

Forecasting global and multilevel thermospheric neutral density and ionospheric electron content by tuning models against satellitebased accelerometer measurements

Forootan, Ehsan; Kosary, Mona; Farzaneh, Saeed; Kodikara, Timothy; Vielberg, Kristin; FernandezGomez, Isabel ; Borries, Claudia; Schumacher, Maike

Published in:
Scientific Reports

DOI (link to publication from Publisher):
[10.1038/s41598-022-05952-y](https://doi.org/10.1038/s41598-022-05952-y)

Creative Commons License
CC BY 4.0

Publication date:
2022

Document Version
Publisher's PDF, also known as Version of record

[Link to publication from Aalborg University](#)

Citation for published version (APA):
Forootan, E., Kosary, M., Farzaneh, S., Kodikara, T., Vielberg, K., FernandezGomez, I., Borries, C., & Schumacher, M. (2022). Forecasting global and multilevel thermospheric neutral density and ionospheric electron content by tuning models against satellitebased accelerometer measurements. *Scientific Reports*, 12, Article 2095 . <https://doi.org/10.1038/s41598-022-05952-y>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal -

Take down policy

If you believe that this document breaches copyright please contact us at vbn@aub.aau.dk providing details, and we will remove access to the work immediately and investigate your claim.



OPEN

Forecasting global and multi-level thermospheric neutral density and ionospheric electron content by tuning models against satellite-based accelerometer measurements

Ehsan Forootan^{1✉}, Mona Kosary², Saeed Farzaneh², Timothy Kodikara³, Kristin Vielberg⁴, Isabel Fernandez-Gomez³, Claudia Borries³ & Maike Schumacher¹

Global estimation of thermospheric neutral density (TND) on various altitudes is important for geodetic and space weather applications. This is typically provided by models, however, the quality of these models is limited due to their imperfect structure and the sensitivity of their parameters to the calibration period. Here, we present an ensemble Kalman filter (EnKF)-based calibration and data assimilation (C/DA) technique that updates the model's states and simultaneously calibrates its key parameters. Its application is demonstrated using the TND estimates from on-board accelerometer measurements, e.g., those of the Gravity Recovery and Climate Experiment (GRACE) mission (at ~ 410 km altitude), as observation, and the frequently used empirical model NRLMSISE-00. The C/DA is applied here to re-calibrate the model parameters including those controlling the influence of solar radiation and geomagnetic activity as well as those related to the calculation of exospheric temperature. The resulting model, called here 'C/DA-NRLMSISE-00', is then used to now-cast TNDs and individual neutral mass compositions for 3 h, where the model with calibrated parameters is run again during the assimilation period. C/DA-NRLMSISE-00 is also used to forecast the next 21 h, where no new observations are introduced. These forecasts are unique because they are available globally and on various altitudes (300–600 km). To introduce the impact of the thermosphere on estimating ionospheric parameters, the coupled physics-based model TIE-GCM is run by replacing the O₂, O₁, He and neutral temperature estimates of the C/DA-NRLMSISE-00. Then, the non-assimilated outputs of electron density (N_e) and total electron content (TEC) are validated against independent measurements. Assessing the forecasts of TNDs with those along the Swarm-A (~ 467 km), -B (~ 521 km), and -C (~ 467 km) orbits shows that the root-mean-square error (RMSE) is considerably reduced by 51, 57 and 54%, respectively. We find improvement of 30.92% for forecasting N_e and 26.48% for TEC compared to the radio occultation and global ionosphere maps (GIM), respectively. The presented C/DA approach is recommended for the short-term global multi-level thermosphere and enhanced ionosphere forecasting applications.

Space weather describes physical processes caused by the Sun's radiation of energy. The manifestations of space weather are multiple, e.g., the variations of the Earth's magnetic field or the changing states of the upper atmosphere—between the altitude of around 100 km up to 2000 km—comprising both the thermosphere and the ionosphere. An accurate quantification of the thermospheric neutral density (TND) and thermospheric temperature

¹Geodesy Group, Department of Planning, Aalborg University, Rendburggade 14, 9000 Aalborg, Denmark. ²School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran, P.O. Box 113654563, Tehran, Iran. ³Institute of Solar-Terrestrial Physics, German Aerospace Center (DLR), Kalkhorstweg 53, 17235 Neustrelitz, Germany. ⁴Institute of Geodesy and Geoinformation, University of Bonn, Nussallee 17, 53115 Bonn, Germany. ✉email: efo@plan.aau.dk

has become increasingly important in the space physics and geodesy community¹. Atmospheric drag has a direct relationship with the TND estimates and represents a considerable impact on the precise orbit determination (POD) and prediction of low Earth orbit (LEO) satellites, for example those with the altitude of less than 1000 km, as well as space debris². The ionosphere variability impacts operational communications and navigation systems, thus, its now-casting and forecasting is extremely important for many applications^{3–7}. Energy and momentum are transferred from the lower to the upper atmosphere and ionosphere through the generation and propagation of waves. These interactions make the thermosphere-ionosphere a coupled system that exhibit non-linear and complex interactions with many inherent time-scales⁸.

The accelerometer sensors on-board of satellites, which measures the non-gravitational forces acting on them, have been emerged as useful tools for estimating TND and its changes⁹. The LEO missions including the Challenging Minisatellite Payload (CHAMP, 2000–2010)¹⁰, the Gravity Recovery and Climate Experiment (GRACE, 2002–2017)¹¹ and its Follow-On mission (GRACE-FO, launched in 2018)¹², Gravity field and steady-state Ocean Circulation Explorer (GOCE, 2009–2013)¹³ and the European Space Agency (ESA)'s Swarm mission (Swarm- A, -B, and -C launched in 2013)¹⁴, are applied by various centers to estimate TNDs, e.g., from the Delft Technical University¹⁵, European Space Agency¹⁶, Mehta et al.¹⁷, and Vielberg et al.¹⁸. However, these measurements are available only along-track of these missions, thus, their application for studying global and multi-level neutral density changes in the thermosphere is limited. In this study, we will provide a technique to address this limitation, and we will demonstrate the benefits of accelerometer-based TND estimates for global and multi-(vertical) level thermosphere and ionosphere studies.

