_{\text{alat}}, x_{\text{hl}}, x_{\text{day}}, x_{\text{Fp}}, x_{\text{Kp}}) \text{ if } 200 \leq x_{\text{alti}} \leq 500 \\
\sigma(x_{\text{alti}} &= 500, x_{\text{alat}}, x_{\text{hl}}, x_{\text{day}}, x_{\text{Fp}}, x_{\text{Kp}}) \text{ if } x_{\text{alti}} > 500
\end{align*}
\]

where \( \sigma(x_{\text{alti}}, x_{\text{alat}}, x_{\text{hl}}, x_{\text{day}}, x_{\text{Fp}}, x_{\text{Kp}}) \) is the function defined in equation (10).

Table 2 displays 1-sigma error at different altitudes (from 100 to 1000 km) and for latitude = 0°, local time = 12 h, day of the year = 180, and two values of solar flux = 65 and 200 (low and high solar activity) and geomagnetic activity \( K_p = 4 \).
and 8 (enhanced and high geomagnetic activity). The 1-σ uncertainty is growing with altitude. The values are five times higher at high solar and low geomagnetic activities passing from 6.84% at 100 km to 30.60% at 1000 km. Likewise, the error is almost two times greater at low altitude (below 200 km) during low solar activity (6.84% at $x_{Fp} = 200$, against 11.37% at $x_{Fp} = 65$). Furthermore, the 1-σ uncertainty can have the same order of magnitude for two different altitudes due to the level of geomagnetic activity (34.19% at 100 km for $x_{Kp} = 8$ against 35.13% at 1000 km for $x_{Kp} = 4/x_{Fp} = 65$).

### 6 Results and discussion

The model computes the 1-sigma error in 6-D parameter space, including geographical coordinates (latitude, altitude, solar local time), season (day of the year), solar and geomagnetic drivers ($F10.7$ proxy and $Kp$ index). As discussed previously, altitude is one of the main drivers in error modeling during quiet and disturbed thermosphere states, and so this parameter is compared with the other physical quantities. The plots are limited to the range of 200–500 km to illustrate the results. Figure 6 shows 3D surface plotting of the 1-sigma uncertainty as a function of latitude, local time, and altitude parameters. Curvilinear shapes are observed due to latitude (left side) and local time (right side) effects. The 1-sigma uncertainty is minimum at the lowest altitude in the tropics (~5–10%), whereas a maximum of 30–35% is localized at the highest altitude (here 500 km) over the poles. It presents a predominantly diurnal variation with a peak uncertainty (~36%) at 500 km altitude at midnight. The error is minimum at the lowest altitude at 15 h solar local time (~6%). Moreover, the ratio between the uncertainty on the night side (22:00–05:00 LT) and dayside (10:00–17:00 LT) linearly decreases with altitude (1.2 at 500 km and 1.6 at 250 km). The latitude effects are also altitude-dependent, with an uncertainty ratio pole/equator of 1.1 at 500 km against 1.4 at 250 km.

The density error exhibits a solar cycle variation, driven by the intrinsic ~11-year variability of magnetic activity in the Sun (Qian & Solomon, 2011), as shown in Figure 7 (right frame). The maximum uncertainty is observed at a given altitude during

