



**HAL**  
open science

# The operational and research DTM-2020 thermosphere models

Sean Bruinsma, Claude Boniface

► **To cite this version:**

Sean Bruinsma, Claude Boniface. The operational and research DTM-2020 thermosphere models. Journal of Space Weather and Space Climate, 2021, 11, pp.47 - (1-15). <10.1051/swsc/2021032>. <hal-03337389>

**HAL Id: hal-03337389**

**<https://hal.science/hal-03337389v1>**

Submitted on 7 Sep 2021

**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.



HAL Authorization

# The operational and research DTM-2020 thermosphere models

Sean Bruinsma\* and Claude Boniface

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

Received 11 June 2021 / Accepted 23 August 2021

**Abstract**—*Aims:* The semi-empirical Drag Temperature Models (DTM) predict the Earth's thermosphere's temperature, density, and composition, especially for orbit computation purposes. Two new models were developed in the framework of the H2020 Space Weather Atmosphere Models and Indices (SWAMI) project. The operational model is driven by the trusted and established F10.7 and  $Kp$  indices for solar and geomagnetic activity. The so-called research model is more accurate, but it uses the indices F30 and the hourly H<sub>po</sub>, which are not yet accredited operationally. *Methods:* The DTM2020 models' backbone comprises GOCE, CHAMP, and Swarm A densities, processed by TU Delft, and Stella processed in-house. They constitute the standards for absolute densities, and they are 20–30% smaller than the datasets used in the fit of DTM2013. Also, the global daily mean TLE densities at 250 km, spanning four solar cycles, were now used to improve solar cycle variations. The operational model employs the same algorithm as DTM2013, which was obtained through fitting all data in our database from 1967 to 2019. Because of the H<sub>po</sub> index, which is not available before 1995, the coefficients linked to the geomagnetic activity of the research model are fitted to data from 2000 to 2019. The algorithm was updated to take advantage of the higher cadence of H<sub>po</sub>. Both models are assessed with independent data and compared with the COSPAR International Reference Atmosphere models NRLMSISE-00, JB2008, and DTM2013. The bias and precision of the models are assessed through comparison with observations according to published metrics on several time scales. Secondly, binning of the density ratios are used to detect specific model errors. *Results:* The DTM2020 densities are on average 20–30% smaller than those of DTM2013, NRLMSISE-00, and JB2008. The assessment shows that the research DTM2020 is the least biased and most precise model compared to assimilated data. It is a significant improvement over DTM2013 under all conditions and at all altitudes. This is confirmed by the comparison with independent SET HASDM density data. The operational DTM2020 is always less accurate than the research model except at 800 km altitude. It has comparable or slightly higher precision than DTM2013, despite using F10.7 instead of F30 as solar activity driver. DTM, and semi-empirical models in general, can still be significantly improved on the condition of setting up a more complete and consistent total density, composition, and temperature database than available at this time by means of a well-conceived observing system.

**Keywords:** Thermosphere / drag

## 1 Introduction

A major application of semi-empirical thermosphere specification models is in the computation of the atmospheric drag force in the orbit determination and prediction of spacecraft as well as debris. The models provide low spatial and temporal resolution average (climatological) predictions of the main constituents' temperature, total and partial densities as a function of location (altitude, latitude, longitude, local solar time), solar and geomagnetic activity, and season. Their accomplishment depends largely on the quality and distribution of the available

atmospheric observations. They reproduce with variable accuracy through a simplified analytical algorithm based on parameterized equations and notably low degree and order spherical harmonic expansions. The spatial resolution that can be reached with the degree and order 3 or 4 expansions is typically 45–60 arc degrees. Consequently, smaller density variations (e.g., gravity waves, certain tides) cannot be reproduced, and they are sources of geophysical noise when comparing to observations.

The thermosphere operational and research Drag Temperature Model (DTM2020) models were developed in the framework of the Space Weather Atmosphere Models and Indices (SWAMI) project (<http://swami-h2020.eu>), which was

\*Corresponding author: [sean.bruinsma@cnes.fr](mailto:sean.bruinsma@cnes.fr)

a European Union Horizon 2020 Framework project. The main objectives of the SWAMI project were updating the DTM2013 thermosphere model (Bruinsma, 2015), extending the upper bound of the altitude range covered by the Met Office Unified Model (MO-UM), creation of a whole atmosphere model by blending of MO-UM output with DTM2020, and taking advantage of the new high-cadence driver for geomagnetic activity Hpo (Jackson et al., 2020).

The DTM2020 and COSPAR International Reference Atmosphere (CIRA) models described and assessed in this paper are constructed by optimally fitting model coefficients to different combinations of density, temperature, and composition measurements. The backbone of the DTM2020 models is the complete Challenging Minisatellite Payload (CHAMP; Doornbos, 2011), Gravity Recovery and Climate Experiment (GRACE; Bruinsma, 2015) and Gravity field and steady-state Ocean Circulation Explorer (GOCE; Doornbos et al., 2014) high-resolution accelerometer-inferred density datasets, as well as Swarm A densities (Van den Ijssel et al., 2020). A major difference with DTM2013 is that the density datasets were not adjusted to the US Space Force High Accuracy Satellite Drag Model (HASDM; Storz et al., 2005) by scaling factors described for GOCE data in (Bruinsma et al., 2014). Instead, CHAMP, GOCE, and Swarm A densities were considered of higher absolute accuracy and used unmodified. Consequently, DTM2020 densities are on average 20–30% smaller than DTM2013 (and JB2008); these smaller absolute densities are consistent with NRLMSIS 2.0 results, which shows the same tendency when compared to NRLMSISE-00.

Another difference with DTM2013 is that two versions of DTM2020 were developed. The proxy for solar Extreme Ultra Violet (EUV) emissions, the solar radio flux at 30 cm F30 (Dudok de Wit et al., 2014), is not yet universally accepted in an operational environment as F10.7 is. The high-cadence Hpo geomagnetic indices (<https://doi.org/10.5880/Hpo.0001>) have been developed very recently and for that reason maybe also not yet be accredited for operational use. Therefore, it was decided to develop an operational model with F10.7 and  $Kp$  as drivers (hereafter referred to as DTM2020\_Oper), and a more accurate but research model driven by F30 and Hp60 (DTM2020\_Res). DTM2020, as well as the CIRA models for atmospheric drag NRLMSISE-00 (Picone et al., 2002), JB2008 (Bowman et al., 2008), DTM2013, and NRLMSIS 2.0 (Emmert et al., 2020), are assessed by comparison to total density data to demonstrate and quantify the improvements made.

The following section briefly reviews the density data and drivers used in this study. The model algorithm is described in Section 3. The assessment is presented in Section 4, and a summary and conclusions are given in Section 5.

## 2 Density data and drivers

The neutral density data used in the construction and validation of DTM2020 are listed in Table 1. The distributions with time (solar activity) and altitude of the main total mass density datasets are displayed in Figure 1, and they are described in more detail in the following subsections, while the last two subsections succinctly describe the new proxies used in DTM2020\_Res (F10.7 and  $Kp$  are not presented).

### 2.1 GOCE and CHAMP

GOCE (Drinkwater et al., 2003) was launched in March 2009 in a 96.5° inclination, dawn-dusk orbit, and re-entered the atmosphere in November 2013. Thanks to drag compensation, the orbit was maintained at 255 km mean altitude for the largest part of the mission, and then it was lowered in four stages, ultimately to 224 km in May 2013. GOCE neutral density is an ESA product (<https://earth.esa.int/eogateway/catalog/goce-thermosphere-data>) with a resolution of 80 km along the orbit and a precision of a few percent or better (Bruinsma et al., 2014). For DTM2020, GOCE data were used as provided by ESA. The processing of GOCE data (by TU Delft under ESA contract) was done using high-fidelity satellite shape and numerical aerodynamic coefficient models, which lead to the most accurate absolute density. In the past, for DTM2009 and DTM2013, density data were adjusted to the High Accuracy Satellite Drag Model (HASDM; Storz et al., 2005), via multiplicative scale factors (e.g., 1.25 for GOCE; Bruinsma et al., 2014) for consistency, compatibility, and comparability. HASDM is an operational model that is corrected in near real-time through data assimilation.

The CHAMP (Reigber et al., 1996) data were also processed using high-fidelity satellite shape and aerodynamic coefficient models by TU Delft (Doornbos, 2011). The scale factor to best fit the HASDM densities along the CHAMP orbit is 1.23, which was applied in the construction of DTM2009 and DTM2013. The absolute densities predicted with a model are linked to the aerodynamic coefficients used in the derivation of atmospheric density. Unfortunately, there is no standard for aerodynamic coefficient modeling, but DTM2020 was constructed with the most accurate data available. Since the GOCE and CHAMP data are not scaled in the DTM2020 development, it is clear that DTM2020 densities will be 20–25% smaller than DTM2013.

### 2.2 Swarm A

The ESA Swarm mission (Friis-Christensen et al., 2008), launched in November 2013, consists of three identical satellites (A and C side-by-side, and B in the higher orbit) in near-polar orbits at about 460 and 510 km initially. They are equipped with accelerometers, but in this study, the GPS-derived densities (Van den Ijssel et al., 2020) are used because of the serious problems and data gaps in the accelerometer data (Siemes et al., 2016). The spatial resolution and precision depend on the level of atmospheric drag experienced in orbit, principally on altitude and solar activity. Only the density data of Swarm A (<http://swarm-diss.eo.esa.int>), one of the lower flying satellites, for high to medium solar activity (a cut-off of F10.7 less than 100 sfu was applied) was used in DTM2020; the spatial resolution of the selected data is estimated to be approximately 4000–8000 km along the orbit, i.e., comparable to the model resolution.

### 2.3 GRACE

Neutral densities inferred from accelerometer measurements on GRACE (Tapley et al., 2004) were computed using the methodology described in Bruinsma et al. (2004). Compared with the TU Delft processing, the satellite shape was a simple

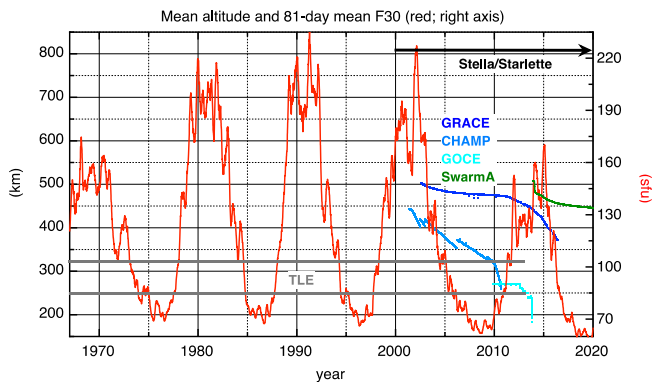
**Table 1.** Data used in the construction DTM2020 ( $\rho$  is total density), and assessment of the CIRA models.

