

Space Weather

RESEARCH ARTICLE

10.1029/2018SW002027

Special Section:

Space Weather Capabilities
Assessment

Key Points:

- We present density data, periods and events, and metrics for comprehensive and standardized thermosphere model assessment
- A short description is given for the models evaluated in this study
- Examples of the assessment are given for TIE-GCM, CTIPE, NRLMSISE-00, JB2008 and DTM2013

Correspondence to:

S. Bruinsma,
sean.bruinsma@cnes.fr

Citation:

Bruinsma, S., Sutton, E., Solomon, S. C., Fuller-Rowell, T., & Fedrizzi, M. (2018). Space weather modeling capabilities assessment: Neutral density for orbit determination at low Earth orbit. *Space Weather*, 16, 1806–1816. <https://doi.org/10.1029/2018SW002027>

Received 23 JUL 2018

Accepted 25 OCT 2018

Accepted article online 29 OCT 2018

Published online 14 NOV 2018

Space Weather Modeling Capabilities Assessment: Neutral Density for Orbit Determination at low Earth orbit

S. Bruinsma¹ , E. Sutton² , S. C. Solomon³ , T. Fuller-Rowell⁴, and M. Fedrizzi⁴

¹CNES, Space Geodesy Office, Toulouse, France, ²Air Force Research Laboratory, Albuquerque, NM, USA, ³National Center for Atmospheric Research, Boulder, CO, USA, ⁴NOAA/CIRES, Boulder, CO, USA

Abstract The specification and prediction of density changes in the thermosphere is a key challenge for space weather observations and modeling, because it is one result of complex interactions between the Sun and the terrestrial atmosphere and also because it is of operational importance for tracking objects orbiting in near-Earth space. For low Earth orbit, neutral density variation is the most important uncertainty for propagation and prediction of orbital elements. A recent international conference conducted under the auspices of the National Aeronautics and Space Administration Community Coordinated Modeling Center included a workshop on neutral density modeling, using both empirical and numerical methods, and resulted in organization of an initial effort in model comparison and evaluation. Here we report on the exploitable density data sets available, the selected years and storm events, and the metrics for complete model assessment. Comparisons between five models (three empirical and two numerical) and neutral density data sets that include measurements by the CHAMP, GRACE, and GOCE satellites are presented as examples of the assessment procedure that will be implemented at Community Coordinated Modeling Center. The models in general performed reasonably well, although seasonal errors sometimes are present, and impulsive geomagnetic storm events remain challenging. Numerical models are still catching up to empirical methods on a statistical basis, but hold great potential for describing these short-term variations.

1. Introduction

The goal of the Neutral Density and Orbit Determination at low Earth orbit working team was to establish metrics for the assessment of thermosphere models, select density data sets with sufficient precision and resolution, as well as periods and specific events for the comparisons. Presently, the fidelity of both semiempirical (SE) and first-principles (FP) models is at best quantified for specific (storm) events, data of a specific satellite mission, or in the form of general statistical numbers. Because often different metrics were used, results are not directly comparable. The ultimate goal of this exercise is to evaluate all thermosphere models available on the CCMC (Community Coordinated Modeling Center: <https://ccmc.gsfc.nasa.gov>) by comparing to the same data and for the same events, applying a consistent and always identical method, in order to establish score cards that can help users make a deliberated choice for their objective.

Operational use of thermosphere models is mainly in determination and prediction of orbits of satellites and orbital debris. The precision of the determination and prediction of orbit positions of objects in low Earth orbit depends on the quality of the force model for atmospheric drag. The atmospheric drag model must be corrected for bias, which is usually done by scaling the thermosphere density, because it causes orbit errors that increase with time. Atmospheric drag depends heavily on the highly variable, both spatially and temporally, total neutral density and also the composition of the thermosphere. The variability is driven by direct solar radiative heating in the ultraviolet (UV) to extreme UV (EUV) range of the spectrum, Joule heating, and particle precipitation due to interaction of the magnetosphere with the solar wind (referred to as *geomagnetic activity* as opposed to *solar activity*), and to some extent due to upward propagating perturbations that originate in Earth's lower atmosphere. Consequently, thermosphere model performance depends strongly on the choice of solar and geomagnetic activity drivers, and in case of density prediction, on the accuracy of the forecast of the drivers. The error due to activity forecast is out of the scope of this paper but has recently been addressed by Bussy-Virat et al. (2018) and Hejduk and Snow (2018).

Three SE models, NRLMSISE-00 (Picone et al., 2002), JB2008 (Bowman et al., 2008), and DTM2013 (Bruinsma, 2015), and two FP models, Thermosphere-Ionosphere-Electrodynamics General Circulation Model (TIE-GCM;

Table 1
Thermosphere and Upper Atmosphere Models Available for Testing

Model	Type	Drivers (solar geomagnetic)	Hor. and time resolution
NRLMSISE-00	SE	$F_{10.7} ap$ (array of seven values)	30°, 3 hr
JB2008	SE	$S_{10}, F_{10.7}, M_{10}, Y_{10} Dst, ap$	30°, 1 hr
DTM2013	SE	$F_{30} K_p$	30 × 15°, 3 hr
TIE-GCM	FP	$F_{10.7} K_p$	2.5 × 2.5°, 1 min
CTIpe	FP	$F_{10.7} K_p$, solar wind, interplanetary magnetic field, hemispheric power	2.0 × 18°, 1 min

Note. SE = semiempirical; FP = first-principles.

Richmond et al., 1992; Roble et al., 1988) and coupled thermosphere-ionosphere-plasmasphere electrodynamics (CTIpe; Fuller-Rowell et al., 1996), are used in this first and incomplete assessment. All five models are implemented at CCMC. The model resolution and drivers are listed in Table 1. Naturally, this is not an exclusive list of models and we encourage modelers that have not submitted their model to do so in order to have a score card as complete as possible. The JB2008 model is somewhat difficult to use for model evaluation because of its frequently-changing solar drivers S_{10} and S_{81} , which are computed and modified by the modelers. As a consequence, repeatability of the assessments cannot be achieved; the solar activity file (SOLFSMY) that was online in the month March 2018 was used for the assessment given in this paper.