<table>
<thead>
<tr>
<th>Altitude (km)</th>
<th>($x_{Fp} = 65, x_{Kp} = 4$)</th>
<th>($x_{Fp} = 200, x_{Kp} = 4$)</th>
<th>($x_{Fp} = 65, x_{Kp} = 8$)</th>
<th>($x_{Fp} = 200, x_{Kp} = 8$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>11.37</td>
<td>6.84</td>
<td>34.19</td>
<td>29.66</td>
</tr>
<tr>
<td>150</td>
<td>11.37</td>
<td>6.84</td>
<td>34.19</td>
<td>29.66</td>
</tr>
<tr>
<td>200</td>
<td>11.37</td>
<td>6.84</td>
<td>34.19</td>
<td>29.66</td>
</tr>
<tr>
<td>225</td>
<td>13.35</td>
<td>8.82</td>
<td>36.17</td>
<td>31.64</td>
</tr>
<tr>
<td>250</td>
<td>15.33</td>
<td>10.80</td>
<td>38.15</td>
<td>33.62</td>
</tr>
<tr>
<td>275</td>
<td>17.31</td>
<td>12.78</td>
<td>40.13</td>
<td>35.60</td>
</tr>
<tr>
<td>300</td>
<td>19.29</td>
<td>14.76</td>
<td>42.11</td>
<td>37.58</td>
</tr>
<tr>
<td>325</td>
<td>21.27</td>
<td>16.74</td>
<td>44.09</td>
<td>39.56</td>
</tr>
<tr>
<td>350</td>
<td>23.25</td>
<td>18.72</td>
<td>46.07</td>
<td>41.54</td>
</tr>
<tr>
<td>375</td>
<td>25.23</td>
<td>20.70</td>
<td>48.05</td>
<td>43.52</td>
</tr>
<tr>
<td>400</td>
<td>27.21</td>
<td>22.68</td>
<td>50.03</td>
<td>45.50</td>
</tr>
<tr>
<td>425</td>
<td>29.19</td>
<td>24.66</td>
<td>52.01</td>
<td>47.48</td>
</tr>
<tr>
<td>450</td>
<td>31.17</td>
<td>26.64</td>
<td>53.99</td>
<td>49.46</td>
</tr>
<tr>
<td>475</td>
<td>33.15</td>
<td>28.62</td>
<td>55.97</td>
<td>51.44</td>
</tr>
<tr>
<td>500</td>
<td>35.13</td>
<td>30.60</td>
<td>57.95</td>
<td>53.42</td>
</tr>
<tr>
<td>600</td>
<td>35.13</td>
<td>30.60</td>
<td>57.95</td>
<td>53.42</td>
</tr>
<tr>
<td>700</td>
<td>35.13</td>
<td>30.60</td>
<td>57.95</td>
<td>53.42</td>
</tr>
<tr>
<td>800</td>
<td>35.13</td>
<td>30.60</td>
<td>57.95</td>
<td>53.42</td>
</tr>
<tr>
<td>900</td>
<td>35.13</td>
<td>30.60</td>
<td>57.95</td>
<td>53.42</td>
</tr>
<tr>
<td>1000</td>
<td>35.13</td>
<td>30.60</td>
<td>57.95</td>
<td>53.42</td>
</tr>
</tbody>
</table>
low solar activity. For instance, it reached 22% in 2008 (minimum of solar cycle 23) while 16% error is deduced in 2002 at 400 km (high solar activity). A 5% uncertainty variation is observed between low and high solar activity in agreement with the range of solar flux variations estimated in Figure 9, left side. We remark that this modulation pattern matches the thermospheric variations due to the solar cycle (Solomon et al., 2011).

Surprisingly, the uncertainty does not exhibit seasonal variations, as illustrated in Figure 7 (left frame). However, as shown in Figure 8, different variations on the 1-sigma error are perceived during low solar activity (minimum solar cycle in 2008) and high (2002).

The 1-sigma uncertainty is minimum at the lowest altitude during high solar activities (~4% for \( F_{10.7} = 200 \) at 200 km), while the maximum uncertainty of 30% is reached at the highest altitude during low solar activities \( (F_{10.7} = 65) \), as shown in Figure 9 (left frame). Whatever the altitude, we noted a 5% decrease in 1-sigma from low to high solar activities. As a final point, the error is significantly enhanced by increasing geomagnetic activity to storm level, as illustrated in Figure 9 (right frame). Indeed, although the 1-sigma uncertainty is minimum at the lowest altitude during quiet geomagnetic activities (~6% for \( K_p = 1 \) at 200 km), the maximum of 65% is obtained at 500 km altitude during extreme geomagnetic storms \( (K_p = 9) \).

We observed a clear transition at \( K_p = 5 \) (low/medium geomagnetic activity \( K_p < 5 \) and high geomagnetic activity \( K_p > 5 \)). Indeed, independent of altitude, the error increases by 6% at low/medium geomagnetic activities, while we observe a 30% uncertainty increase from \( K_p = 5 \) to \( K_p = 9 \) (high/severe magnetic activities). It is well-known that thermospheric density variations and uncertainties are enhanced in response to geomagnetic activity by the energy deposition from the solar wind and magnetosphere (Knipp et al., 2013; Bruinsma et al., 2021a).

We focus on the influence of geographical coordinates and solar/geomagnetic activities in Figure 10. The 1-sigma uncertainty is plotted as the latitude and local time function on the left side. A reverse saddle-like shape is observed. The maximum uncertainty occurs near the 24 LT in the pole region (~32%). In contrast, a minimum uncertainty of 20% is expected in the late afternoon near the equatorial zone. Therefore, the largest uncertainties are at high latitudes in the night/early morning (22:00–04:00 LT), whereas they are smallest in late morning and afternoon (10:00–18:00 LT) from the low to mid-latitudes. Figure 10 (right frame) shows that the 1-sigma uncertainty is minimum during high solar and low geomagnetic activity (~21% for \( F_{10.7} = 200 \) and \( K_p = 1 \)). The maximum uncertainty of 60% is observed for \( F_{10.7} = 65 \) and \( K_p = 9 \). A clear transition always appears at \( K_p = 5 \). Moreover, independent of solar activity, an increase of 5% occurs at low geomagnetic activity \( K_p < 5 \) whereas the uncertainty increases by a factor of 30% during high magnetic activities.