	Measurement	Period	Used in DTM2020
CHAMP	$\rho$	01/2001 – 09/2010	Operational, research
GRACE	$\rho$	08/2002 – 12/2016	Operational, research
GOCE	$\rho$	11/2009 – 11/2013	Operational, research
Swarm A	$\rho$	11/2013 – 02/2016	Operational, research
Stella	$\rho$	01/2000 – 12/2019	Operational, research
Starlette	$\rho$	01/2000 – 12/2016	No
TLE 250/325 km	$\rho$	01/1967 – 08/2013	Operational, research*
HASDM database	$\rho$	01/2000 – 12/2019	No
SETA2	$\rho$	05/1982 – 11/1982	No
SETA3	$\rho$	07/1983 – 03/1984	No
AE-C MESA	$\rho$	12/1973 – 09/1976	No
AE-E MESA	$\rho$	12/1975 – 01/1978	No
DE-2**	T, He, O, N <sub>2</sub>	08/1981 – 02/1983	Operational
AE-C	N <sub>2</sub>	01/1974 – 04/1977	Operational
AE-E	T, O	12/1975 – 05/1981	Operational
EDR	$\rho$	1/2000 – 12/2010	No
SUSIM***	O <sub>2</sub>	10/1991 – 2/2005	Operational

\* 2000–2013;

\*\* Dynamics Explorer;

\*\*\* Solar Ultraviolet Spectral Irradiance Monitor.

**Fig. 1.** The distribution of the total density data in time, mean solar activity, and mean altitude.

macro model, and the analytical Sentman aerodynamic coefficient model (Sentman, 1961) was employed. The GRACE data cover the period 08/2002 through 12/2016 but starting in 2011, 2–3 gaps of 4 weeks each per year occurred, and data quality is slightly worse due to battery issues on the spacecraft. To make the GRACE data more consistent with the more accurate TU Delft densities, scaling factors were determined through comparison, after normalization to the mean altitude, with CHAMP and Swarm A densities when the orbital planes were nearly co-planar (less than 1-hour local time difference). The GRACE density scaling factors thus determined a decrease over the mission: 0.76 (2002–2005), 0.73 (2006), 0.70 (2007–2016).

#### 2.4 TLE daily-mean global densities

The daily-mean global average mass densities at 250 and 325 km altitude derived from Two-Line orbital Element (TLE) sets covering the years 1967–2013 (Emmert, 2015) were

used or partly used (from 2000 to 2013) in the case of the research model. They are at the low end of the ten altitude levels (250–575 km) on which the TLE densities were computed; the higher altitude data were not used because precision decreases and noise grows rapidly. Although data is provided with daily cadence, the temporal resolution is less because of the 3-day cubic spline smoothing involved in the density retrieval procedure (Emmert, 2009). Even if the noise is much larger than that in GOCE data at 250 km, the TLE dataset is essential because it covers four solar cycles. Because the scale of the TLE data is not consistent with the accelerometer-inferred and Swarm A densities, this dataset is used in the adjustment procedure to estimate relative density variations and not the absolute value. Consequently, a model bias with respect to this dataset is to be expected.

#### 2.5 Stella and Starlette

The spherical satellites Stella and Starlette are in a 96° inclination and near-circular orbit at approximately 813 km altitude, and 49° inclination slightly eccentric orbit ( $e = 0.02$ ) with a perigee altitude of 800 km, respectively. Because of their shape (no attitude-related errors), knowledge of the satellite characteristics (mass, surface, reflectivity), and the very accurate laser tracking by the International Laser Ranging Service (ILRS; Pearlman et al., 2002), they are suitable for density derivation despite their relatively high altitude. The densities were inferred from the analysis of orbit perturbations (Jachia & Slowey, 1963), which essentially ties the observed decay to a mean density over the orbit in the case of Stella and mean density close to perigee for Starlette and span the period 2000–2019. Starlette was not used in the construction of DTM2020 and was reserved for validation. The scale of these two satellites is also considered accurate because the analytical closed form equations used are actually as accurate as the numerical models for spheres.

## 2.6 SET HASDM density database

Space Environment Technologies (SET) has made 20 years of densities available 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). These densities are not observations but the output of a model corrected in near real-time through data assimilation. Densities are given every 3 h from 2000 to 2019 on  $[10^\circ \text{ latitude} \times 1 \text{ h local time}]$  grids, and from 175 to 825 km altitude in steps of 25 km. The SET HASDM data on the altitudes 250, 400, 475, and 800 km are only used for model validation for the entire 20-year time span.

The scale of the SET HASDM data (and the data in the next section) is not consistent with the accelerometer-inferred and Swarm A densities. A model bias concerning this dataset is, therefore, to be expected.

## 2.7 Energy Dissipation Rate (EDR) densities

Daily-mean densities inferred from satellite drag data from the Cosmos satellites, Venuslander, Ov-3, and Hitchhiker-1, all in elliptical orbits with perigee heights in the 200–500 km range, are also reserved for assessment only. The densities, covering the period 2000–2010, were computed from the estimated energy dissipation rates derived from the orbits fitted directly to radar and optical tracking data. For a complete description of the method and the data, we refer to (Bowman et al., 2004).

## 2.8 The F30 radio flux

DTM2013 and DTM2020\_Res are driven by F30 because the observed densities can be reconstructed with higher fidelity than with F10.7, most likely because of the much higher contribution of thermal Bremsstrahlung (Dudok de Wit et al., 2014). Similar to F10.7, it is a ground-based radio measurement, and therefore more robust and more accurately calibrated than satellite measurements of solar emissions (Vourlidis & Bruinsma, 2018). Still, the absolute calibration also of ground-based radio instruments remains a challenging issue. And although F30 is a better proxy for solar EUV than F10.7, it is still a proxy.

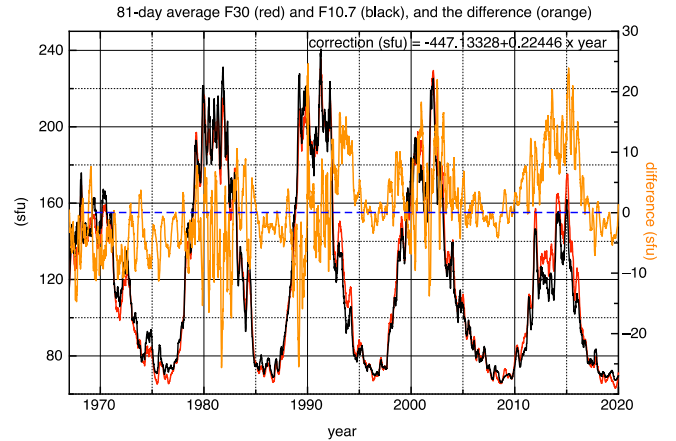
For convenience and to facilitate mission design and other simulations, F30 was used in DTM2013 after scaling to F10.7 via the linear regression shown in eq. (1):

$$F30_{\text{rescaled to F10.7}} = -1.5998 + 1.553755 \times F30_{\text{observed}}. \quad (1)$$

During the development of DTM2020, problems with fitting the density data over the complete 1967–2019 interval led to a renewed inspection of F30. The difference with F10.7 revealed a positive trend, as can be seen in Figure 2. Correcting F30 for the drift of about 10 sfu over the 53 years using the linear regression formula printed in Figure 2 resolved the density data fitting issue. The drift is most likely due to imperfect calibration of F30. In the following, and different from DTM2013, the drift-corrected F30 given below is used:

$$F30_{\text{rescaled to F10.7}} = -1.5998 + 1.553755 \times F30_{\text{observed}} - 447.13328 + 0.22446 \times \text{year} \quad (2)$$

in which year is the decimal calendar year (e.g., 2010.4928).



**Fig. 2.** The 81-day mean observations of the solar activity proxies F30 and F10.7, and the difference F30 minus F10.7 (orange; right axis).

## 2.9 The Hp60/ap60 hourly planetary geomagnetic index

The research DTM2020 model uses a new driver for the geomagnetic activity of the Hpo index family (<http://swami-h2020.eu/wp-content/uploads/2021/04/AGU-iPoster-Matzka.pdf>) with higher temporal resolution than  $Kp$ , the hourly Hp60 index (<https://doi.org/10.5880/Hpo.0001>). Up to values of  $Kp = 9$ –, it employs the same (operational) algorithm used to derive the  $Kp$  index and is conceived to have the same frequency distribution, i.e., the same number of quiet days and geomagnetic storms. Compared with the  $Kp$  algorithm, another essential improvement besides the higher resolution is that it is no longer capped at 9 (all geomagnetic events above a certain threshold are  $Kp = 9$ ). The Hp60 index is open-ended, and very high levels of geomagnetic activity can now be more accurately classified. For instance,  $Kp$  reached a value of 9 during the Halloween storm in 2009, while the Hp60 reaches a value of 12. However, it is not available before 1995 because it requires digital high-cadence magnetograms.

## 3 The development of DTM2020

### 3.1 The algorithms

In DTM, total density in the altitude range from 120 to approximately 1500 km is reconstructed by summing the partial densities of the main thermosphere constituents ( $N_2$ ,  $O_2$ ,  $O$ ,  $He$ ,  $H$ ), under the hypothesis of independent static diffuse equilibrium. The total mass density  $\rho$  at altitude  $z$  is calculated as follows:

$$\rho(z) = \sum_i \rho_i(120 \text{ km}) f_i(z) \exp(G_i(L)). \quad (3)$$

The partial densities of the  $i$ th constituent,  $\rho_i$ , are specified at 120 km altitude and are propagated to higher altitudes employing their specific height function  $f_i(z)$ , which is based on temperature. Variations in the temperature at 120 km, the exospheric temperature, and the partial densities are modeled using the

spherical harmonic function  $G(L)$ , which takes the effect of environmental parameters  $L$  (latitude, longitude, local solar time, season, solar flux, and geomagnetic activity) into account.