Empirical and physical models are developed to address the demands for precise estimation of thermospheric variables such as the TND and its individual neutral compositions, thermospheric temperature, as well as ionospheric parameters such as total electron content (TEC), the three-dimensional electron density, and its evolution in time (fourth dimension)¹⁹. Physical models are based on physical laws and principles such as continuity, energy and momentum equations and require the solution of partial differential equations^{20,21}, while empirical models are driven by fitting mathematical equations to actual measurements, as provided from one or more observation systems, during an appropriate time period^{22–25}. The current status of measuring and modeling techniques related to thermosphere and ionosphere is reviewed by Palmroth et al. (2021)²⁶.

A logical step to make the best use of observations and models can be realized through the mathematical merging frameworks that build a connection between them. For example, (1) correction fields are applied as a ratio of TNDs from LEO satellites and those of models^{27–33}. However, this method can only be used for reanalyzing and now-casting the thermosphere, and the reliability of these ratio fields, derived by a limited number of satellite tracks, might be treated with caution. (2) Statistical data-model integration frameworks apply decomposition techniques³⁴ to extract the dominant structures of the upper atmosphere dynamics. Then, predicting steps are performed using state-space techniques such as the Kalman filter (KF)³⁵. Previous studies^{36–40} successfully applied this approach during geomagnetic storms because the pronounced temporal and spatial changes during these events enhance the model-data integration. This technique might be less efficient during calm periods, where the level of uncertainties of models and data is comparable. (3) Sequential data assimilation (DA) techniques are found to be efficient for merging observations and model outputs, while decreasing the model uncertainties^{41–45}. However, DA techniques mostly focus on updating the model states and they are not very efficient to extend the along-track measurements to cover the whole globe. Besides, they are not very efficient for forecasting. (4) The calibration and data assimilation (C/DA) can provide the opportunity to update the model's states (similar to DA) and can simultaneously calibrate the model's key parameters^{46,47}. The advantage of this approach is that the calibrated parameters can be used to simulate TNDs in locations outside of the orbital tracks of LEO satellite, or to forecast TNDs.

The methodology of this study follows that of Forootan et al.⁴⁷, categorized as (4), who proposed a C/DA method based on the sequential ensemble Kalman filter (EnKF) to calibrate the NRLMSISE-00 model²². In their study, CHAMP and GRACE TND estimates, normalised at 400 km altitude, were used as observation. They applied C/DA to estimate the daily calibrated key parameters of NRLMSISE-00, which referred to the solar radiation, geomagnetic activity, and calculation of the exospheric temperature. Then the calibrated parameters were used to forecast the global TND fields of the next day at 400 km altitude.

Here, we extend the previous studies by applying C/DA while using along-track TNDs without height-dependent normalisation, i.e., using GRACE data at its altitude, e.g., ~ 410 km during February 2015. The new model that contains the re-calibrated parameters through C/DA is named here as 'C/DA-NRLMSISE-00'. The effect of this (re-)calibration is examined by running the C/DA-NRLMSISE-00 up to 21-h in forecasting-mode (i.e., without introducing new observations) and estimating the global TND as well as individual neutral mass densities covering the altitudes of 300–600 km. TND estimations from the POD analysis of Swarm data at the altitude range of 470–520 km (during February 2nd–28th, 2015) are then used for validation.

To investigate the thermosphere-ionosphere interactions, the impact of the C/DA-derived neutral mass density estimations is studied on the ionosphere parameters, e.g., electron density and total electron content (TEC). This investigation was missing in most of the existing literature. For this, the physics-based model TIE-GCM²⁰ is chosen as basis. This model was run twice, i.e., first with its original primary history fields. In the second run, we replaced some of the history fields, which are mass mixing ratios of atomic oxygen (O), molecular oxygen (O₂), and helium (He), as well as the neutral temperature (TN) using the forecasting estimates derived from C/DA-NRLMSISE-00. The second run is considered to take advantage of the improved representation of thermospheric parameters thus the new runs are named 'TIE-GCM-I'. For validation, the electron density and TEC estimates of the original TIE-GCM and TIE-GCM-I are validated against the measurements from COSMIC⁴⁸ and the global ionosphere maps (GIM)⁴⁹. Thus, the investigations related to the ionosphere part (i.e., during February 2015) are completely independent from the type of measurements that we used in the C/DA procedure.