The following section provides short descriptions of the five models tested in this first part of the assessment. The third section presents the procedure to assess the models, which is applied to the five models in an example given in section 4. The final section describes the developments and actions required to accomplish the complete assessments of the models listed in Table 1, as well as all models that can be run on the future CCMC testbed.

2. Model Descriptions

2.1. Semiempirical Thermosphere Models: NRLMSISE-00, JB2008, and DTM2013

The SE models are mainly used in orbit computation and mission design. They are easy to use and computationally fast thanks to their simplified modeling algorithm, providing pointwise estimates (i.e., what is required along an orbit). They are climatology (or *specification*) models that have a low spatial and temporal resolution of the order of thousands of kilometers, which is due to the low maximum degree of the spherical harmonic expansion used in the algorithm, and hours, imposed by the cadence of the geomagnetic indices, respectively. As a consequence, the SE models cannot reproduce wave-like activity due to, for example, geomagnetic storms and the large scale traveling atmospheric disturbances it causes, the complex dynamics in the polar caps, or the effects of tidal perturbations propagating from the lower atmosphere. The minimum altitude of JB2008 and DTM2013 is 120 km, whereas it is 0 km for NRLMSISE-00, and they can be used to approximately 1,500 km.

SE models are constructed by optimally estimating the unknown model coefficients to their respective underlying databases (density, temperature, and composition measurements) in the least squares sense. The main sources of density data are satellite drag inferred total densities by means of orbit perturbation analysis (Jacchia & Slowey, 1963) or accelerometers (e.g., Champion & Marcos, 1973) and neutral mass spectrometers (e.g., Nier et al., 1973), which provide composition measurements. Both techniques have in common that they do not provide absolute measurement of density, due to calibration issues in case of mass spectrometers. The most recent and precise accelerometer-inferred density data sets are also not absolute, but their scale depends on the satellite model (and in particular the aerodynamic coefficient) that was used in the computation. As a consequence, the SE thermosphere models are in a certain fashion scaled to the adopted satellite models—which are rarely the same between modelers. JB2008 (up to 2008 at least) and DTM2013 have a scale that is close to the U.S. Air Force operational thermosphere model HASDM (Storz et al., 2005).

2.2. TIE-GCM

The National Center for Atmospheric Research TIE-GCM is a FP upper atmospheric general circulation model that solves the Eulerian continuity, momentum, and energy equations for the coupled thermosphere-ionosphere system (Richmond et al., 1992; Roble et al., 1988). It uses pressure surfaces as the vertical coordinate and extends in altitude from approximately 97 to 600 km. The normal resolutions of the model are 5° horizontal and 0.5 scale height in the vertical or 2.5° horizontal and H/4 vertical. Tidal forcing at the lower boundary is specified by the Global Scale Wave Model (Hagan et al., 2001), and semiannual and annual density periodicities are enhanced by applying seasonal variation of the eddy diffusivity coefficient at the lower boundary. Solar inputs are described by Solomon and Qian (2005). The low-latitude electrodynamic potential field is internally generated using the model conductivities, densities, and neutral winds; it is merged with a high-latitude magnetospheric potential from the Heelis et al. (1982) empirical formulation, which is driven by the *K_p* index, following the method described in Solomon et al. (2012). There is also an option to use the Weimer (2005) empirical model of magnetospheric potential, which uses upstream solar wind and interplanetary magnetic field as input, but that option was not employed in the runs conducted for this work. Recent developments, addition of helium as a major species, and lower boundary options are described in Qian et al. (2014), Sutton et al. (2015), and Maute (2017). Version 2.0 of the TIE-GCM is a community release that was issued in March 2016.

2.3. CTIPe

The CTIPe is a global, three-dimensional, time-dependent, nonlinear, self-consistent model that solves the momentum, energy, and composition equations for the neutral and ionized atmosphere (Fuller-Rowell et al., 1996; Millward et al., 2001). The global atmosphere in CTIPe is divided into a series of elements in geographic latitude, longitude, and pressure. The latitude resolution is 2°, the longitude resolution is 18°, and each longitude slice sweeps through local time with a 1-min time step. In the vertical direction, the atmosphere is divided into 15 levels in logarithm of pressure from a lower boundary of 1 Pa at 80- to more than 500-km altitude. The magnetospheric input is based on the statistical models of auroral precipitation and electric fields described by Fuller-Rowell and Evans (1987) and Weimer (2005), respectively. Auroral precipitation is keyed to the hemispheric power index, based on the TIROS/National Oceanic and Atmospheric Administration auroral particle measurements. The Weimer electric field model is keyed to the solar wind parameters impinging the Earth's magnetosphere, and its input drivers include the magnitude of the interplanetary magnetic field in the *y-z* plane, together with the velocity and density of the solar wind. The (2,2), (2,3), (2,4), (2,5), and (1,1) propagating tidal modes are imposed at 80-km altitude (Fuller-Rowell et al., 1991; Müller-Wodarg et al., 2001) with a prescribed amplitude and phase. In this paper, the lower boundary conditions in CTIPe simulations are specified using monthly averaged wind and temperature fields from the Whole Atmosphere Model (Akmaev et al., 2008; Fuller-Rowell et al., 2008). CTIPe uses time-dependent estimates of nitric oxide (NO) obtained from Marsh et al. (2004) empirical model based on Student Nitric Oxide Explorer satellite data rather than solving for minor species photochemistry self-consistently. For higher-altitude applications, helium needs to be included in the model. Solar heating, ionization and dissociation rates, and their variation with solar activity are specified by Solomon and Qian (2005) solar EUV energy deposition scheme for upper atmospheric general circulation models.

3. Model Assessment Procedure

The following three subsections describe the density data and the necessary preprocessing, the time intervals and storms selected for comprehensive model evaluation, and the metrics. All models, and also model upgrades, will be tested according to the same standards allowing unambiguous quantification of improvement.