The function  $G(L)$  used in DTM2013 and DTM2020\_Oper is identical, and the model improvement is due to fitting to new data only. For the research model, it was updated to take advantage of the higher cadence of the Hp60 index. The modifications made were both logistic, to handle a larger number of indices, and algorithmic. Presently ten values of Hpo are used: Hpo delayed by 0–4 h (5 indices), the mean of the previous 24 h, and four mean values of three indices delayed by 6 (mean of Hpo delayed by 5, 6, and 7 h), 10, 15, and 20 h. The delay applied to the five Hp60 indices is a function of latitude, and it varies from 1-h at the pole to 4-h at the equator. The above delays and delayed mean values were determined by trial and error on a selection of storms. A further improvement is achieved by coupling the amplitude of the density enhancement due to geomagnetic activity with the mean solar activity.

### 3.2 Sequential modeling procedure

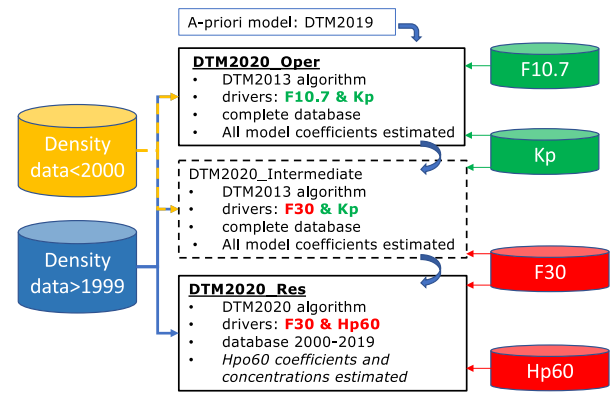
The DTM2020 models were constructed in a sequential procedure because Hp60 was unavailable before 1995, i.e., not for the entire density database (Table 1). In a preliminary step of the procedure, using DTM2013 as a background model, the temperature and partial densities were re-estimated via least-squares adjustment to better fit the approximately 20–25% smaller observed densities than those used in DTM2013. That unpublished model, DTM2019, was the prior model used in the sequential development depicted in Figure 3. The first model developed was DTM2020\_Oper, which is driven by F10.7 and  $Kp$ , and therefore all density data from the sixties to the present-day can be assimilated. Then in a second step, using DTM2020\_Oper as background, DTM2020\_Intermediate was constructed with F30 and  $Kp$ . Since F30 covers the entire period for which density data are available, all solar activity-related model coefficients are accurately estimated. The DTM2020\_Res model is developed in the third step, using DTM2020\_Intermediate as a-priori, and only re-estimating those model coefficients linked to Hp60 using the recent high-resolution density data (2000–2019), thereby preserving the accuracy of the coefficients linked to solar activity.

### 3.3 The density uncertainty estimate

An optional new feature of DTM2020 is the  $1-\sigma$  uncertainty estimate for the density given as a percentage. The uncertainty model was constructed by binning and analyzing all accelerometer-inferred densities and Swarm A densities using manifold approximation, described in the companion paper by Boniface and Bruinsma (2021)

## 4 Results

All models are evaluated by computing the mean and standard deviation (StD) of the density ratios, defined as



**Fig. 3.** Schematic representation of the sequential development of the DTM2020 models.

“observed-over-calculated” or “O/C”, in natural log space ( $\ln O/C$ ). The metrics are described in (Bruinsma et al., 2018). The mean of the density ratios represents the bias of a model with respect to the dataset used in the assessment. The model bias thus defined is not an entirely objective quality indicator, and one must consider whether the dataset was ingested into the thermosphere model, and if that was done in an absolute or relative manner. Also, the processing standards applied in the production of the dataset, in particular, the satellite shape and aerodynamic coefficient, because the absolute values of the densities depend on them. The standard deviation of the density ratios, computed as a percentage of the observation, represents a combination of the ability of the model to reproduce observed density variations and (geophysical and instrumental) noise in the observations. The signal-to-noise ratio of the density data is inversely proportional to the density level, i.e., it increases with altitude.

In the following sections, the relative precisions of the models are evaluated by comparing to the complete datasets in terms of the mean and StD of the density ratios computed with all observations of a dataset, on a yearly and daily basis. The yearly statistics are computed by comparing to all data per calendar year, i.e., computing one density ratio per year. The StD is then computed with the yearly density ratios in log space. The daily statistics are obtained similarly. More specific information on errors can be obtained by binning the density ratios in latitude, local time, 81-day mean F10.7/F30, and  $Kp$ . Analysis of the mean and StD of the ratios may give clues about specific modeling errors or data inconsistencies. To eliminate obvious outliers, density ratios are rejected in all assessments if they do not lie in the interval from 0.4 to 2.5, and 0.333 to 3.0 for TLE data.

The density datasets may be inconsistent due to errors and approximations in the satellite geometry models, the aerodynamic force (drag and lift) coefficients, and the surface-to-mass ratios used in their derivation. This will cause the model to be biased concerning those datasets. This inconsistency will diminish when a more physically accurate calculation for the drag force is developed and applied consistently, and all datasets reprocessed using such a standard. However, it is a vast topic for which no consensus is reached, and it cannot be addressed in this paper.

**Table 2.** Mean and StD, the latter expressed as percentage (smallest values within 0.5% in bold), of the density ratios of the models comparing to full datasets. “DTM2020 Intermed” is the intermediate model displayed in Figure 3.

	MSIS00	JB2008	DTM2013	DTM2020 Oper	DTM2020 Intermed	DTM2020 Research	MSIS2.0
CHAMP	0.69/23.1	0.78/20.2	0.76/19.8	0.98/19.0	1.00/ <b>18.6</b>	0.98/ <b>18.4</b>	0.79/25.0
GRACE	0.75/28.8	0.83/26.9	0.76/ <b>23.2</b>	1.04/25.9	1.02/23.9	0.99/ <b>23.1</b>	0.80/31.9
GOCE	0.82/14.3	0.81/12.4	0.79/11.6	0.99/10.5	0.98/ <b>9.7</b>	1.00/ <b>9.5</b>	0.95/15.7
GOCE reentry	0.92/10.4	0.81/10.2	0.84/11.5	1.04/ <b>8.3</b>	1.04/ <b>8.7</b>	1.06/8.9	1.16/10.8
Swarm A	0.89/25.8	0.92/21.2	0.76/21.6	1.13/22.4	1.02/ <b>19.3</b>	1.00/ <b>19.1</b>	0.99/25.6
Stella	0.86/21.2	0.93/55.2	0.76/20.4	1.01/20.8	1.00/20.3	0.97/ <b>19.5</b>	0.85/21.7
TLE 250 km	0.99/19.0	**	1.01/22.9	1.21/18.1	1.23/ <b>18.2</b>	**	1.13/20.3
TLE 325 km	1.00/ <b>31.1</b>	**	1.09/36.4	1.35/33.2	1.34/32.0	**	1.13/32.2
TLE 250 km*	0.94/20.4	0.98/20.5	0.93/21.4	1.14/18.8	1.17/ <b>18.4</b>	1.17/19.4	1.06/21.7
TLE 325 km*	0.94/ <b>33.5</b>	1.02/34.0	0.99/35.1	1.29/34.4	1.29/ <b>33.1</b>	1.27/34.3	1.06/34.5
Below: not used in DTM2013 and DTM2020							
SETA2	0.97/ <b>13.2</b>	**	0.97/21.3	1.16/14.4	1.26/15.6	**	1.18/13.9
SETA3	0.91/16.8	**	0.88/17.7	1.06/17.2	1.11/ <b>16.2</b>	**	1.10/23.3
AE-C MESA	1.08/ <b>19.5</b>	**	1.11/22.3	1.24/21.0	1.30/22.4	**	1.26/22.4
AE-E MESA	1.06/ <b>18.1</b>	**	1.13/20.8	1.20/21.4	1.23/19.3	**	1.22/20.5
Starlette	0.90/ <b>18.5</b>	0.92/39.3	0.78/20.0	1.04/ <b>18.9</b>	1.04/19.1	1.01/19.5	0.90/19.9
HASDM 250 km	0.95/17.0	1.04/19.8	0.95/12.5	1.17/13.0	1.22/ <b>11.6</b>	1.23/ <b>11.9</b>	1.10/17.9
HASDM 400 km	0.87/27.4	0.99/20.5 <sup>†</sup>	0.95/18.9	1.26/18.8	1.28/ <b>18.4</b>	1.25/ <b>18.0</b>	0.97/27.3
HASDM 475 km	0.83/33.1	0.99/24.2 <sup>†</sup>	0.90/ <b>23.2</b>	1.21/23.6	1.24/23.5	1.20/ <b>22.8</b>	0.92/33.4
HASDM 800 km	0.99/50.4	1.07/24.2 <sup>†</sup>	0.88/49.0	1.09/48.9	1.09/49.0	1.07/48.5	1.01/46.9
EDR: Cosmos	0.87/22.7	0.98/16.3	0.97/16.8	1.28/16.3	1.29/ <b>14.9</b>	1.26/ <b>14.7</b>	0.99/22.9
Venuslander	0.95/12.7	0.99/8.3	0.93/12.0	1.11/12.2	1.15/11.4	1.17/ <b>11.6</b>	1.08/14.0
Ov-3-1	0.87/20.3	0.98/14.4	0.97/15.5	1.26/14.3	1.27/ <b>13.3</b>	1.25/13.5	0.99/20.2
Hitchhiker-1	0.89/17.9	0.98/13.7	0.96/14.0	1.23/13.3	1.25/ <b>12.3</b>	1.24/ <b>12.4</b>	1.01/18.5

\* 2000–2013;

\*\* no indices;

† 7–8% data rejected.