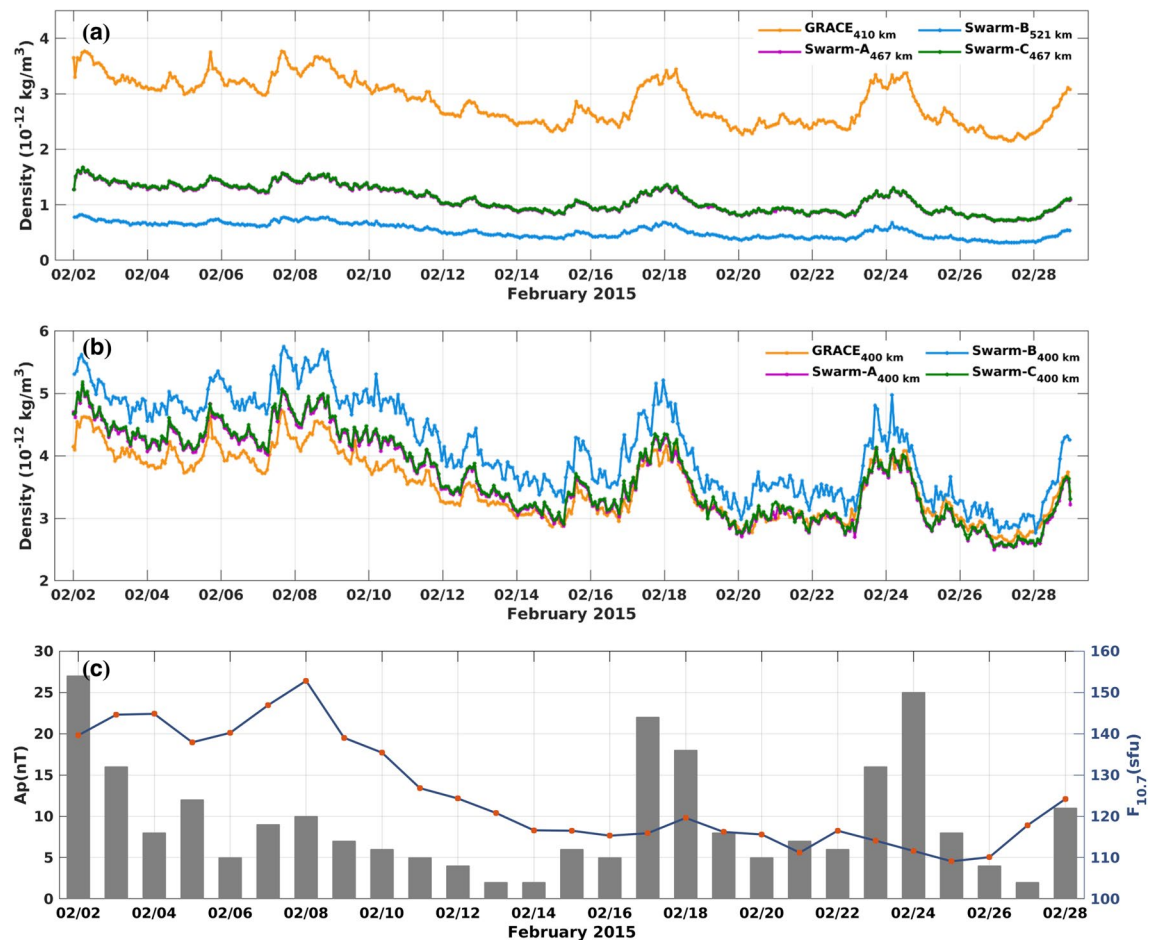


Figure 1. From top to bottom: (a) orbit-averaged TND from GRACE, Swarm-A, -B, and -C. on their original altitude. (b) Time series of orbit-averaged TND from GRACE, Swarm-A, -B, and -C normalised on a common altitude of 400 km. (c) Time series of the solar ($F_{10.7}$) and geomagnetic activity (A_p) indices during February 2015. The plots were generated using MATLAB (version R2021a, <https://www.mathworks.com/>).

Data and model

Space-borne thermospheric neutral density. Space-borne TND estimates, which will be used as observations in the C/DA procedure and for validation are derived from GRACE¹¹ and the ESA's Swarm mission¹⁴, respectively. GRACE-TNDs with 30 s sampling rate are available by Vielberg et al. (2021)¹⁸. The TNDs of Swarm POD (at 30-s sampling rate) are downloaded from the Delft Technical University's website¹⁵. More details about these data can be found in the Supplementary file (Section S2 online). Figure 1 provides an overview of the data used in this study. Panel (a) presents time series of the mean TND estimates derived from GRACE and Swarm missions along their orbits at their original altitude during February 2015, and (b) represents them at the common altitude of 400 km. The height-dependent function and vertical profile of NRLMSISE-00 model was used to transform the TNDs to a common altitude of 400 km (see the Supplementary file Section S3 online), which makes it easier to compare the consistency of TNDs derived from GRACE and Swarm-A/-B/-C⁴⁷. In fact, this figure demonstrates that a bias exists between the TND estimates of various satellite missions and the mean bias of $-1.67 \times 10^{-13} \text{ kg/m}^3$ is estimated between the four TND estimates of Fig. 1b. Based on these results, one may conclude that the estimation of TND from Swarm-B that is on much higher altitude (on average 520 km) contains bigger biases than the other missions. Thus, comparisons with this estimation might be done with caution. It is worth mentioning that in our study the multi-mission data are not used within the C/DA and the procedure is implemented only for GRACE. But for validation we use the Swarm data while taking into account the uncertainties that are shown in Fig. 1b. In Panel (c), the variation of solar and planetary geomagnetic activity are shown using $F_{10.7}$ and A_p indices, respectively.

Figure 1 indicates that the trend of changes in TNDs depends on the variation of solar activity. For example, compare the peaks of (b) and (c), where the maximum value of TNDs corresponds to the day with the high solar activity (i.e., February 8th, 2015). Sudden peaks in A_p for example, during February 18–19th and 23–24th, 2015 also caused changes in the magnitude of TNDs (compare the gray bars of Fig. 1c with the time series in Fig. 1b).

Electron density profile from radio occultation measurements. To validate the ionosphere related parameters, N_e profiles of the original TIE-GCM model and TIE-GCM-I are compared with the COSMIC radio

occultation⁴⁸. The COSMIC constellation includes six satellites, which was launched on April 15th, 2006. Their initial orbit altitude was 500 km, but it was gradually raised to 800 km. In this study, the second-level ‘ionPrf’ data products⁴⁸ is used, whereas their accuracy is evaluated to be around 10^4 – 10^5 cm⁻³, see⁴⁸. Before assessing the TIE-GCM-I with RO data, it is necessary to apply some quality control tests on the individual ionospheric electron density profiles. For this, the least squares method is applied to fit a two-layer Chapman function to each profile⁵⁰. This makes the best fit with RO electron density profiles at the F2 layer. In addition, we estimated the mean deviation of these profiles to quantify the effect of ionospheric plasma irregularities on the height-dependent variability of the electron density⁵¹.

Global ionosphere maps (GIM). As our second assessment, related to the ionosphere parameters, the TEC forecasts of TIE-GCM-I are compared with those of the IGS (International GNSS Service) products⁴⁹. Since 1998, Based on GNSS dual-frequency code and phase measurements from globally distributed IGS tracking stations and seven Ionospheric Associate Analysis Centers (IAACs) have established products known as the global ionosphere maps (GIMs) in the IONEX (ionosphere exchange) format. GIMs provide vertical TEC (VTEC) estimates (here we simply call it TEC) in terms of the spherical harmonics expansion up to degree and order 15 or in the grid domain with the spatial resolution of $2.5^\circ \times 5^\circ$ in latitude and longitude, respectively. Their temporal resolution is 15 min to 2 h. These products are available with a latency of less than 24 h and approximately 11 days in the rapid and final solution modes, respectively⁵². In this study, the rapid global TEC maps with 15 min time interval obtained from the UPC center⁵³ are used for validating TEC estimates of TIE-GCM-I, globally.