3.1. Selected Density Data

The density variability in the thermosphere is large (tens to hundreds of percent), and it can be rapid (few hours) in the event of geomagnetic storms. It depends on latitude, local solar time, longitude, season, and solar and geomagnetic activity, and this dependence changes with altitude (i.e., composition). However, most density measurements are in situ and taken along satellite orbits, and therefore, the coverage provided by a single satellite is in fact rather poor. The latitudinal extent depends on the orbital inclination, the local

Table 2
Data Sets Selected for the Model Assessment

Satellite	Period	Altitude (km)	i	LST24 hr LST	Cadence	Precision (%)
CHAMP	May 2001 to August 2010	450–250	87°	0–24120–130	80 s	1–4
GRACE	August 2002 to July 2016	490–300	89°	0–24120–160	80 s	2–6
GOCE	November 2009 to October 2013	270–180	96°	6–8 and 18–20	80 s	1–3
Swarm A	June 2014 to May 2017	450–440	89°	0–24135	10 s	5 ^a
TLE (Emmert)	January 1967 to December 2013	250, 400, 550	—	—	24 hr	2
Stella	January 2000 to December 2016	815	93°	9–15 and 21–3	24 hr	5–15
Starlette	January 2000 to December 2016	800	49°	0–24200	24 hr	5–20

Note. (i is inclination, and LST is the local solar time coverage and the approximate period, in days, to cover 24 hr). The orbital period of the satellites varies from 90 (GOCE) to 100 min (Stella). TLE = two-line element.

^aAfter averaging over one orbit during low solar activity.

time coverage is essentially limited to the time of its ascending and descending pass and the precession of the orbital plane, and the temporal resolution is one orbital period of roughly 1.5 hr. Precise density data sets of recent satellite missions and mean data sets that comprise more than one solar cycle are selected to cover as much of the relevant temporal and spatial scales as possible. However, with the currently available data one cannot reconstitute the complete picture of the state of the thermosphere at any given time.

Table 2 lists the essential information of the selected data sets. All density data sets except the two-line element (TLE) data were processed by one of the authors (Bruinsma) and are available on National Aeronautics and Space Administration/CCMC. The original GOCE and Swarm densities can be obtained, after registration, on the ESA server (earth.esa.int). The GOCE densities used in this study were first rescaled to the HASDM model (using a scale factor of 1.24 for version v1.5; Bruinsma et al., 2014), then filtered to suppress scales smaller than 600 km, and finally downsampled to 80-s cadence (the CHAMP and GRACE densities were similarly filtered and downsampled; they are also consistent with HASDM model scale). The global mean, daily mean densities derived from TLE data are available in the supporting information of Emmert (2015).

3.2. Selected Long Time Intervals and Storm Events

It was proposed and accepted to run the models over a number of complete years in order to assess model performance on seasonal and solar-rotation period time scales, under high (2002), low (2007), and medium (2012) solar activity conditions. Figure 1 presents the 81-day mean of the solar radio flux $F_{10.7}$, the most used proxy for solar UV/EUV activity, and the insert shows in addition the daily $F_{10.7}$ for 2012 with strong 27-day variations; the years selected for assessment are within the blue frames.

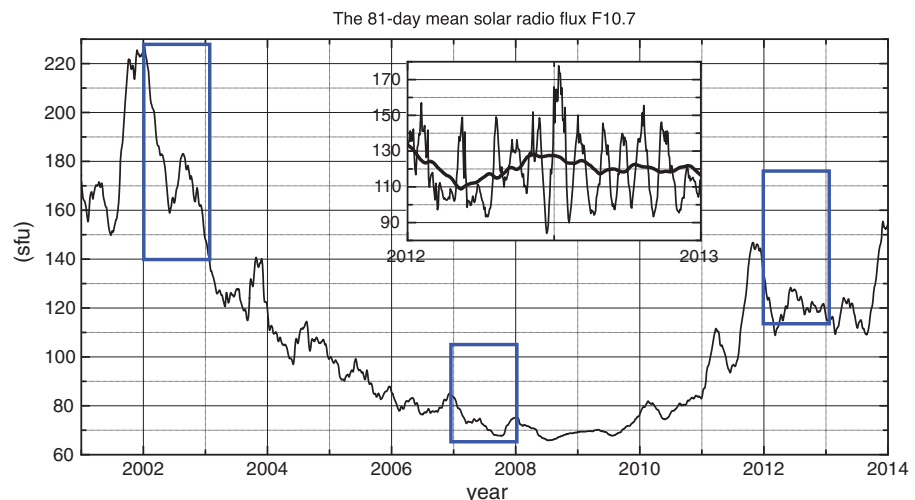


Figure 1. The 81-day mean solar radio flux $F_{10.7}$. The inset shows the daily $F_{10.7}$ for 2012.

Table 3
Selected Storm Intervals, Minimum *Dst*, and Maximum *ap/Kp*

Date	Min <i>Dst</i> (nT)	Max <i>ap/Kp</i>
29 March to 3 April 2001	−387	300/9−
18–31 July 2004	−170	300/9−
17–20 January 2005	−103	179/8−
20–23 January 2005	−105	207/8
7–10 October 2005	−127	236/8+
14–17 May 2005	−263	236/8+
29 May to 1 June 2005	−138	179/8−
8–14 July 2005	−92	94/6+
23–26 August 2005	−216	300/9−
8–19 September 2005	−147	179/8−
8–11 March 2012	−131	207/8
16–20 March 2013	−132	111/7−
31 May to 4 June 2013	−119	132/7
21–24 June 2015	−204	236/8+

Fourteen storm events were selected: Seven of these were also selected by the Ionosphere Team, whereas eight storms in 2005 are the so-called problem storms for the U.S. Air Force (Knipp et al., 2013). Table 3 lists the dates, minimum *Dst*, and maximum *ap/Kp* of the storms. The intervals are selected to cover, if possible, a storm sequence that returns to low geomagnetic activity (*Kp*); for some storms, it results in rather long intervals.