#### 4.1 Model precision: overall

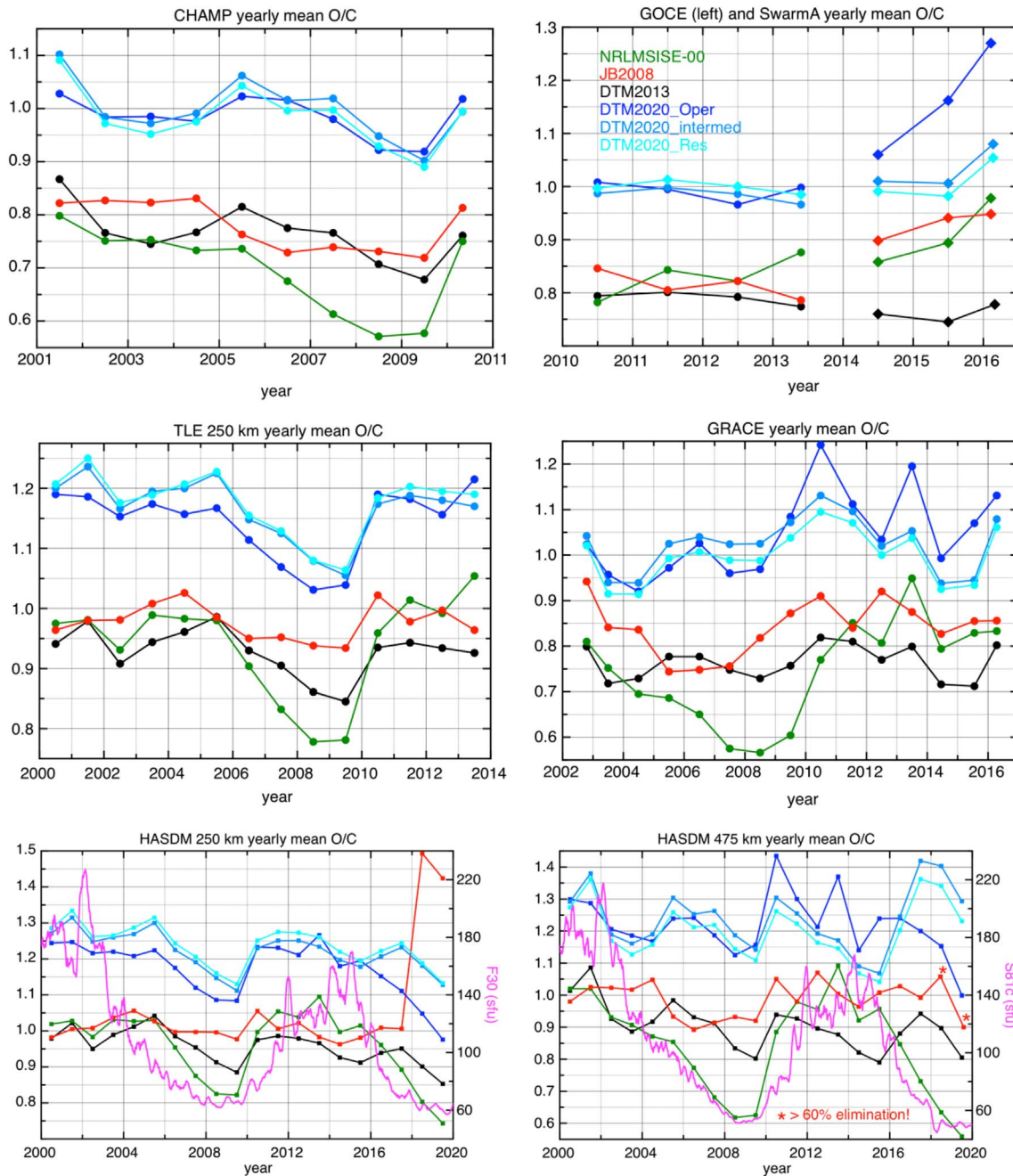
In this first overall assessment, the relative precisions of the models are calculated by comparing the full datasets in terms of mean and StD of the density ratios computed with all observations of the datasets in Table 1. The results, listed in Table 2, are also given for the intermediate DTM2020 model constructed with F30 and  $Kp$  to quantify the better performance of DTM2020\_Res compared to DTM2013 before 2000. Most of the DTM2020 intermediate and research model coefficients are identical, particularly those related to solar activity, and, therefore, the performance under pre-year 2000 conditions is nearly identical. The table is divided into two parts: the top part shows results of comparisons with ingested datasets, while the bottom part presents comparisons to datasets that were not used to construct the DTM models.

As explained in Section 2, in DTM2020, the CHAMP, GOCE, and Swarm A density datasets are presently ingested without scaling factors. The results in Table 2 are obtained using smaller densities, and the 0.2–0.3 bias in the MSIS00, JB2008, and DTM2013 density ratios attests to this. DTM2020 fits the data well, with mean density ratios very close to unity. The StD of the density ratios is the smallest (in bold) for the DTM2020 models. The biggest improvements are seen in comparisons with GOCE and Swarm A data, and the former densities at an approximate altitude of 250 km are now reproduced with less than 10% StD.

The TLE densities, on the contrary, fit NRLMSISE-00, JB2008, and DTM2013 well on average, whereas the DTM2020 models, as expected, underestimate them by 20–35%. The JB2008 and DTM2020\_Res models cannot be compared to the complete TLE datasets because the drivers are not available that far back in the past, and therefore, the comparison was made from 2000 to 2013. DTM2020 reproduces the TLE densities with significantly smaller StD than DTM2013.

The rightmost column of Table 2 shows NRLMSIS 2.0, which except for Swarm A, always has a slightly higher StD than NRLMSISE-00. This is not surprising since essentially new stratosphere and mesosphere data were ingested, but no new thermosphere data. However, one of the consequences of the assimilation of lower atmosphere data in NRLMSIS 2.0 is the lower density in the thermosphere. While still higher than DTM2020, it is important to note that the models converge independently to lower densities.

The low-altitude Satellite Electrostatic Triaxial Accelerometer (SETA) and Atmosphere Explorer Miniature ElectroStatic Accelerometer (AE MESA) densities (below 250 km, and down to 130 km) are presently more precisely reconstructed with DTM2020 than with DTM2013, and the smallest StD is even achieved with the intermediate DTM2020. The best results are obtained with NRLMSISE-00, which was fitted to those datasets. Note that the biases of NRLMSIS 2.0, from 10–26% (for AE-C MESA), are in rather good agreement with



**Fig. 4.** The yearly mean density ratios with CHAMP (top left), GOCE, and Swarm A (top right), TLE 250 km (center left), GRACE (center right), HASDM 250 km (bottom left), and HASDM 475 km (bottom right).

DTM2020. Starlette densities at 800 km altitude are well-modeled with DTM2020 and best with NRLMSISE-00; the StD of JB2008 is high, 39.3%, and this is due to the absence of predicting a winter Helium bulge (Keating & Prior, 1968). The StD with Stella data, also at about 800 km, understandably is high for the same reason.

The HASDM data from 2000 to 2019, selected at four altitudes (250, 400, 475, and 800 km) typically covering GOCE, CHAMP, GRACE, and Stella, is used in the assessment. The data was partly used (2001–2005) in the construction of

JB2008, which has small biases. Below 500 km, DTM2020 densities are 17–28% smaller on average, whereas DTM2013 is 5–10% overestimating. The NRLMSISE-00 and NRLMSIS 2.0 mean density ratios decrease considerably from 250 to 475 km. The StD is smallest with DTM2020, in line with the results for GOCE, CHAMP, and GRACE. The unexpectedly high StD with JB2008 and the high data rejection result from very bad predictions shown in the next section for 2018 and 2019. At 800 km, the results need some clarification. The latitude-local time structure of the HASDM densities at

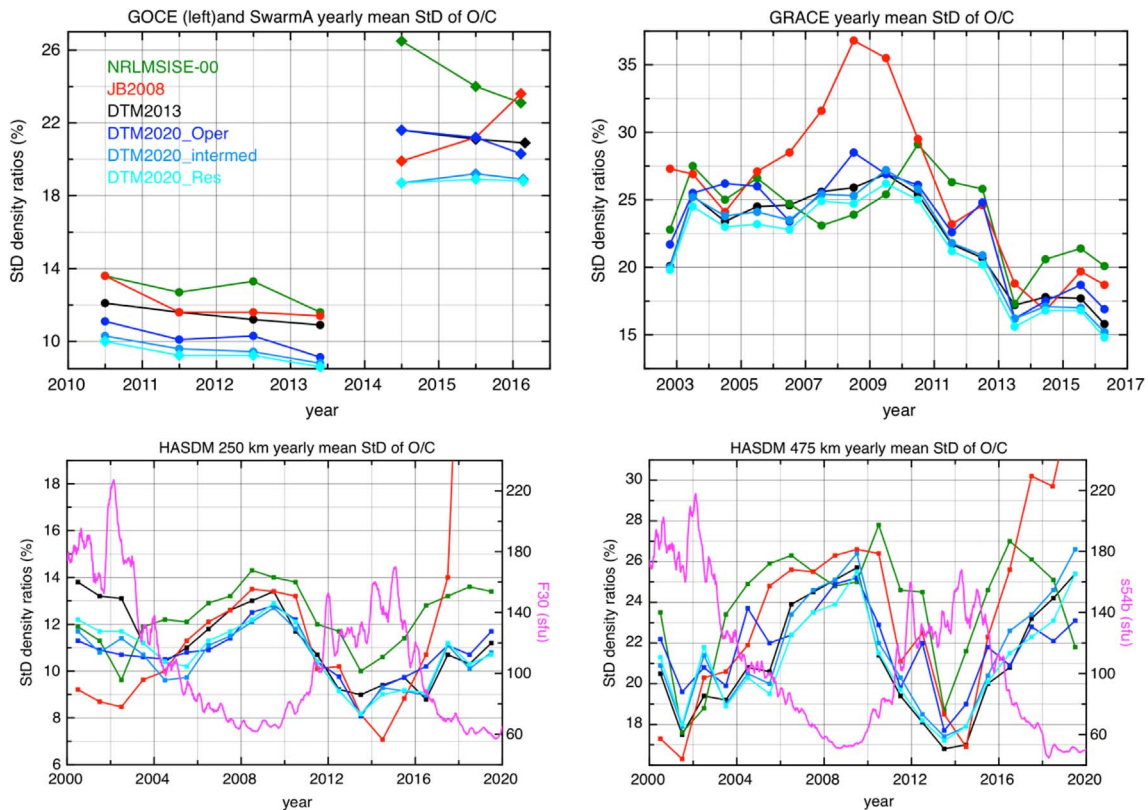
**Table 3.** StD (%; smallest values within 0.5% in bold) of the yearly mean density ratios of the models.