NRLMSISE-00 model. The Naval Research Laboratory’s Mass Spectrometer Incoherent Scatter Extension 2000 (NRLMSISE-00) model is an empirical model to simulate thermospheric variables²². The difference between NRLMSISE-00 and new version NRLMSISE2.0 can be found in the model outputs at low altitudes due to the consideration of new data below 100 km, so the impact of new model at the GRACE and Swarm orbits is negligible. The model uses various sources of data, see Fig. 2.9 in Kodikara (2019)⁵⁴ including drag measurements and satellite-borne accelerometer data, incoherent scatter radar, mass spectrometer and pressure gauge over several decades (1961–1998)²². The model estimates depend on the altitude, geodetic latitude and longitude, local apparent solar time, solar flux for the previous day ($F_{10.7}$)⁵⁵ and its 81-day average ($F_{10.7A}$), geomagnetic index (A_p)⁵⁶ and 3205 model coefficients: *pavgm*, *pd*, *pdl*, *pdm*, *pma*, *ps*, *pt*, *ptl*, *ptm* and *sam*, which are adjusted during the model’s calibration period, and used to compute the number densities of He, O, N₂, O₂, Ar, H, N, anomalous oxygen and quantities such as TND, as well as neutral temperature.

To increase the efficiency of C/DA, the global sensitivity analysis (GSA) method⁵⁷ is applied to identify key parameters with high impact on simulating TNDs in NRLMSISE-00. These parameters can be effectively calibrated by the accelerometer-derived TND estimates, see also Forootan et al. (2020)⁴⁷. The covariances between the TND estimates and nineteen parameters of NRLMSISE-00 are shown in the Supplementary file (Section S4 online (Fig. S1)), which indicate the consistency between the GSA results from Forootan et al. (2020)⁴⁷ and the empirical covariances between ensembles and model parameters. These results show that the first component of *ptm* and *pt* model coefficients, which are associated with estimating exospheric temperature, have the largest impact on estimating TNDs. To account for the sudden changes in solar and magnetic activities, which are not well reflected in the model coefficients, we replace the two constants of 150 and 4 in the model structure that are used to compute the index anomalies, i.e., $dF_{10.7A} = F_{10.7A} - 150$ and $dA_p = A_p - 4$ are changed to $dF_{10.7A} = F_{10.7A} - C_{F_{10.7A}}$ and $dA_p = A_p - C_{A_p}$. Therefore, these four variables (*ptm*(1), *pt*(1), $C_{F_{10.7A}}$ and C_{A_p}) are estimated within the C/DA procedure. The GRACE derived TND estimates during February 2015 are used as an example to test the performance of the C/DA approach.

Thermosphere-ionosphere-electrodynamics general circulation model (TIE-GCM). TIE-GCM is a coupled thermosphere-ionosphere model that uses a finite difference scheme to solve the nonlinear equations of conservation of mass, energy, and momentum for neutral and ion species^{58,59}. This study is based on the TIE-GCM version 2.0 (released on March 21st, 2016). The horizontal resolution of this model is set to $5^\circ \times 5^\circ$, and the vertical resolution is two levels per scale height. The altitude of the model extends from approximately 97 km to 450–600 km depending on the solar activity⁵⁴. In the TIE-GCM, the EUVAC (Extreme Ultraviolet Flux model for Aeronomic Calculations) empirical solar proxy model^{60,61} provides the solar irradiance inputs via the daily $F_{10.7}$ and its 81-day averaged ($F_{10.7A}$) time series. This model uses the *Kp* index⁶² instead of the *Ap* index for reflecting geomagnetic activity. Throughout this work, the Heelis model is been used to specify the high latitudes ion convection⁶³. The Global Scale Wave Model (GSWM) provides the lower boundary condition, which is related to the atmospheric tides⁶⁴.

To run TIE-GCM, one needs to introduce primary history files, which include the prognostic fields that are necessary to start the model. These files contain variables such as the neutral and ion temperature, neutral zonal and meridional wind, molecular and atomic Oxygen, Nitric Oxide, Helium, Argon, O⁺ ion, electron temperature and density, O₂⁺ ion, vertical motion, geopotential height and electric potential⁵⁹. The differences between thermospheric constituents derived from TIE-GCM and C/DA-NRLMSISE-00 is discussed in the Supplementary file (Section S5 online). The density of atomic oxygen (O), molecular oxygen (O₂), helium (He), and neutral temperature (TN) from the C/DA-NRLMSISE-00 will replace the primary history files of TIE-GCM (thus named as TIE-GCM-I) to test their impact on the forecasts of TEC and electron density variables.

Methods

Calibration and data assimilation (C/DA) framework to form C/DA-NRLMSISE-00. In this study, the simultaneous C/DA^{46,47} technique is applied to integrate accelerometer-derived TND measurements into NRLMSISE-00. The C/DA approach contains a model-state equation, where the model derived integral of all neutral mass estimates, i.e., TNDs, and four model parameters are the unknowns. These are computed simultaneously through a sequential ensemble Kalman filter (EnKF)⁶⁵ procedure. First, let us assume that our original model, i.e., NRLMSISE-00²², is mathematically represented as:

$$\text{Original model, i.e., NRLMSISE-00 : } F(\Theta) = F(\Theta_P, \Theta_R, \Theta_I), \quad (1)$$

where Θ is a vector of model parameters and input values that enable the model to be run and simulate TND, individual mass densities, and thermospheric temperature at arbitrary locations and time steps. In Eq. (1), we consider that Θ contains $\Theta_{P4 \times 1}$ that are the four key parameters from the global sensitivity analysis (discussed in the Supplementary file, Section S6 online) and will be updated during the C/DA procedure. Those parameters that remain unchanged during C/DA is shown by Θ_R . Finally, Θ_I indicates the input variables such as the values of solar and geomagnetic indices, location (longitude, latitude, and height), and time.