3.3. Metrics for Model-Data Comparison

The models and data can be compared by computing density residuals, which is an absolute difference (observed minus computed), but the observed-to-computed density ratio is a better quantity to express a model's skill to reproduce the observations, that is, reality. Density ratios of one indicate perfect duplication of the observations, that is, an unbiased model that reproduces all features, deviation from unity points to under (larger than 1) or overestimation (less than 1). Because of the very large and dynamic range of density, mainly as a function of altitude and solar cycle, it is rather difficult to analyze and interpret model performance in absolute

values, whereas the relative precision of a density ratio is always simple to comprehend. A model bias, that is, the mean of the density ratios differs from unity, is most damaging to orbit extrapolation because it causes position errors that increase with time. The standard deviation (StD) of the density ratios or residuals represents a combination of the ability of the model to reproduce observed density variations, and the geophysical noise (e.g., waves, the short-duration effect of large flares) and instrumental noise in the observations. The mean, StD, and sometimes the root-mean-square of the density ratios, due to their distribution, are computed in log-space (Sutton, 2018):

$$\text{Mean } O/C = \text{Exp}(\langle \ln(O/C) \rangle); \text{StD } O/C (\%) = (\text{Exp}(\text{StD}(\ln(O/C))) - 1) * 100$$

where $\langle \rangle$ indicates computing the mean. The correlation coefficients R are also computed. The correlation coefficient is independent of model bias, and R^2 represents the fraction of observed variance captured by the model.

Mean, StD, and correlation are computed on specific time scales, which are selected based on solar variability and operational forecast horizons (year, 27-day rotation, daily mean, and orbit average). Figure 2, which shows GOCE observations and model predictions (DTM2013), illustrates the assessment results for a time scale of one orbit. The model performance is computed using the observations listed in Table 2, that is, filtered and then downsampled to 80 s. The ratio can be computed in two ways: as the average of all individual ratios (0.986) in the time interval or as the ratio of the sum of all observations to the sum of all model

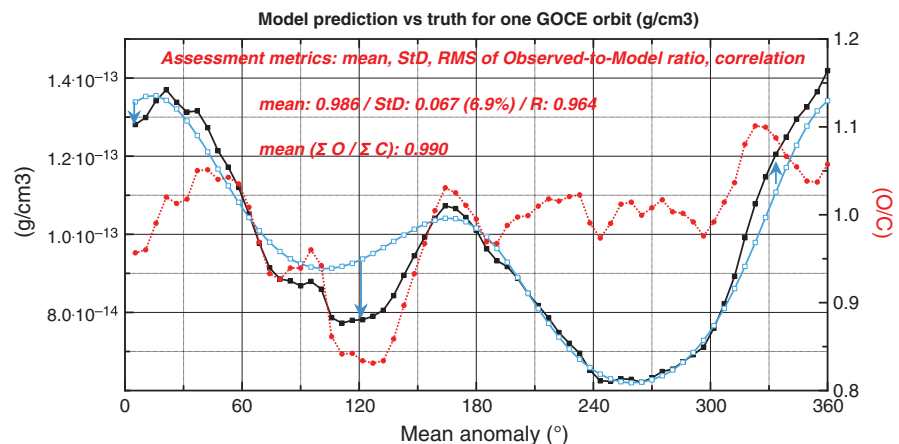


Figure 2. Observed (GOCE: black) and model (blue) density, and the resulting density ratios (red; right axis). The formal uncertainty of the GOCE densities here is 1–2%. StD = standard deviation; RMS = root-mean-square.

Table 4

Average (Total) Density Ratio and StD (%) of the Density Ratios (Log Space), and Correlation Coefficient, With the Models DTM2013, NRLMSISE-00, JB2008, TIE-GCM, and CTIpe for GOCE Only

Model	CHAMP	GRACE	GOCE
DTM2013 (2002)	1.01(1.01) 17.6 0.94	1.05(1.06) 20.2 0.91	
NRLMSISE-00	1.04(1.06) 18.6 0.95	1.09(1.11) 23.7 0.89	
JB2008 ^a	1.09(1.08) 18.2 0.94	1.24(1.22) 27.0 0.86	
TIE-GCM	1.05(1.09) 22.7 0.94	1.14(1.18) 25.7 0.90	
DTM2013 (2007)	1.00(1.00) 20.4 0.92	1.06(1.05) 27.3 0.91	
NRLMSISE-00	0.81(0.82) 22.9 0.90	0.80(0.80) 29.6 0.89	
JB2008 ^a	0.96(0.98) 22.0 0.91	1.06(1.02) 33.7 0.89	
TIE-GCM	0.91(0.96) 30.9 0.85	1.02(1.11) 41.1 0.82	
DTM2013 (2012)		1.10(1.13) 20.9 0.94	0.98(0.99) 11.2 0.94
NRLMSISE-00		1.21(1.27) 25.3 0.93	1.04(1.05) 12.2 0.94
JB2008 ^a		1.31(1.33) 24.3 0.93	1.02(1.02) 11.6 0.95
TIE-GCM		1.30(1.38) 30.4 0.88	0.99(1.00) 13.3 0.92
CTIpe			1.20(1.19) 12.1 0.95

Note. StD = standard deviation; TIE-GCM = Thermosphere-Ionosphere-Electrodynamics General Circulation Model; CTIpe = coupled thermosphere-ionosphere-plasmasphere electrodynamics.

^aWith solar activity file SOLFSMY downloaded in March 2018.

predictions (0.990). The average of the ratios informs us if the model is under or overestimating most of the time (i.e., for most observations) regardless of the absolute error, whereas the ratio of sums represents the total model error over the interval, that is, the error as experienced in satellite orbit computation over the same interval. Ideally, these numbers are the same. Both values, average density ratio, and total density ratio, respectively, will be computed.

4. Examples of Assessment Results

This is only an example of model assessment and the adopted metrics, and not all data listed in Table 2 are used nor will the models be evaluated for all storms listed in Table 3.