To conclude this section, we underline the chief impact of the altitude on the 1-sigma uncertainty (ranging from 10 to 30%) compared to the other parameters excepting high geomagnetic activity, i.e., \( K_p > 5 \) (from 30 to 80%). Latitudinal and diurnal effects on the density uncertainties are observed. Finally, the variation in solar activity over a cycle impacts the neutral density uncertainty even if its amplitude is rather modest (variation in the uncertainty of 5%) and on a long time scale.

7 Validation using HASDM uncertainty model

HASDM densities at 250, 325, 400, and 475 km covering almost two solar cycles have been used to assess our uncertainty model developed with accelerometer-inferred neutral density data (called UM in the following). For that, we used the same process to construct an uncertainty model but with HASDM (called HUM in the following) by binning the vast amount of data (more than 15 GB) to deduce a 6-D polynomial function in the altitude range 200–500 km. The comparisons are shown
in Figures 11–13 (left and right frames for the UM and HUM model, respectively). The two models present overall the same tendencies. For instance, they exhibit similar cubic shapes as a function of altitude and the local hour in Figure 11 (bottom), although a delay of two hours is observed on the maximum (15 LT for UM and 17 LT for HUM models). Besides, the long-term variation due to the solar cycle is well reproduced by the two models, as illustrated in Figure 12 (bottom frames).
The 1-sigma error, as a function of altitude and the magnetic index, presents the same shape with a transition at $K_p = 5$, as remarked in Figure 13 (bottom frames). One of the main features is an overall underestimation by nearly 5% of the HUM model. This difference increases significantly with the $K_p$ index and reaches 20% at the highest altitude (Fig. 13 – bottom frames). We remind that the datasets have different sampling rates (and so also different spatial resolutions). The accelerometer measurements have a 10s temporal resolution, while HASDM’s is 3 h, which explains lower uncertainties due to the smoothing effect.

However, some differences need to be clarified. As shown in Figure 11 (top frames), latitude effects are not visible in the HUM model. This is due to the much lower spatial and temporal resolution of HASDM data, which does not capture the rapid and short spatial-scale density variations, especially at auroral latitudes. Then, the HUM reveals seasonal effects not observed in the UM model (Fig. 10, top frames). Indeed, the
UM model presents a constant 1-sigma error at any given altitude, whereas the HUM uncertainty exhibits seasonal variation. The principle maximum is just after the June summer solstice (~24% at 450 km), and the secondary maximum is just after the winter solstice (~22% at 450 km). The absence of this variation in UM is most likely due to its much sparser database, but the amplitude of the variation around the mean uncertainty is only ±2%. Finally, a linear decrease as a function of solar flux is modeled with UM, whereas HUM is nearly constant on an altitude level (Fig. 13, top frames). We inferred that it is due to a DTM2020 biases effect during low solar activities (Bruinsma & Boniface, 2021) that overestimated the uncertainty value of UM model at low solar activities \((F_p < 150)\).

To summarize, accelerometers deliver the most precise and high-resolution measurements to infer neutral density. Comparison with HUM confirms that the UM model provides realistic estimates of the 1-sigma uncertainty.

8 Conclusion

We developed a semi-analytical model in 6-D space parameters using the DTM2020 thermosphere model that characterizes the uncertainty of its neutral density estimates. The model is constructed with density data from 200 to 500 km altitude, but it was extended to cover 100–1000 km by making presumptions for altitudes without data coverage. The model is founded on a new combined statistical data binning and local least square fitting approach. Because the distribution of the high-resolution density data is relatively sparse, the model was successfully validated in the 200–500 km altitude range using the recently released SET HASDM density data.

The principal results of the model are:

- the main parameter for the uncertainty during low and moderate magnetic activity \((K_p \leq 5)\) is altitude;
- the largest uncertainty is reached during severe or extreme geomagnetic storms;
- the solar cycle causes a long-term modulation of uncertainty;
- the model predicts latitudinal and diurnal effects on the density uncertainties.

Further study is needed to validate the uncertainty model in the extended range (100–1000 km), but that depends on future data availability.

Acknowledgements. The SWAMI project has been granted from the European Union’s Horizon 2020 research and innovation program under grant agreement No. 776287. CNES APR equally supports SB and CB. The HASDM database is made available by SET (https://spacewx.com/hasdm/). GOCE and Swarm density data is made available by ESA (https://earth.esa.int/eogateway/). CB would like to acknowledge Julien Laurent-Varin for their fruitful discussions on the mathematical method used in this work. The DTM models and the density uncertainty estimate model are available on Github via the SWAMI website (http://swami-h2020.eu). The editor thanks David Vallado and an anonymous reviewer for their assistance in evaluating this paper.

References