During the C/DA procedure, accelerometer-derived TNDs are used as observations at the time step to correct the TND states of NRLMSISE-00 and to calibrate its key (here 4) parameters through minimizing the following cost function:

$$J(\mathbf{X}) = \frac{1}{2}[\mathbf{X} - \bar{\mathbf{X}}^b]^T (\mathbf{P}^b)^{-1} [\mathbf{X} - \bar{\mathbf{X}}^b] + \frac{1}{2}[\mathbf{H}\mathbf{X}^b - \mathbf{Y}]^T \mathbf{R}^{-1} (\mathbf{H}\mathbf{X}^b - \mathbf{Y}), \quad (2)$$

where N is the ensemble member ($N = 75$), and $\mathbf{X}_k^b = [\mathbf{x}_k^{b(1)}, \dots, \mathbf{x}_k^{b(N)}]$ is the ensemble of background state (i.e., $\mathbf{x}^{b(i)} \in \mathfrak{R}^n$) and is composed of two parts: the ensemble of model state and model parameters. $\bar{\mathbf{X}}^b$ is the ensemble mean vector and \mathbf{P}_k^b is the background error covariance. Ensembles of observations are stored in $\mathbf{Y}_k = [\mathbf{y}^1, \dots, \mathbf{y}^N]$ (i.e., $\mathbf{y}^i \in \mathfrak{R}^m$), which are perturbed by \mathbf{R}_k , which is the covariance matrix of observations. Assuming that the measurements of TNDs to be independent, the magnitude of biases in Fig. 1b, transformed to the satellites' altitudes, is considered as uncertainty.

The analysis state estimates (\mathbf{X}^a) are obtained while taking into account the measurements and the cross-correlations of model states and parameters as:

$$\mathbf{X}^a = \mathbf{X}^b + \mathbf{K}(\mathbf{Y} - \mathbf{H}\mathbf{X}^b), \quad (3)$$

where \mathbf{K} is the Kalman gain, and the design matrix $\mathbf{H} : \mathfrak{R}^n \rightarrow \mathfrak{R}^m$ connects the model states and parameters to the observations. Equation (3) was evaluated at time k to obtain the ensemble of analysis estimate \mathbf{X}^a . The state can then be used for forecasting \mathbf{X}^b of the next step and the C/AD process continues until the observations are available. The C/DA procedure of this study is performed using 3 h of GRACE-TNDs. The last set of key parameters derived from the last ensemble mean (in Eq. (3)) are considered as the updated key parameters shown by $\hat{\Theta}_P$. These parameters then replace the default values of the original NRLMSISE-00 model (Eq. (1)) to form the C/DA-NRLMSISE-00 model as:

$$\text{C/DA-NRLMSISE-00 model, i.e., : } F(\hat{\Theta}_P, \Theta_R, \Theta_I), \quad (4)$$

C/DA-NRLMSISE-00 can be applied to estimate TNDs globally (i.e., even the regions that are not covered by GRACE orbits) and for forecasting the global map of TNDs, individual mass densities and thermospheric temperature of the next 1 h to 21 h (see the next section). The mathematical details of C/DA are described in the Supplementary file (Section S6 online).

Statistical measures to examine the quality of model simulations. To evaluate the performance of models, the following statistical metrics are applied.

- The Root-Mean-Squares of Error (RMSE) is applied as a scale-dependent measure to extract how well these models agree with the measurements as:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (\text{Obs} - \text{Model})^2}{n}}, \quad (5)$$

where Obs and Model denote observation and model estimates, respectively and n is the number of observations.

- The reduction of RMSE is interpreted as 'improvement', which is computed as: where RMSE_1 is the differences between the original model (e.g., NRLMSISE-00 or TIE-GCM) and observations (e.g., neutral densities from GRACE, Swarm-A /-B and -C, and electron density profiles from RO), and RMSE_2 is determined from differences between C/DA-NRLMSISE-00 or TIE-GCM-I models and observations.

$$\text{Improvement} = 100 \times \frac{\text{RMSE}_1 - \text{RMSE}_2}{\text{RMSE}_1}, \quad (6)$$

The Average Absolute Percentage Deviations (AAPD) is used as a scale-independent measure, i.e.,

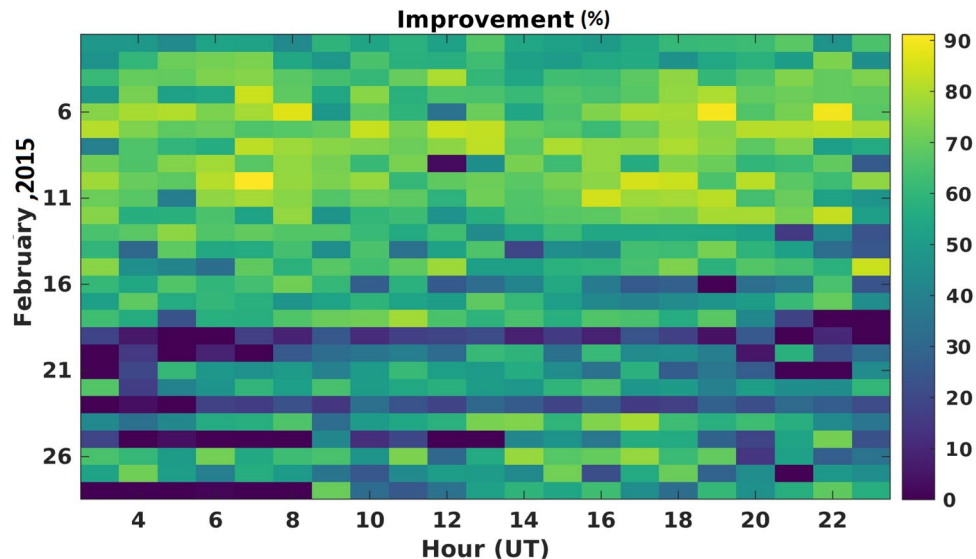


Figure 2. An overview of the improvements (from Eq. (6)) in forecasting TNDs during February 2015. The results are estimated by running the original NRLMSISE-00 and C/DA-NRLMSISE-00. The C/DA is processed using GRACE-TNDs as observations, then the 3-hourly calibrated parameters are used to form the C/DA-NRLMSISE-00 model and forecasting TNDs along Swarm-C during different times of the month. The quality of forecasts is found to be 50.62% improved, on average. The plot was made in MATLAB (version R2021a, <https://www.mathworks.com/>).

$$AAPD = \frac{\sum_{i=1}^n (100 \times \frac{|\text{Obs}_i - \text{Model}_i|}{\text{Obs}_i})}{n}, \quad (7)$$

where, the minimum (maximum) values for AAPD indicate that the model provides on average best (worst) performance in prediction.