	MSIS00	JB2008	DTM2013	DTM2020 Oper	DTM2020 research
CHAMP	12.9	6.3	7.0	<b>4.1</b>	5.9
GRACE	16.7	7.4	<b>5.0</b>	9.1	5.9
GOCE	4.9	3.1	<b>1.5</b>	1.8	<b>1.2</b>
Swarm A	5.2	2.9	<b>2.2</b>	9.4	4.7
Stella	9.1	5.4 <sup>†</sup>	6.5	<b>4.1</b>	5.9
Starlette	7.1	<b>4.6</b> <sup>†</sup>	7.1	<b>4.9</b>	7.5
TLE 250 km*	10	<b>3.0</b>	4.4	5.4	4.7
HASDM 250 km	11.3	12.4 (2.6 <sup>+</sup> )	5.1	7.0	<b>4.7</b>
HASDM 400 km	18.1	5.0 <sup>†</sup> (4.5 <sup>+</sup> )	6.8	6.6	6.6
HASDM 475 km	21.5	5.9 <sup>†</sup> (5.5 <sup>+</sup> )	8.3	8.0	7.7
HASDM 800 km	<b>5.3</b>	11.6 <sup>†</sup>	12.3	7.2	9.2

\* 2000–2013;

+ without 2018–2019;

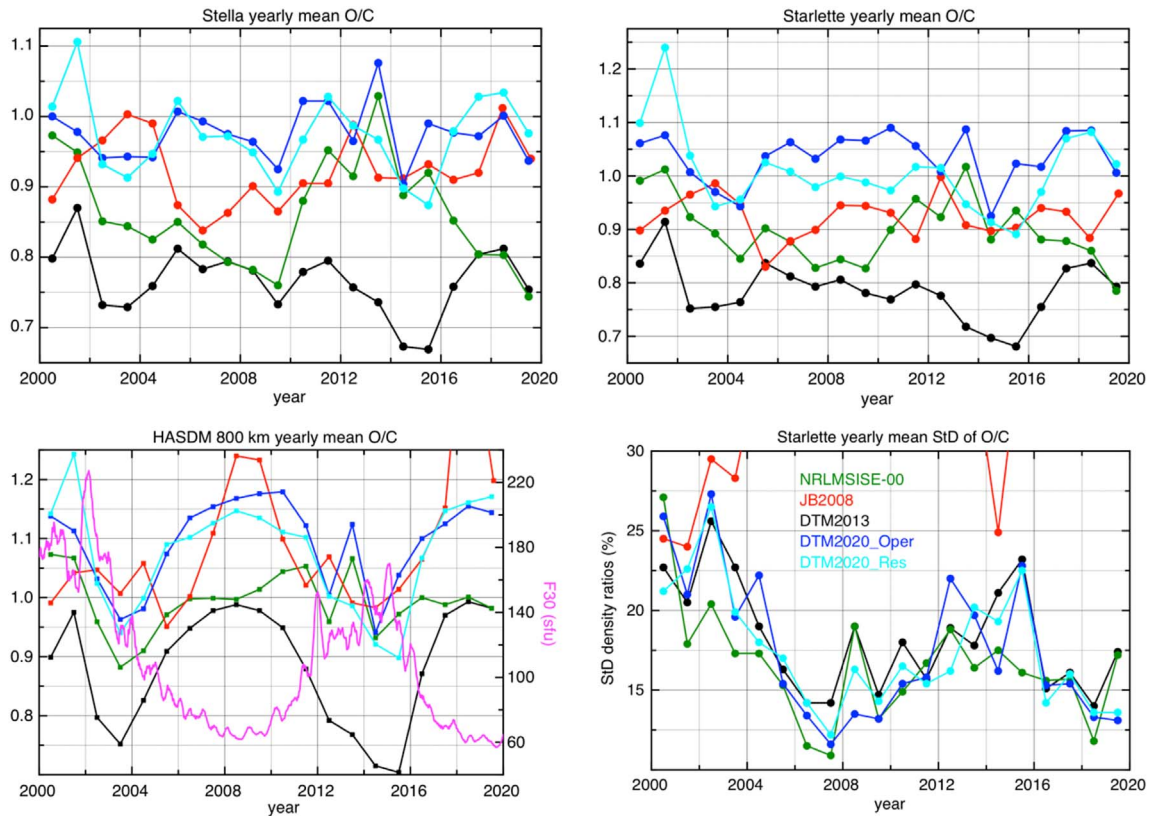
† 60% or more data rejected for 2018–2019.



**Fig. 5.** StD of the yearly mean density ratios with GOCE and Swarm A (top left), GRACE (top right), HASDM 250 km (bottom left), and HASDM 475 km (bottom right).

800 km does not agree with the MSIS and DTM models. Under solar minimum conditions, the winter Helium bulge (Keating & Prior, 1968) is absent in HASDM, which explains its absence in JB2008. The effect of Helium maximizing in the winter hemisphere, causing the total density maximum to be located in the winter hemisphere under (very) low solar activity conditions, can be directly observed even in Swarm and GRACE

densities. Therefore, when comparing to the MSIS and DTM models, which both predict maximum density in the winter hemisphere, the StD reflects this altered structure which explains why it is so large. However, the mean density (on time scales of years) should be correctly determined in HASDM. The NRLMSIS models have negligible biases, and the DTM2013 bias of 12% is much reduced now with DTM2020.



**Fig. 6.** The yearly mean density ratios with Stella (top left), Starlette (top right), HASDM 800 km (bottom left), and the StD of the density ratios with Starlette (bottom right). JB2008 for the years 2018 and 2019 again eliminates more than 60% of the data.

**Table 4.** StD (%; smallest values within 0.5% in bold) of the daily mean density ratios of the models.

	MSIS00	JB2008	DTM2013	DTM2020 oper	DTM2020 research
CHAMP	18.4	<b>11.6</b>	13.1	12.6	<b>12.1</b>
GOCE	9.9	6.8	6.3	6.7	<b>5.6</b>
TLE 250 km*	12.6	<b>9.7</b>	11.4	<b>10.0</b>	<b>10.0</b>
HASDM 250 km	13.9	17.4	8.6	10.0	<b>7.9</b>

\* 2000–2013.

Finally, the EDR densities are reproduced with biases ranging from 15 to 29% with DTM2020, but the StD is smallest and meaningful progress compared to DTM2013 has been made.

#### 4.2 Model precision: per year

The comparison in the previous section gives a good general impression of the model performance. Computing mean and StD of the density ratios per year makes more detailed analysis possible, particularly the performance over the solar cycle and as a function of altitude. The yearly mean density ratios are shown in Figure 4 for GOCE, CHAMP, GRACE, Swarm A, TLE 250 km, and HASDM at 250 and 475 km. The assessment is not done for NRLMSIS 2.0 because the results in Section 4.1

were shown only to place the lower DTM2020 densities in a present-day perspective. Observations before 2000 were not used because the JB2008 and DTM2020\_Res drivers are nonexistent.

The offset between DTM2020 and NRLMSISE-00, JB2008 and DTM2013 is evident in all frames of Figure 4 and reflects the 20–25% lower density estimates of the new models. Another salient feature common to all models is a dip, more or less deep, from 2007 to 2009 during the minimum transit from solar cycle 23 to 24. Due to the low solar EUV emissions, which were overestimated with all proxies, and most with F10.7 (Solomon et al., 2010). Consequently, all models overestimate density and NRLMSISE-00 most because it uses F10.7 and is not fitted to data in that period. F30, shown in the bottom left frame of Figure 4, also overestimated the

EUV flux during the last minimum, but it is smaller than F10.7. The dip is less obvious in JB2008 because one of its drivers, S10 and its 81-day centered mean S81c, is modified (decreased) for the solar minimum years to better-fit orbit data. A similar scheme seems to have been employed for the minimum years 2017–2019. But the extremely low values of S81c (less than 55 sfu) cause serious underestimation at 250 km so that applying our always identical thresholding in the comparison procedure led to more than 60% rejection of the 2018–2019 HASDM data at 475 km (also at 400 km; not shown). The JB2008 assessment presented in the DTM2013 paper (Bruinsma 2015), using the specific indices files made available, also revealed similar density under-estimation issues for 2010 and 2011.

In general, DTM2020 behaves well under all solar activity conditions without having recourse to adjusting the single solar activity driver F30. Below 400 km during solar minima, density is approximately 10% overestimated. This is a consequence of using a proxy for EUV emissions. The DTM2020\_Oper model struggles with GRACE and Swarm A and is not as stable over the years as DTM2020\_Res or DTM2013 because it uses the F10.7 proxy. The comparison with independent HASDM data confirms the good performance of DTM2020 and DTM2013 (which uses the F30 proxy as well). The StD of the yearly mean density ratios displayed in Figure 4 (i.e., of the 10 CHAMP density ratios, 4 GOCE density ratios, etc.) is listed in Table 3. The overall lowest StD, indicative of the smallest scatter of the density ratios from year to year, is achieved with DTM2020\_Res, which is slightly lower than DTM2013.

Figure 5 displays the StD per year for GOCE, Swarm A, and GRACE at the top, and HASDM at 250 and 475 km at the bottom. The StD in the top frames, which shows a comparison to ingested data, is nearly always smallest with DTM2020\_Res, and a significant improvement compared to DTM2013 is achieved. It also reveals a high StD of JB2008 for GRACE during solar minimum years 2007–2009, during which NRLMSISE-00 has the smallest StD, probably due to its more accurate modeling of Helium (consistent with Starlette results in Table 2). The HASDM results shown in the bottom frames of Figure 5 validate the low StD of GOCE with DTM2020. They also clearly show that StD and solar activity are inversely proportional. The external validation confirms that DTM2020 has the lowest StD for low to medium solar activity – which are the conditions for most ingested data – whereas JB2008 has the smallest StD under high solar activity (also for CHAMP in 2001; not shown), especially at 250 km. However, for the decaying phase of solar cycle 24 to the minimum from 2017–2019, the StD of JB2008 goes off the chart.