Table 4 lists the average (total) density ratio, StD, and R using the high-resolution accelerometer-inferred densities for the 3 years selected in section 3.2. Performance of the models is highly dependent on the year (solar activity) and the altitude. DTM2013 has smallest bias and StD for all years and missions, which is an expected

CHAMP 27-day mean density ratios: DTM2013 (blue), NRLMSISE-00 (green), JB2008 (red)

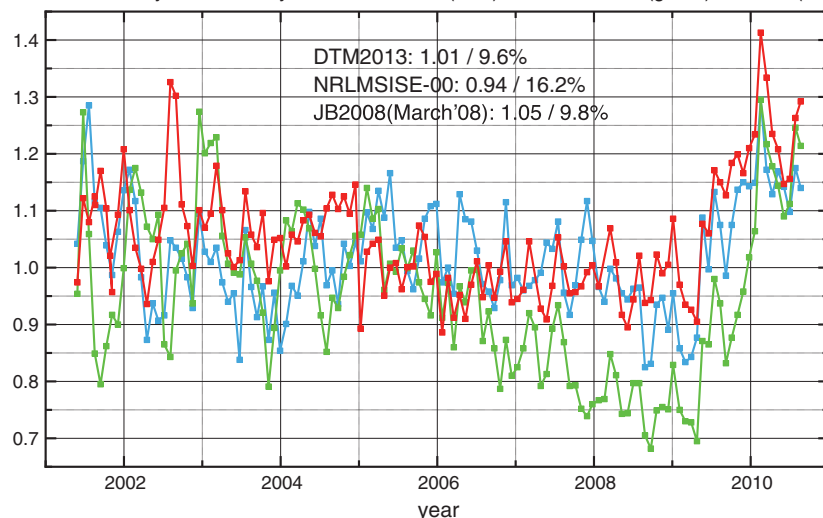


Figure 3. Density ratios computed per 27 days with CHAMP data. The mean and standard deviation of the density ratios are given for the models.

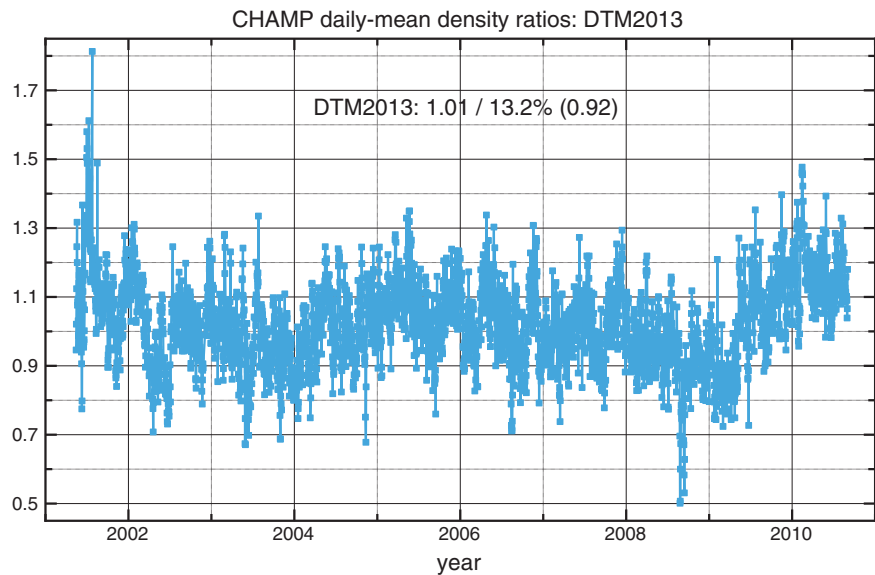


Figure 4. DTM2013 density ratios computed per day with CHAMP data. The mean and standard deviation of the density ratios (correlation) is given for DTM2013. The corresponding numbers for NRLMSISE-00 are 0.94/19.8% (0.92) and for JB2008 1.05/12.9% (0.91).

result because these data were assimilated with internally consistent scale. All models perform best at low altitude, with StDs ranging from 11% to 13%. CTIPe estimates density approximately 20% lower than the other models, which in fact agrees well with the original GOCE data, that is, before scaling to HASDM. NRLMSISE-00 overestimates density in 2007, which is probably due to the solar proxy $F_{10.7}$. TIE-GCM has the highest StD and lowest correlation coefficients for 2007.

Figure 3 displays the NRLMSISE-00, JB2008, and DTM2013 density ratios for CHAMP over the entire mission, computed per 27 days. The models have rather different performance, as expected, due to their underlying databases. DTM2013 is accurate with a StD of less than 10% because of the assimilation of the CHAMP data. This particular realization of JB2008 predictions (i.e., using proxies from March 2018) is accurate as well, but it

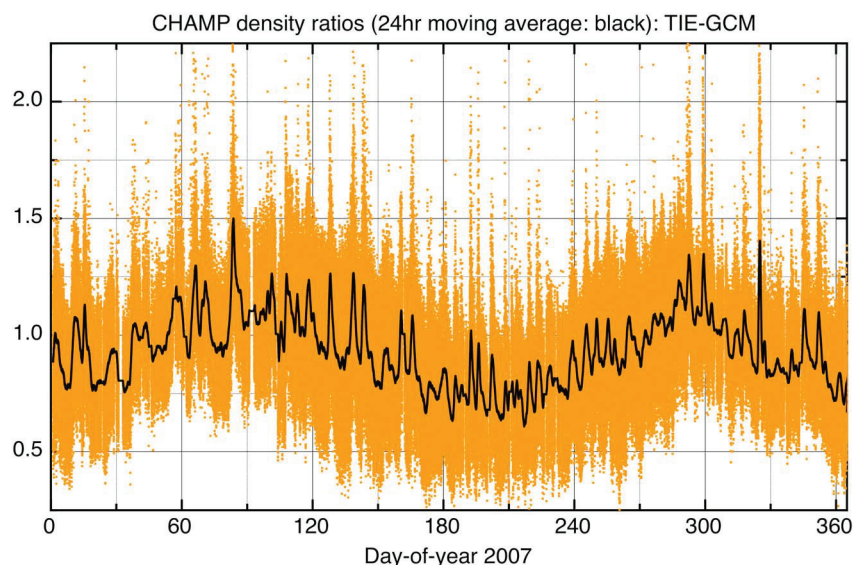


Figure 5. TIE-GCM density ratios computed with 2007 CHAMP data (black: 24-hr moving average). TIE-GCM = Thermosphere-Ionosphere-Electrodynamics General Circulation Model.