- The Normalized Root Mean Square Error (NRMSE) is defined by comparing the models against the observations to represent the fraction of data-variance that can be explained by these models. The computation follows:

$$\text{NRMSE} = 1 - \frac{\sqrt{(\text{Obs} - \text{Model})^2}}{\sqrt{(\text{Obs} - \bar{\text{Obs}})^2}}, \quad (8)$$

where $\bar{\text{Obs}}$ is defined as the mean of observations. In contrast to AAPD, the minimum (maximum) values for NRMSE indicate that the model provides on average worst (best) performance in estimating the variable of interest.

- The expression of difference between model and observation in percentage is computed based on the 'Relative Error (RE)' as:

$$\text{RE} = 100 \times \frac{|\text{Obs} - \text{Model}|}{\text{Obs}}, \quad (9)$$

where $|\cdot|$ computes the absolute values.

Setup of the C/DA-NRLMSISE-00 model. The specifics of the C/DA configuration and parameter settings used in this study are as follows: the ensemble size is selected to be 75. The assimilation window size is determined experimentally by changing its size from 1 to 5 h. In each experiment, the calibrated parameters were used to predict the TNDs of the next hours. The minimum RMSE (Eq. 5), against Swarm-A, -B, and -C TNDs, was found to be related to the window size of 3 h. To define an optimal period of the forecast, we implemented different prediction periods from next 1 h to the next 21 h. This means that, for example, the TNDs of 00:00:00 UT to 02:59:30 h UT are used as observations, then the C/DA-NRLMSISE-00 predicts the neutral density values for the next 21 h from 03:00:00 UT to 23:59:30 UT. Figure 2 indicates the improvements in the RMSE values of forecasts for all the 24 h of each day during February 2015. The results indicate that the quality of forecast are generally high (50.62% on average) in most of the days. During days with low (e.g., February 20th, 2015) or sudden changes in the geomagnetic index (e.g., February 24th, 2015), the quality is slightly less (32.34 and 13.46% on average), which can be related to the uncertainties of F10.7 and A_p indices in the forecasting mode. But overall, the C/DA-NRLMSISE-00 is found to be better than the original NRLMSISE-00 in simulating TNDs.

The initial ensembles of C/DA are generated by perturbing the key parameters (i.e., $ptm(1)$, $pt(1)$, $C_{F,107}$ and C_{Ap}) for each ensemble member N and the parameters are sampled from a normal distribution. The mean values for the key parameters $ptm(1)$, $pt(1)$, $C_{F,107}$ and C_{Ap} are 1041.3, 0.9865, 150 and 4, respectively, corresponding

to the default values in the original NRLMSISE-00 model. The variances of the perturbation for $ptm(1)$, $pt(1)$, $C_{F,107}$ and C_{Ap} are selected to be 10.41, 0.0098, 2 and 1, respectively (i.e. around 10% of their default values). During the 3 h of C/DA window, four parameters of $\hat{\Theta}_P$ are estimated, then they are used in Eq. (4) for forecasting thermospheric variables.

Results and discussion

An overview of the work-flow of this study to apply C/DA on NRLMSISE-00 and testing its performance in forecasting thermospheric and ionospheric variables is presented in Fig. 3. The procedure is started by (1) using the GRACE-TND observations to re-calibrate the four key parameters of NRLMSISE-00 (i.e. $ptm(1)$, $pt(1)$, $C_{F,107}$ and C_{Ap}) through C/DA. In (2), the TND values during the 1-h to 21-h forecast mode are compared with the independent measurements derived from Swarm. In (3), some thermospheric constituents derived from the C/DA-NRLMSISE-00 model are used to replace the primary history files of TIE-GCM to test whether the interactions between thermosphere and ionosphere can be improved. Another validation step is performed in (4), the original TIE-GCM and TIE-GCM-I (with improved primary history files) is compared with ionospheric observations such as TEC maps⁶⁶ and RO profiles⁶⁷.

Along-track validation of TNDs from C/DA-NRLMSISE-00 during the analysis phase. The focus of this study is on forecasting TNDs based on C/DA. However, it is essential to validate whether this method can improve NRLMSISE-00 during the data assimilation window (analysis phase). The validation can be carried out using the data that were not used in the C/DA procedure. Thus, here, we use the GRACE-TNDs as observation of the C/DA, and the TNDs along the Swarm-A/-B and -C orbits are used for validation. Figure 4(left) presents the improvements in terms of RMSEs between the TNDs of NRLMSISE-00 and C/DA-NRLMSISE-00. In addition, Fig. 4(right) shows the time series of TNDs (during February 2015) used for these estimations. Based on the statistical results shown in Table 1, after implementing the C/DA, the overall RMSE during the entire month is reduced by 53, 59, and 56%, respectively, along Swarm-A, Swarm-B, and Swarm-C orbits compared to the original model.

These numerical results illustrate the ability of C/DA-NRLMSISE-00 in estimating TNDs with high accuracy compared to the original model. Please note that during these experiments only calibrated parameters derived from the C/DA procedure are used to run the model in the analysis and forecasting phases. We do not present the updated TNDs that are co-estimated with the four parameters, because those values are well fitted to GRACE measurements during the analysis phase due to the least squares optimization. Therefore, from the results presented in this paper, the fit of the estimates during the analysis phase is not 100%. For instance the fitting parameter between model and observation is increased (in terms of NRMSE) from -0.09, 0.16 and -0.07 to 0.48, 0.64 and 0.53 along the Swarm-A, Swarm-B, and Swarm-C orbits, respectively. Nevertheless, the fitted estimates validated with the data from the Swarm mission is presented in Fig. 4 and Table 1. We also found a drop in the estimated improvements around February 19th, 2015 but this does not correspond to the lack of the skill of the two models. This is caused by the small differences between the absolute values of models and measurements (see Fig. 4(right)). A statistical T-test is performed between the observation and TNDs from NRLMSISE-00 and C/DA-NRLMSISE-00 along the Swarm-A, Swarm-B, and Swarm-C orbits. The significance level was set at 90%. The test decision indicates that the changes in the performance of C/DA is statistically significant. In the light of these positive results, we further assess C/DA-NRLMSISE-00 during the forecast phase.