The assessment at 800 km altitude is done with Stella (ingested) and the external Starlette and HASDM at 800 km data. Figure 6 presents the yearly mean density ratios for Stella, Starlette, and HASDM at 800 km and the StD per year for Starlette only. In this comparison, DTM2020 agrees with Stella and Starlette within  $\pm 10\%$  except for 2001, which seems to be due to a problem in F30 (it is also visible in DTM2013). Comparison with HASDM shows a strong signal correlated with solar activity in JB2008 and the least in NRLMSISE-00. The StD of the yearly mean density ratios displayed in Figure 6 are also listed in Table 3; DTM2020 is considerably more accurate at 800 km than DTM2013. The yearly mean StD with Starlette confirms the highest precision of NRLMSISE-00 at

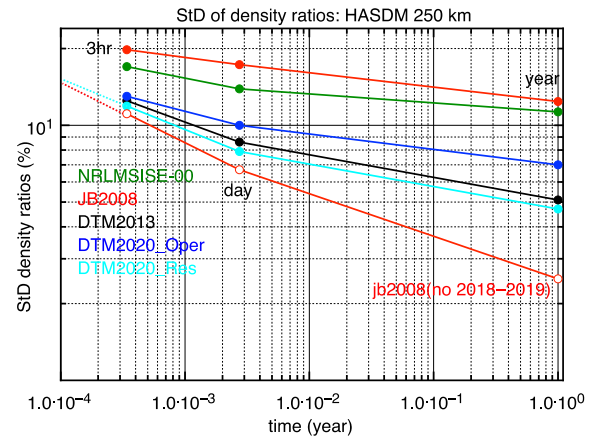


Fig. 7. The StD of the density ratios with HASDM 250 km as a function of time scale. The dotted lines are extrapolations.

800 km, but the DTM models are close, whereas JB2008, which does not predict the winter Helium bulge, has a very large StD.

#### 4.3 Model precision: per day

The StD of the daily mean density ratios is computed per day and listed in Table 4 for GOCE, CHAMP, TLE 250 km (from 2000 to 2013), and HASDM 250 km. DTM2020\_Res has the lowest values, and a significant improvement of around 10% concerning DTM2013 has been made in reproducing daily variability. As expected, values lie between the overall (highest, using all high-cadence data) and yearly mean StD values. The yearly and overall statistics enable us to quantify how StD evolves on 3 h to yearly time scales and shorter to multi-year times scales through extrapolation (but this is a risky operation). An example is shown in Figure 7 using HASDM 250 km data. The JB2008 StD is the highest, varying from 17.4% at 3-h time scales to 12.4% at yearly scales. However, it is lowest when the years 2018 and 2019 are rejected, which have erroneous indices, varying from 2.9% to 11.1%. The figure clearly shows the climatological nature of the models, reflected by the decreasing StD when increasing the time scales.

#### 4.4 Model precision: geomagnetic storms

The assessment under geomagnetic storm conditions is done according to the procedure and on the 13 storms described in (Bruinsma et al., 2021), but with four additional storms at the solar cycle maximum (31/03/2001; 06/11/2001; 23/05/2002; 07/09/2002). In the procedure, storms are divided into four phases relative to the time of minimum Dst: 12 h preceding onset = phase 1; onset = phase 2; recovery = phase 3; post-storm = phase 4). The models are first de-biased concerning the observations via a scaling factor, then applied to the model densities. The scaling factor is the ratio of the sum of all observations to the sum of all model densities in phase 1. The de-biasing procedure aims to minimize the effect of scaling issues between datasets and non-storm related model errors on the assessment

The mean and StD of the density ratios for each phase and the complete interval, as well as the amplitude and timing of the

**Table 5.** The mean and StD of the density ratios, the amplitude error and StD (%; negative is underestimation), and the timing error and StD (h; negative means model peak too early) and correlation coefficient  $R$  for 28 storm cases. The best values are printed in bold.

	MSIS00	JB2008	DTM2013	DTM2020 Oper	DTM2020 Research
$\mu/\sigma$ (all phases)	1.10/30%	1.02/27%	0.98/25%	1.06/25%	<b>1.00/22%</b>
Storms 2001–2003	1.02/28%	0.97/24%	0.96/23%	1.02/24%	<b>1.01/21%</b>
$\mu/\sigma$ (phase 2 + 3)	1.15/33%	1.03/29%	0.98/28%	1.09/28%	<b>1.01/24%</b>
Amplitude error/ $\sigma$ storms 2001–2003	–52%/23%	27%/22%	47%/33%	–32%/19%	– <b>24%/15%</b>
	–40%/26%	35%/ <b>10%</b>	65%/25%	–34%/25%	– <b>28%/19%</b>
Timing error/ $\sigma$	–1.1/1.9	–0.3/3.0	–0.6/2.0	– <b>0.1/2.2</b>	0.7/2.6
$R$	0.84	0.86	0.87	0.87	<b>0.90</b>

maximum density peak, are computed. Orbit-averaged CHAMP, GRACE, and GOCE density data for 17 storms are used in this assessment. The results are given in Table 5, which lists the mean and StD of the density ratios for the complete events (phases 1–4), the storm phases 2 + 3, amplitude, the timing of the peak, and correlation coefficient. DTM2020\_Res reproduces the storms significantly more accurately than DTM2013, with smaller biases and standard deviations, more accurate amplitude, and higher correlation. The density peak amplitude error at high solar activity is halved. The timing error is relatively important, unexpected since the model uses geomagnetic indices with hourly instead of 3-hourly cadence. DTM2020\_Oper underestimates density and has nearly identical standard deviations and correlation as DTM2013, but more accurate amplitudes.

### 4.5 Specific comparisons (binning)

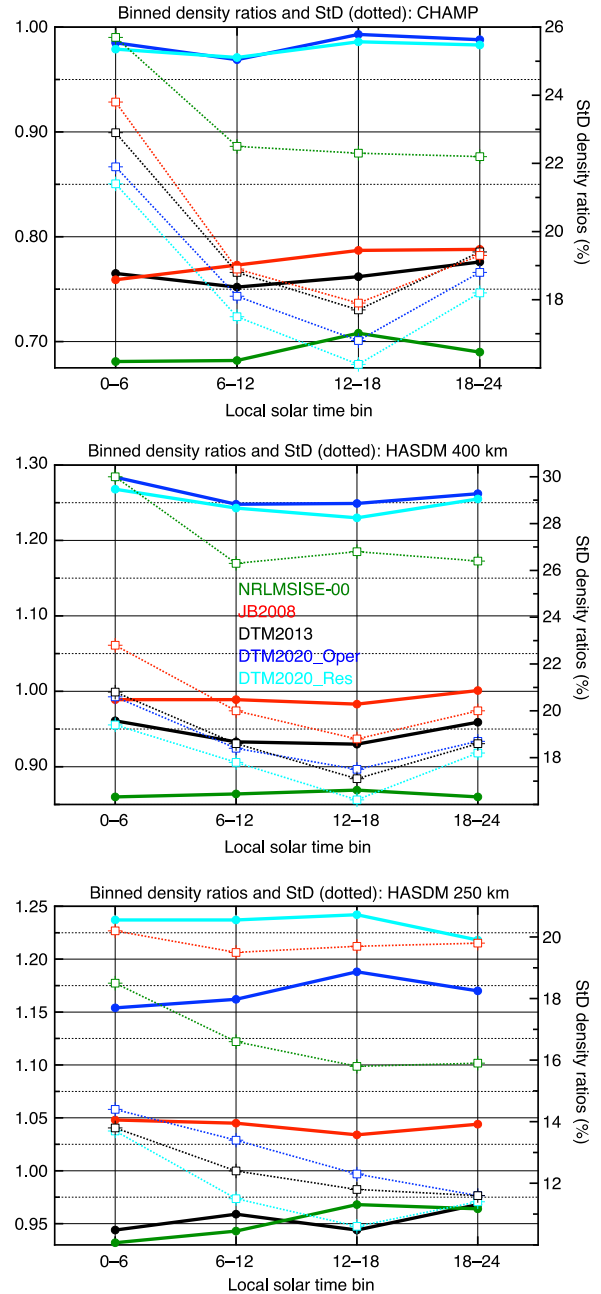
Binning the density ratios in certain model input parameters facilitates testing the model quality concerning specific variations. In the following sections, the most relevant and revealing comparisons are shown. Note that binning provides basic but incomplete information on modeling errors because sorting data in specific bins only partly averages out model errors of different origins.

#### 4.5.1 Solar local time bins

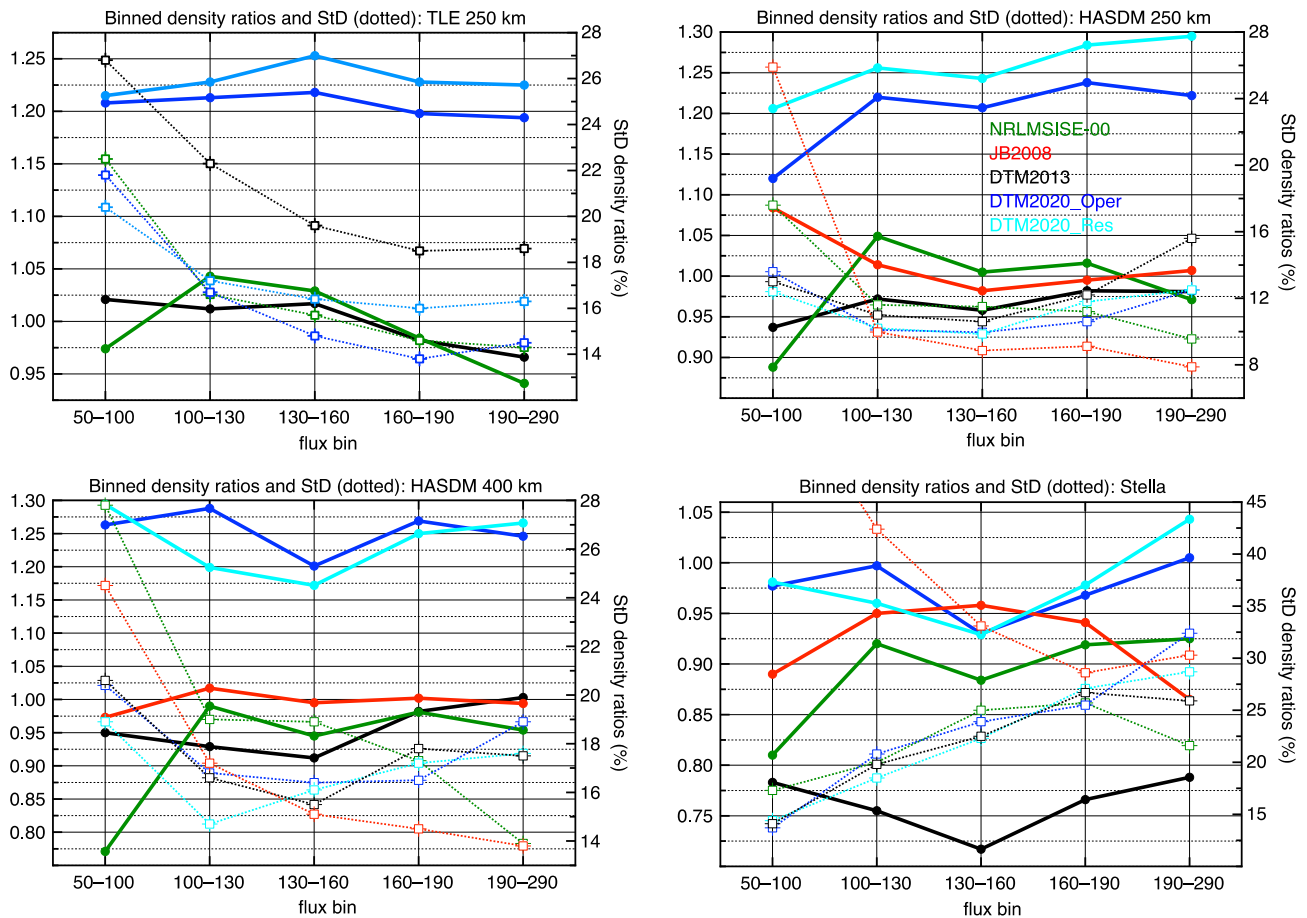
In Figure 8, the density ratios of CHAMP and HASDM at 400 and 250 km are binned per 6-h local solar time. There are no systematic errors visible for any of the models. The StD of all models is always highest in the midnight to early morning sector. DTM2020\_Res has the smallest StD in all sectors. Also, it has a clear minimum StD in the afternoon sector, which is in line with expectations about errors due to density variability at spatial scales smaller than the models' resolution (Bruinsma & Forbes, 2008).