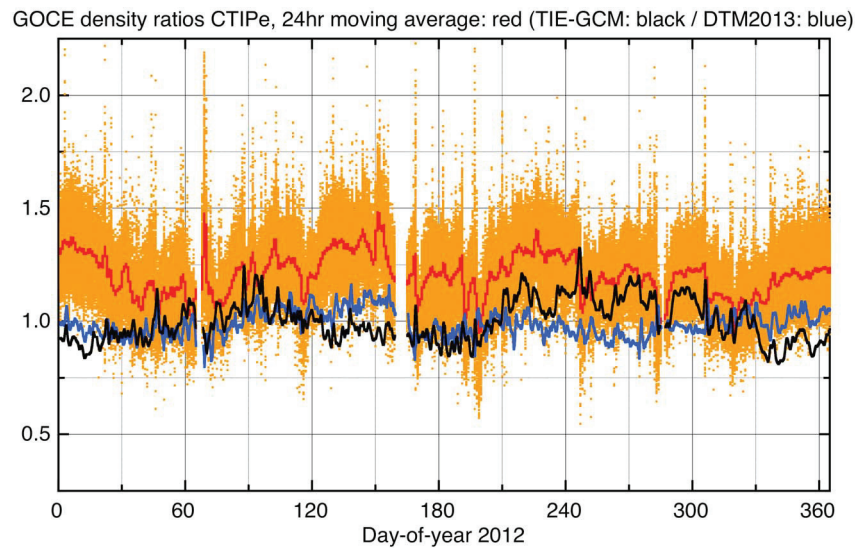


Figure 6. CTIPe density ratios computed with 2012 GOCE data (red: 24 hr moving average). The black and blue line represent the 24-hr moving average of TIE-GCM and DTM2013, respectively. TIE-GCM = Thermosphere-Ionosphere-Electrodynamics General Circulation Model; CTIPe = coupled thermosphere-ionosphere-plasmasphere electrodynamics.

is due to modification of the solar drivers from 2008 to 2010 for a large part. NRLMSISE-00 and DTM2013 proxies are not modified to improve the model fit to density data for those low solar activity years, which causes their overestimation of density for those solar minimum years. The correlation coefficients are about the same for the three models.

Figure 4 displays the DTM2013 density ratios for CHAMP, but now computed per day, and lists the statistics for the three SE models. The bias of course remains identical for all models, while the StD increases by 60–70%. The considerable overestimation during the solar cycle minimum is clearly visible. Besides that, considerable underestimation in 2001 and an approximately 10% systematic underestimation from mid-2009 to the end of the mission is revealed in Figures 3 and 4. Clear signatures of model errors at seasonal time scales are not visible.

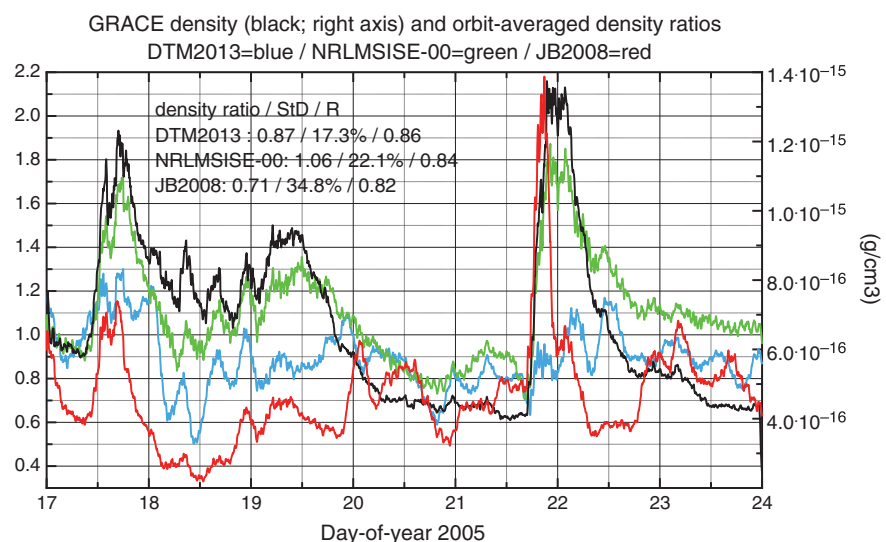


Figure 7. Orbit-averaged density ratios computed with GRACE density data (black; right axis) for the 17 and 21 January 2005 storms. The mean density ratio, the standard deviation, and correlation coefficient, respectively, are printed for the three models.

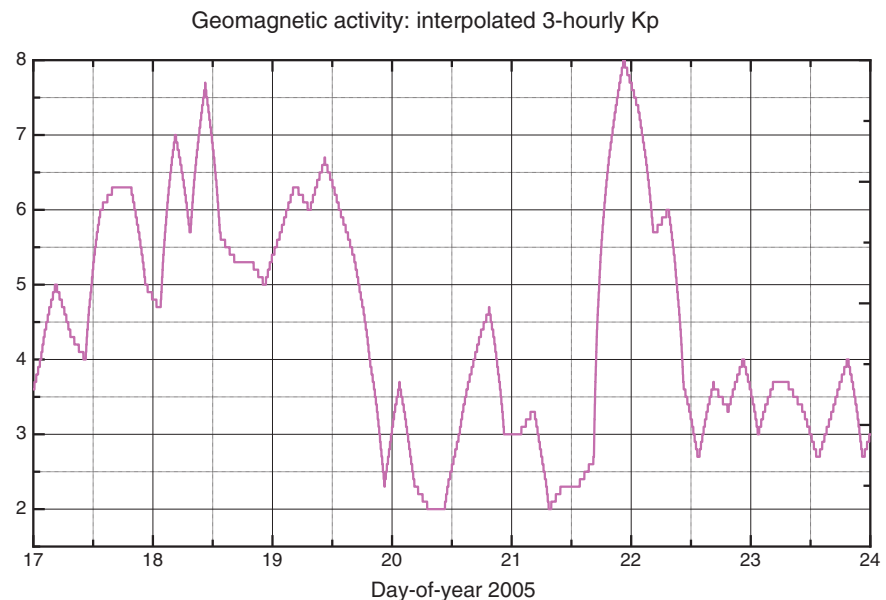


Figure 8. The 3-hourly K_p index, as used in DTM2013 after linear interpolation, for the two January 2005 storms.