Forecasting of C/DA-NRLMSISE-00 TNDs along the daily tracks of GRACE and Swarm. In this section, the estimated model parameters for each data assimilation window (i.e., 3 h) are applied to forecast TNDs for the next hour. Figure 5 presents the forecast of TND maps derived from the original NRLMSISE-00, C/DA-NRLMSISE-00 and those of GRACE and Swarm as a function of time and the argument of latitude (the angle along the orbital path from the ascending node to the spacecraft's position in the direction of the spacecraft's motion) during 27 days (February 2nd–28th, 2015). A visual comparison indicated that the TNDs of C/DA-NRLMSISE-00 (in the forecast phase) are much closer to the observations. Higher spatial correlation of 98.28, 98.25, 97.53, and 98.25% are found between C/DA-NRLMSISE-00 and GRACE, Swarm-A, -B, and -C compared to those of the original NRLMSISE-00, i.e., 88.34, 90.76, 92.01, and 90.75%.

An average of the daily assessment during the forecast mode (based on Eq. (4)) is presented in Fig. 6, where an average improvement of 62.98% is obtained during February 2nd–17th, 2015 and an average 39.54% during February 18th–28th, 2015. The small improvement value during the second period corresponds to the considerable geomagnetic variability, reflected in the A_p values (see Fig. 1) that negatively affects the optimization using a short-length C/DA window. Nevertheless, these improvements show that C/DA is very effective during days with medium and high solar activity (see Fig. 1, panel (c)), where the original NRLMSISE-00 shows considerable uncertainties⁶⁸. The monthly average of statistical measures during the forecasting mode is provided in Table 2.

Multi-level variations of the thermospheric temperature and neutral density. The four calibrated parameters do not only affect the simulation of thermospheric variables at the altitude of the assimilated observations (here GRACE TND data), their impacts can also be sensed on other altitudes mainly because of the relationship between the two parameters ($ptm(1)$ and $pt(1)$) and the calculation of thermospheric temperature. To illustrate the changes, temperature and density profiles of 100–500 km are presented in Fig. 7, where they correspond to two days with low and high solar activity on February 26th and 8th, 2015 at 12:00 UT. These profiles are shown by the grey and red lines, respectively, and they are associated with the latitude 0° and longitude 0°. These results indicate that by applying C/DA both profiles are modified, whereas the temperature during high (low) solar activity was on average increased by 34.34 (43.26) degree covering the altitude of 200 to 500 km. This