#### 4.5.2 81-day mean solar flux bins

The longest dataset, TLE 250 km, is used to bin the density ratios in 81-day mean solar radio flux. The DTM2020\_Res model cannot be directly tested due to the absence of Hp60 over the entire interval, but the solar activity-related model coefficients are the same as in the intermediate DTM2020 model. Five mean solar flux bins are defined: 50–100, 100–130, 130–160, 160–190, and 190–290 sfu, and the results are shown



**Fig. 8.** Density ratios in 6-h local solar time bins (solid lines) and the StD of the density ratios (dotted lines) for CHAMP (top), HASDM at 400 km (middle), and 250 km (bottom).



**Fig. 9.** Density ratios in 81-day mean solar flux bins (solid lines) and the StD of the density ratios (dotted lines) for TLE 250 km (top right), HASDM at 250 km (top left), and 400 km (bottom left), and Stella (bottom right).

in Figure 9 (top left frame). DTM2013 reveals a small trend of about 5%, while the biases of the DTM2020 models hardly vary with solar activity. NRLMSISE-00 density ratios exhibit a clear dependence on solar activity level, and the trend from medium (100–130 sfu) to high (190–290 sfu) solar activity is about 10% (from 1.04 to 0.94). The StD of the density ratios is highest under low activities for all models. The DTM2020\_Oper model has the lowest StDs except under low activity, for which the best result is obtained with DTM2020\_Intermed. It is unclear presently why F30 does not achieve lower StD at high activity.

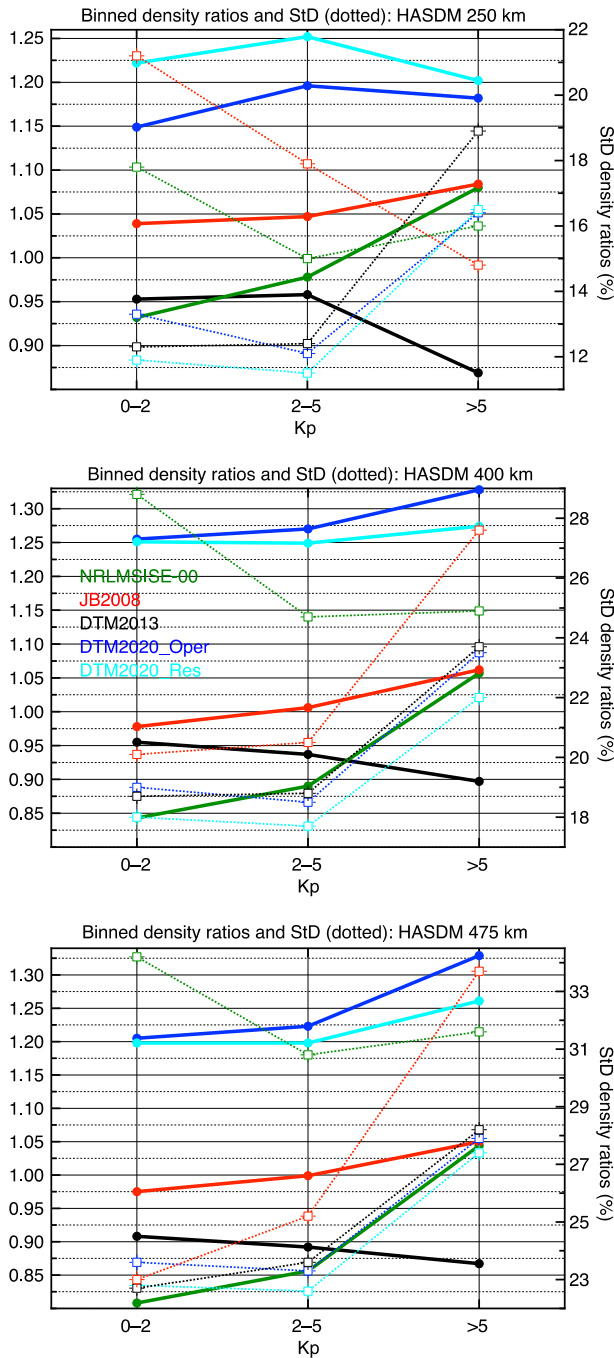
HASDM densities at 250, 400, and 475 km covering almost two solar cycles have also been used in this assessment. The HASDM 250 km and TLE 250 km results do not fully agree, the small trends from low to high activity in particular, but differences are less than  $\pm 5\%$ . The StD, while much larger with TLE 250 km, is largest for low solar activity and reveals a minimum at the medium or medium-high activity for the DTM models, while the StD of NRLMSISE-00 and JB2008 decreases from low to high activity monotonously. Except under low solar activity, JB2008 has the smallest StD. At 400 km, the DTM2020 density ratios do not present a trend, but a v-shaped signature of  $\pm 5\%$  around the mean is visible; JB2008 displays a variation of a few percent. The StD of the models presents

rather comparable behavior to the 250 km results, except that DTM2020\_Res now clearly has the smallest StD for low and low-medium solar activity. The results at 475 km are comparable (not shown).

Comparison with Stella at about 800 km leads to similar v-shaped signatures in the DTM density ratios as with HASDM at 400 and 475 km, namely a minimum for the 130–160 sfu bin and higher values for low and very high activity. The StD, excepting JB2008, is now the smallest for low solar activity. This non-intuitive result may be due to the data weighting scheme (Stella weights optimized for Helium, i.e., low solar activity) and slight incompatibilities with GRACE and Swarm A for DTM, but it cannot explain the same tendency in NRLMSISE-00. The StD of JB2008 does exhibit the highest values for low activity since Helium variability as a function of solar activity is modeled incorrectly (i.e., Helium maximizing in the winter hemisphere).

#### 4.5.3 Geomagnetic activity bins

The HASDM density ratios shown in Figure 10 at 250, 400, and 475 km are binned in  $Kp$  representing low ( $Kp < 2$ ), medium ( $2 < Kp < 5$ ), and high ( $Kp > 5$ ) geomagnetic activity. At 250 km, the DTM2020 models do not have a large difference



**Fig. 10.** Density ratios in low, medium, and high geomagnetic activity bins (solid lines) and the StD of the density ratios (dotted lines) for HASDM at 250, 400, and 475 km (top to bottom).

in bias between low-medium and high activity, and the StD at high geomagnetic activity is 2% lower than for DTM2013. DTM2020\_Res displays the same behavior at 400 and 475 km, and it has the lowest StD, whereas DTM2020\_Oper underestimates density at the highest activity. The comparison with the independent HASDM data confirms the good DTM2020 results of the geomagnetic storm assessment presented in Section 4.4.

## 5 Summary and conclusions

Both DTM2020 models have been constructed using the same accelerometer-inferred densities and daily-mean density data after 2000, but DTM2020\_Oper is also directly fitted to historical data already assimilated in DTM-2000 (mass spectrometer data in particular; Bruinsma et al., 2003), and the global mean densities at 250 km inferred from TLEs spanning four solar cycles. The CHAMP, GOCE, and Swarm A densities are the most accurate total density datasets thanks to detailed satellite shape models and high fidelity numerical models for the aerodynamic coefficients. They constitute the most accurate absolute density scale, 20–30% smaller than DTM2013 and JB2008 (HASDM) densities. Smaller absolute densities are also predicted with NRLMSIS 2.0 but through assimilation of data in the mesosphere instead of the accelerometer-inferred densities.

Two combinations of solar and geomagnetic indices were used to develop DTM2020: the classic combination of F10.7 and Kp is used in DTM2020\_Oper, whereas the drivers for DTM2020\_Res are the 30-cm radio flux F30 and the new planetary geomagnetic index Hp60. The latter model is more precise, but the F30 measurement and production are not considered sufficiently permanent and unailing for use in an operational environment.

DTM2020\_Res is the most precise when comparing to assimilated data. For CHAMP, GOCE, Swarm A, and TLE 250 km, it performs appreciably better than DTM2013, with StD of the density ratios 8%, 22%, 13%, and 9% smaller, respectively (Table 1). For Stella data at higher altitude, DTM2020\_Res (10%) and DTM2020\_Oper (59%) perform better than DTM2013, too, when considering the yearly means (Table 3).

DTM2020 also reproduces the external density datasets more precisely than DTM2013. The EDR density datasets are precisely reproduced with DTM2020\_Res, while DTM2020\_Oper is significantly less precise but still better than DTM2013. Comparison to HASDM data at 250, 400, and 475 km altitudes confirms that DTM2020\_Res is the most precise DTM model. Comparison to HASDM 800 km densities showed that DTM2020\_Oper performs better than DTM2020\_Res at higher altitude, that both are more accurate than DTM2013 (71 and 34 % more accurate, respectively), and that MSIS is the most stable over the 20 years (Table 3). We speculate that this is brought about by more realistic modeling of Helium.