Figure 5 displays the density ratios computed with TIE-GCM for CHAMP in 2007. One reason for the large StD of the FP models relative to that of the SE models is due to the incompletely modeled annual and semiannual variations. The configuration of TIE-GCM used here attempts to correct this by implementing a variation in eddy diffusivity at the lower boundary of the model and does reasonably well reproducing the annual variation during 2007. However, there is still noticeable disagreement between TIE-GCM and CHAMP during this year when it comes to the semiannual variation. In fact, such errors are visible in the comparisons with all three accelerometer data sets (not shown). This kind of information, the type of error, is valuable to both modelers as well as users and will be mentioned on the scorecards.

Figure 6 displays the density ratios computed with CTIPe, DTM2013, and TIE-GCM for GOCE in 2012. CTIPe has a 20% bias with respect to the GOCE data used in this assessment, but the StD and correlation are good. Compared to the original ESA GOCE data, that is, not rescaled to HASDM as was done here, CTIPe is unbiased. Due to the absence of a clear density data processing standard, and notably the aerodynamic coefficient modeling, bias is not an objective variable in model evaluation. TIE-GCM still has a visible error at the semiannual period, but much smaller than at CHAMP altitude (Figure 5; same Y axis). DTM2013, like CTIPe, has no discernible seasonal error.

An example of (SE) model performance during a storm event is given by comparison with GRACE data for two of the selected events, the two, in quick succession, January 2005 storms. Figure 7 presents the density ratios and the GRACE observations, all smoothed over one orbit. The performance of the three models is inadequate for the first storm, although in different ways. DTM2013 reproduces the second storm much better on average than the other two models, for which the onset of the storm starts too late. The geomagnetic activity, the 3-hourly K_p index, is displayed in Figure 8. The raw (5-s) density ratios and the corresponding StD are much larger, as can be seen in Table 4. The models under and overestimate density by up to a factor 5 locally, and factors of 3 and 2, respectively, for orbit-averaged densities.

5. Summary and Conclusions

The density data, periods and special events, and the metrics for thermosphere model assessment have been described in this paper. Examples of the assessment on several time scales have been given but only for models that were easily accessible to us. The density data do not provide an absolute scale for the densities due to differences and errors in the satellite model used in their calculation. As a consequence, model bias is not an objective variable, and it must be interpreted with caution; an example of this is given in Figure 6, which presents comparison with GOCE data. In absence of a standard approach to model satellites (i.e.,

their shape and aerodynamic coefficients), thermosphere model performance is objectively represented by StD and correlation. Unfortunately, this means that the orbit computation community has to estimate a scaling factor per satellite, which cannot be exchanged between groups using different orbit determination software. For atmospheric studies or for thermosphere modeling, it means that total density and partial densities of the constituents are not absolute values.

The next step is to perform the assessment of all models available at CCMC according to the procedure presented in this paper. Their score cards will be published on the CCMC website.

Acknowledgments

We thank ESA (GOCE, Swarm), NASA (GRACE), and DLR (CHAMP and GRACE) for making their data publicly available. The density measurements, excepting TLE densities, are available on NASA/CCMC (<https://webserver1.ccmc.gsfc.nasa.gov/camel/>). S. B. received support from CNES/SHM. We acknowledge the reviewers for their valuable comments that helped to improve the manuscript.