Geomagnetic storms are also reproduced much more accurately with DTM2020\_Res thanks to the new Hp60 indices than with DTM2020\_Oper and DTM2013, which have equal performance (both use Kp). The mean and StD of the density ratios and the amplitude of the storm peak all support this conclusion (Table 5).

Taking all of the above into account, DTM2020\_Res reproduces total mass neutral density most accurately overall, and it is a significant improvement over DTM2013 under all conditions. Applying corrections to F30 during solar cycle minimum conditions (i.e., decreasing F30 by a few percent), which was not done in this project, could improve model accuracy further by reducing or even eliminating biases. Because of the sparseness of precise and high-resolution density data, (semi-empirical) model improvement is still possible and necessary, particularly

from 100 to 250 km and around 800 km (where most Earth observation satellites operate). However, high-resolution density data will remain a scarce commodity in the immediate future, and its provision will remain dependent on missions with other objectives, i.e., data of opportunity. Such high-resolution data is not available at 800 km at all. Secondly, it is essential that all density datasets available presently are processed according to a common standard to make them consistent. That will facilitate the adjustment and improve the accuracy of the resulting model because the model coefficients can be fitted more precisely.

**Acknowledgements.** We thank the reviewers for their time and considerable assistance. The SWAMI project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 776287. SB and CB are equally supported by CNES APR grant METEOESP.

The Hpo indices are distributed by GFZ (<ftp://ftp.gfz-potsdam.de/pub/home/obs/Hpo>)

F30 is distributed by CLS (<ftp://ftpsedr.cls.fr/pub/previsol/solarflux/observation/>)

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/>), or via TU Delft directly (<http://thermosphere.tudelft.nl>)

The DTM models and the density uncertainty estimate model are available on Github via the SWAMI website (<https://swami-h2020.eu>) The editor thanks Eelco Doornbos and Christian Siemes for their assistance in evaluating this paper.

## References

- Boniface C, Bruinsma S. 2021. Uncertainty quantification of the DTM2020 thermosphere model. *J Space Weather Space Clim.*
- Bowman BR, Tobiska WK, Marcos F, Huang CY, Lin CS, Burke WJ. 2008. A new empirical thermospheric density model JB2008 using new solar and geomagnetic indices. In: *AIAA 2008-6438*. AIAA/AAS Astrodynamics Specialist Conference, Honolulu, Hawaii, 2008.
- Bowman BR, Marcos FA, Kendra MJ. 2004. A method for computing accurate daily atmospheric density values from satellite drag data. In: *AAS 04-173*, 14th AAS/AIAA Space Flight Mechanics Conference. Maui, Hawaii. pp. 2004.
- Bruinsma S, Sutton E, Solomon SC, Fuller-Rowell T, Fedrizzi M. 2018. Space weather modeling capabilities assessment: Neutral density and orbit determination at LEO. *Space Weather*. <https://doi.org/10.1029/2018SW002027>.
- Bruinsma SL, Doornbos E, Bowman BR. 2014. Validation of GOCE densities and thermosphere model evaluation. *Adv Space Res* **54**: 576–585. <https://doi.org/10.1016/j.asr.2014.04.008>.
- Bruinsma SL. 2015. The DTM-2013 thermosphere model. *J Space Weather Space Clim*. <https://doi.org/10.1051/swsc/2015001>.
- Bruinsma SL, Forbes JM. 2008. Medium to large-scale density variability as observed by CHAMP. *Space Weather* **6**: S08002. <https://doi.org/10.1029/2008SW000411>.
- Bruinsma S, Tamagnan D, Biancale R. 2004. Atmospheric densities derived from CHAMP/STAR accelerometer observations. *Planet Space Sci* **52**: 297–312. <https://doi.org/10.1016/j.pss.2003.11.004>.
- Bruinsma S, Thuillier G, Barlier F. 2003. The DTM-2000 empirical thermosphere model with new data assimilation and constraints. *J Atmos Sol-Terr Phys* **65**: 1053–1070. [https://doi.org/10.1016/S1364-6826\(03\)00137-8](https://doi.org/10.1016/S1364-6826(03)00137-8).
- Doornbos E, Bruinsma S, Fritsche B, Koppenwallner G, Visser P, Den Van, IJssel J, de Teixeira de Encarnação J. 2014. ESA contract 4000102847/NL/EL, GOCE+ Theme 3: Air density and wind retrieval using GOCE data. *Final Report*, TU Delft.
- Doornbos E. 2011. Thermospheric density and wind determination from satellite dynamics. *Ph.D. Dissertation*. University of Delft, 188 pp. Available at <http://repository.tudelft.nl/>.
- Drinkwater MR, Floberghagen R, Haagmans R, Muzi D, Popescu A. 2003. GOCE: ESA's First Earth Explorer Core Mission. *Space Sci Rev* **108**: 419–432. <https://doi.org/10.1023/A:1026104216284>.
- Dudok de Wit T, Bruinsma S, Shibasaki S. 2014. Synoptic radio observations as proxies for upper atmosphere modelling. *J Space Weather Space Clim*. <https://doi.org/10.1051/swsc/2014003>.
- Emmert JT, Drob DP, Picone JM, Siskind DE, Jones M Jr, et al. 2020. A whole-atmosphere empirical model of temperature and neutral species densities. *Earth Space Sci* **8**(3): e2020EA001321. <https://doi.org/10.1029/2020EA001321>
- Emmert JT. 2015. Altitude and solar activity dependence of 1967–2005 thermospheric density trends derived from orbital drag. *J Geophys Res Space Phys* **120**: 2940–2950. <https://doi.org/10.1002/2015JA021047>.
- Emmert JT. 2009. A long-term data set of globally averaged thermospheric total mass density. *J Geophys Res* **114**: A06315. <https://doi.org/10.1029/2009JA014102>.
- ESA. 1999. The Four Candidate Earth Explorer Core Missions – Gravity field and steady-state ocean circulation explorer. *Report SP-1233(1)*, European Space Agency. [https://earth.esa.int/goce04/Documents/goce\\_sp1233\\_1.pdf](https://earth.esa.int/goce04/Documents/goce_sp1233_1.pdf).
- Friis-Christensen E, Lühr H, Knudsen D, Haagmans R. 2008. Swarm – An Earth Observation Mission investigating Geospace. *Adv Space Res* **41**(1): 210–216. <https://doi.org/10.1016/j.asr.2006.10.008>.
- Jacchia LG, Slowey J. 1963. Accurate drag determinations for eight artificial satellites; Atmospheric densities and temperatures. *Smithsonian Astrophys Obs Spec Rep* **100**: <https://doi.org/10.5479/si.00810231.8-1.1>.
- Jackson DR, Bruinsma S, Negrin S, Stolle C, Budd CJ, et al. 2020. The Space Weather Atmosphere Models and Indices (SWAMI) Project: Overview and First Results. *J Space Weather Space Clim*. <https://doi.org/10.1051/swsc/2020019>.
- Keating GM, Prior EJ. 1968. The winter helium bulge. *Space Res* **8**: 982–992.
- Pearlman MR, Degnan JJ, Bosworth JM. 2002. The international laser ranging service. *Adv Space Res* **30**: 135–143. [https://doi.org/10.1016/S0273-1177\(02\)00277-6](https://doi.org/10.1016/S0273-1177(02)00277-6).
- Picone JM, Hedin AE, Drob DP, Aikin AC. 2002. NRLMSISE-00 empirical model of the atmosphere: Statistical comparisons and scientific issues. *J Geophys Res* **107**(A12): 1468. <https://doi.org/10.1029/2002JA009430>.
- Reigber C, Bock R, Förste C, Grunwaldt L, Jakowski N, Lühr H, Schwintzer P, Tilgner C. 1996. CHAMP Phase B: Executive Summary. *Scientific Technical Report STR; 96/13*. GFZ German Research Centre for Geosciences, Potsdam, 24 p. <https://doi.org/10.2312/gfz.b103-96131>.

- Sentman LH. 1961. Comparison of the exact and approximate methods for predicting free molecule aerodynamic coefficients. *ARS J* **31**: 1576–1579.
- Siemes C, Encarnação J, Doornbos E, Van den IJssels J, Kraus J, Perešty R, Grunwaldt L, Apelbaum G, Flury J, Holmdahl Olsen PE. 2016. Swarm accelerometer data processing from raw accelerations to thermospheric neutral densities. *Earth Planet Space* **68**: 92. <https://doi.org/10.1186/s40623-016-0474-5>.
- Solomon SC, Woods TN, Didkovsky LV, Emmert JT, Qian L. 2010. Anomalous low solar extreme-ultraviolet irradiance and thermospheric density during solar minimum. *Geophys Res Lett* **37**: <https://doi.org/10.1029/2010GL044468>.
- Storz MF, Bowman BR, Branson MJI, Casali SJ, Tobiska WK. 2005. High accuracy satellite drag model (HASDM). *Adv Space Res* **36**: 2497–2505. <https://doi.org/10.1016/j.asr.2004.02.020>.
- Tapley BD, Bettadpur S, Watkins M, Reigber C. 2004. The gravity recovery and climate experiment: Mission overview and early results. *Geophys Res Lett* **31**: L09607. <https://doi.org/10.1029/2004GL019920>.
- Tobiska WK, Bowman BR, Bouwer SD, Cruz A, Wahl K, Pilinski MD, Mehta PM, Licata RJ. 2021. The SET HASDM density database. *Space Weather* **19**: e2020SW002682. <https://doi.org/10.1029/2020SW002682>.
- Van den IJssel J, Doornbos E, Iorfida E, March G, Siemes C, Montenbruck O. 2020. Thermosphere densities derived from Swarm GPS observations. *Adv Space Res* **65**: 7. <https://doi.org/10.1016/j.asr.2020.01.004>.
- Vourlidas A, Bruinsma S. 2018. EUV irradiance inputs to thermospheric density models: open issues and path forward. *Space Weather* **16**: 5–15. <https://doi.org/10.1002/2017SW001725>.

**Cite this article as:** Bruinsma S & Boniface C 2021. The operational and research DTM-2020 thermosphere models. *J. Space Weather Space Clim.* **11**, 47. <https://doi.org/10.1051/swsc/2021032>.