References

- Akmaev, R. A., Fuller-Rowell, T. J., Wu, F., Forbes, J. M., Zhang, X., Anghel, A. F., Iredell, M. D., et al. (2008). Tidal variability in the lower thermosphere: Comparison of Whole Atmosphere Model (WAM) simulations with observations from TIMED. *Geophysical Research Letters*, 35, L03810. <https://doi.org/10.1029/2007GL032584>
- Bowman, B., Tobiska, W. K., Marcos, F., Huang, C., Lin, C., & Burke, W. (2008). A new empirical thermospheric density model JB2008 using new solar and geomagnetic indices. AIAA/AAS Astrodynamics Specialist Conference. Honolulu, HI: AIAA.
- Bruinsma, S. L. (2015). The DTM-2013 thermosphere model. *Journal of Space Weather and Space Climate*, 2. <https://doi.org/10.1051/swsc/2012005>
- Bruinsma, S. L., Doornbos, E., & Bowman, B. R. (2014). Validation of GOCE densities and thermosphere model evaluation. *Advances in Space Research*, 54(4), 576–585. <https://doi.org/10.1016/j.asr.2014.04.008>
- Bussy-Virat, C. D., Ridley, A. J., & Getchius, J. W. (2018). Effects of uncertainties in the atmospheric density on the probability of collision between space objects. *Space Weather*, 16, 519–537. <https://doi.org/10.1029/2017SW001705>
- Champion, K. S. W., & Marcos, F. A. (1973). The triaxial-accelerometer system on Atmosphere Explorer. *Radio Science*, 8(4), 297–303.
- Emmert, J. T. (2015). Altitude and solar activity dependence of 1967–2005 thermospheric density trends derived from orbital drag. *Journal of Geophysical Research: Space Physics*, 120, 2940–2950. <https://doi.org/10.1002/2015JA021047>
- Fuller-Rowell, T. J., Akmaev, R. A., Wu, F., Anghel, A., Maruyama, N., Anderson, D. N., Codrescu, M. V., et al. (2008). Impact of terrestrial weather on the upper atmosphere. *Geophysical Research Letters*, 35, L09808. <https://doi.org/10.1029/2007GL032911>
- Fuller-Rowell, T. J., & Evans, D. S. (1987). Height integrated Pedersen and Hall conductivity patterns inferred from the TIROS-NOAA satellite data. *Journal of Geophysical Research*, 92(A7), 7606–7618.
- Fuller-Rowell, T. J., Rees, D., Quegan, S., & Moffett, R. J. (1991). Numerical simulations of the sub-auroral F-region trough. *Journal of Atmospheric and Terrestrial Physics*, 53(6–7), 529–540.
- Fuller-Rowell, T. J., Rees, D., Quegan, S., Moffett, R. J., Codrescu, M. V., & Millward, G. H. (1996). A coupled thermosphere-ionosphere model (CTIM). In R. W. Schunk (Ed.), *Handbook of ionospheric models*, (pp. 217–238). Logan, Utah: Utah State Univ.
- Hagan, M. E., Roble, R. G., & Hackney, J. (2001). Migrating thermospheric tides. *Journal of Geophysical Research*, 106(A7), 12,739–12,752. <https://doi.org/10.1029/2000JA000344>
- Heelis, R. A., Lowell, J. K., & Spiro, R. W. (1982). A model of the high-latitude ionosphere convection pattern. *Journal of Geophysical Research*, 87(A8), 6339. <https://doi.org/10.1029/JA087iA08p06339>
- Hejduk, M. D., & Snow, D. E. (2018). The effect of neutral density estimation errors on satellite conjunction serious event rates. *Space Weather*, 16, 849–869. <https://doi.org/10.1029/2017SW001720>
- Jacchia, L. G., & Slowey, J. (1963). Accurate drag determinations for eight artificial satellites: Atmospheric densities and temperatures. *Smithsonian Contribution Astrophysical*, 8(1), 1–99.
- Knipp, D., Kilcommons, L., Hunt, L., Mlynczak, M., Pilipenko, V., Bowman, B., Deng, Y., et al. (2013). Thermospheric damping response to sheath-enhanced geospace storms. *Geophysical Research Letters*, 40, 1263–1267. <https://doi.org/10.1002/grl.50197>
- Marsh, D. R., Solomon, S. C., & Reynolds, A. E. (2004). Empirical model of nitric oxide in the lower thermosphere. *Journal of Geophysical Research*, 109(A7), A07301. <https://doi.org/10.1029/2003JA010199>
- Maute, A. (2017). Thermosphere-Ionosphere-Electrodynamics General Circulation Model for the Ionospheric Connection Explorer: TIEGCM-ICON. *Space Science Reviews*, 212(1–2), 523–551. <https://doi.org/10.1007/s11214-017-0330-3>
- Millward, G. H., Müller-Wodarg, I. C. F., Aylward, A. D., Fuller-Rowell, T. J., Richmond, A. D., & Moffett, R. J. (2001). An investigation into the influence of tidal forcing on F region equatorial vertical ion drift using a global ionosphere-thermosphere model with coupled electrodynamics. *Journal of Geophysical Research*, 106(A11), 24,733–24,744.
- Müller-Wodarg, I. C. F., Aylward, A. D., & Fuller-Rowell, T. J. (2001). Tidal oscillations in the thermosphere: A theoretical investigation of their sources. *Journal of Atmospheric and Solar - Terrestrial Physics*, 63(9), 899–914.
- Nier, A. O., Potter, W. E., Hickman, D. R., & Mauersberger, K. (1973). The open-source neutral-mass spectrometer on Atmosphere Explorer-C, -D, -E. *Radio Science*, 8(4), 271–276.
- Picone, J. M., Hedin, A. E., Drob, D. P., & Aikin, A. C. (2002). NRLMSISE-00 empirical model of the atmosphere: Statistical comparisons and scientific issues. *Journal of Geophysical Research*, 107(A12), 1468. <https://doi.org/10.1029/2002JA009430>
- Qian, L., Burns, A. G., Emery, B. A., Foster, B., Lu, G., Maute, A., et al. (2014). The NCAR TIE-GCM: A community model of the coupled thermosphere/ionosphere system. In J. Huba, R. Schunk, & G. Khazanov (Eds.), *Modeling the ionosphere-thermosphere system*, *Geophysical Monograph Series* (Vol. 201, 73 pp.). Washington, DC: American Geophysical Union. <https://doi.org/10.1002/9781118704417.ch7>
- Richmond, A. D., Ridley, E. C., & Roble, R. G. (1992). A thermosphere/ionosphere general circulation model with coupled electrodynamics. *Geophysical Research Letters*, 19(6), 601. <https://doi.org/10.1029/92GL00401-604>
- Roble, R. G., Ridley, E. C., Richmond, A. D., & Dickinson, R. E. (1988). A coupled thermosphere/ionosphere general circulation model. *Geophysical Research Letters*, 15(12), 1325. <https://doi.org/10.1029/GL015i012p01325-1328>
- Solomon, S. C., Burns, A. G., Emery, B. A., Mlynczak, M. G., Qian, L., Wang, W., Weimer, D. R., et al. (2012). Modeling studies of the impact of high-speed streams and co-rotating interaction regions on the thermosphere-ionosphere. *Journal of Geophysical Research*, 117(A9), A00L11. <https://doi.org/10.1029/2011JA017417>
- Solomon, S. C., & Qian, L. (2005). Solar extreme-ultraviolet irradiance for general circulation models. *Journal of Geophysical Research*, 110, A10306. <https://doi.org/10.1029/2005JA011160>
- Storz, M. F., Bowman, B. R., Branson, M. J. I., Casali, S. J., & Tobiska, W. K. (2005). High accuracy satellite drag model (HASDM). *Advances in Space Research*, 36(12), 2497–2505. <https://doi.org/10.1016/j.asr.2004.02.020>

- Sutton, E. K. (2018). A new method of physics-based data assimilation for the quiet and disturbed thermosphere. *Space Weather*, 16, 736–753. <https://doi.org/10.1002/2017SW00178>
- Sutton, E. K., Thayer, J. P., Wang, W., Solomon, S. C., Liu, X., & Foster, B. T. (2015). A self-consistent model of helium in the thermosphere. *Journal of Geophysical Research: Space Physics*, 120, 6884–6900. <https://doi.org/10.1002/2015JA021223>
- Weimer, D. R. (2005). Improved ionospheric electrodynamic models and application to calculating Joule heating rates. *Journal of Geophysical Research*, 110(A5), A05306. <https://doi.org/10.1029/2004JA010884>