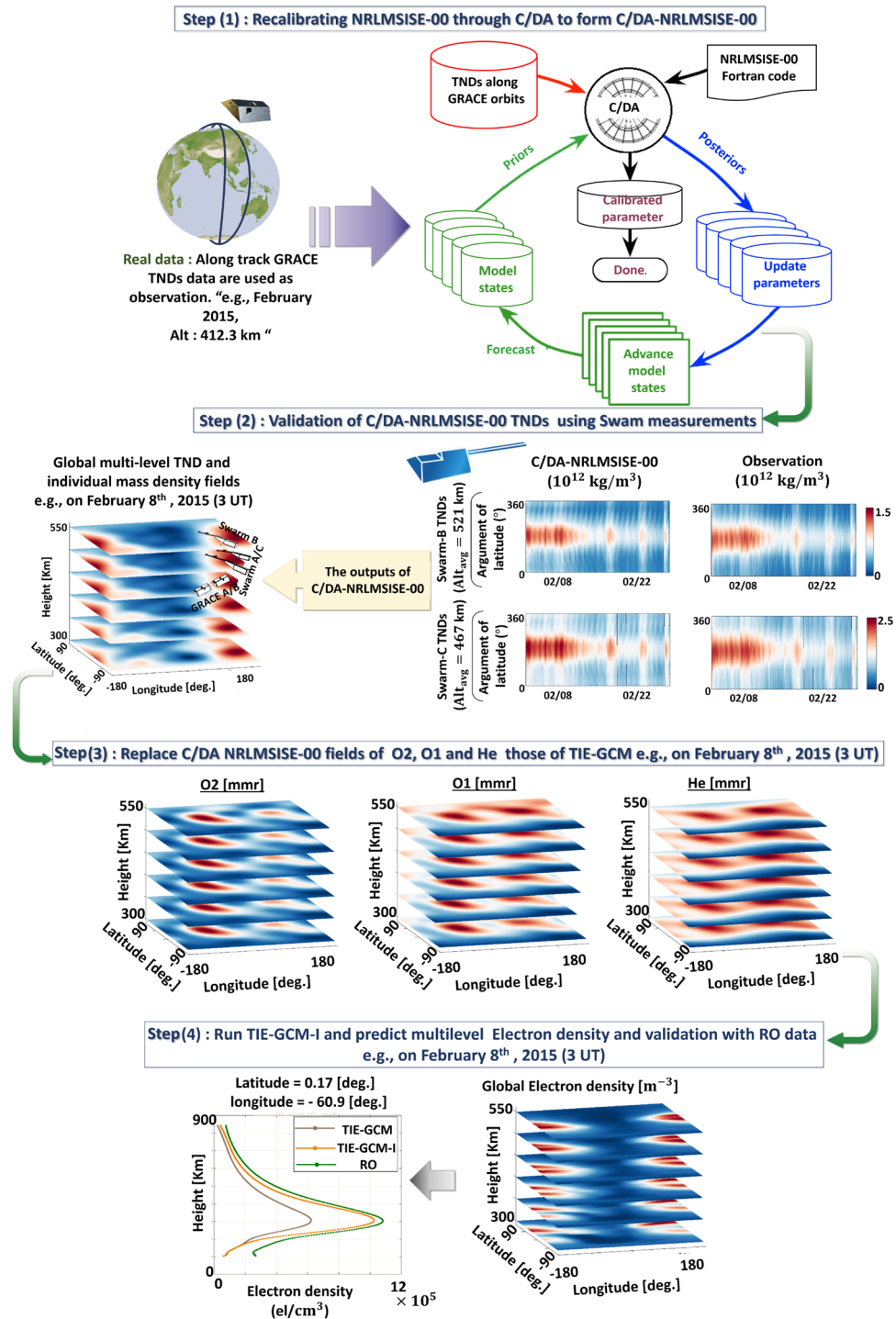


Figure 3. An overview of the proposed C/DA procedure and validation experiments. The procedure is divided into four steps: (1) Applying the C/DA on the NRLMSISE-00 model to form C/DA-NRLMSISE-00, (2) Validating TNDs against Swarm measurements, (3) Improving the primary history files of TIE-GCM using the forecasting values provided by C/DA-NRLMSISE-00, and (4) validating the ionospheric variables using GNSS and RO measurements. The plots were generated using MATLAB (version R2021a, <https://www.mathworks.com/>) and PowerPoint (2016, <https://www.microsoft.com/>) was used for producing the figure.

is equivalent with 7.52×10^{-12} (8.50×10^{-12}) changes in density estimates, respectively. It is worth mentioning that the changes in thermosphere are more complex during storms. Especially, variations due to winds become more pronounced and the barometric processes get violated, see e.g., Lie et al.⁶⁹. Therefore, adding wind and

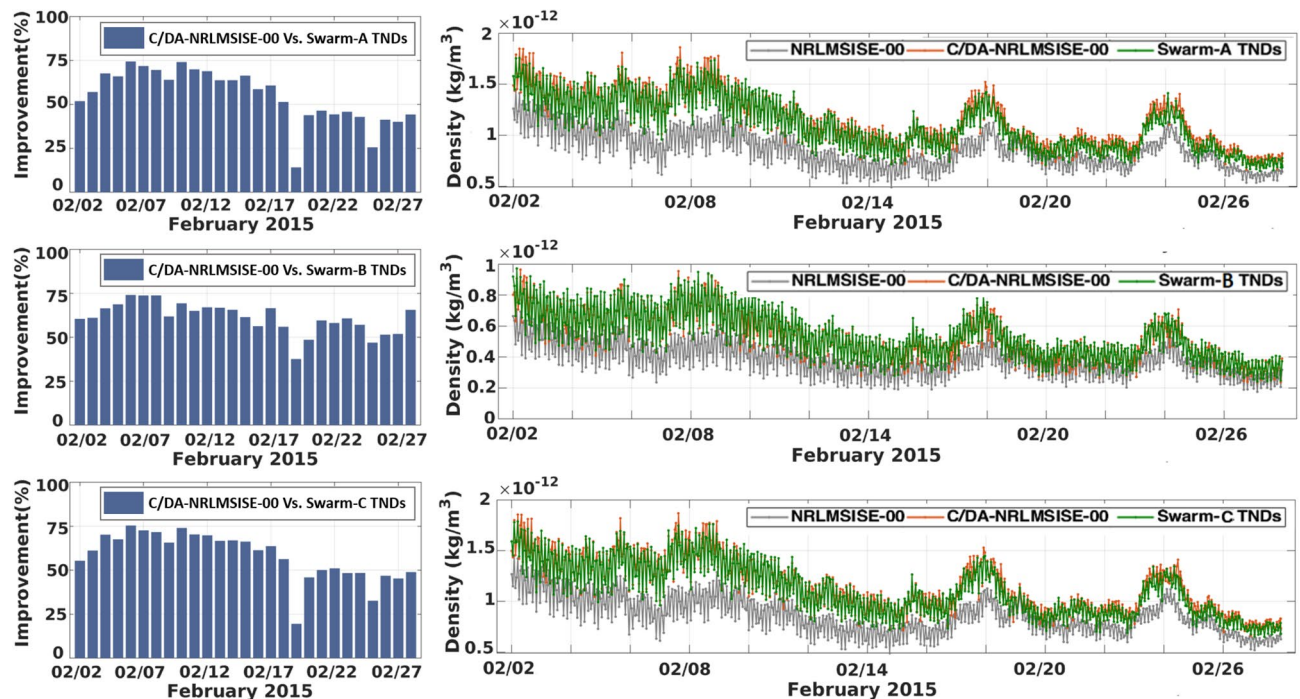


Figure 4. A comparison of the estimated improvements (Eq. (6)) and the time series of TNDs derived from the original NRLMSISE-00 and C/DA-NRLMSISE-00 models along with Swarm TNDs during February 2015. The C/DA is processed using GRACE-TNDs as observations, then the C/DA-NRLMSISE-00 densities are evaluated along the Swarm-A, -B and -C orbits. The plots were made in MATLAB (version R2021a, <https://www.mathworks.com/>).

Mission	Model					
	RMSE (kg/m ³)		NRMSE		AAPD (%)	
	NRLMSISE-00	C/DA-NRLMSISE-00	NRLMSISE-00	C/DA-NRLMSISE-00	NRLMSISE-00	C/DA-NRLMSISE-00
Swarm A	2.76×10^{-13}	1.16×10^{-13}	− 0.09	0.48	22.23	9.17
Swarm B	1.77×10^{-13}	7.05×10^{-14}	0.16	0.64	28.61	12.29
Swarm C	2.86×10^{-13}	1.14×10^{-13}	− 0.07	0.52	22.76	8.93

Table 1. A summary of the statistical measures derived during the analysis mode between the original NRLMSISE-00 and C/DA-NRLMSISE-00 models. Here GRACE-TNDs are used as observations within the C/DA procedure, and the results are evaluated with those of the Swarm mission during February 2015.

temperature measurements to the C/DA might further improve the results during these events. In this study, our assumption is that the density changes derived from accelerometer measurements are de-coupled from the wind component because this is the methodology that is applied for estimating accelerometer-derived TNDs (see Doornbos⁷⁰). Further investigations will be performed in future to apply C/DA on the coupled neutral mass density and wind models.

In Fig. 8, global maps of TND at different altitudes of 300, 350, 400, 450 and 500 km at the universal time 12:00 are shown. The maps shown from left to right correspond to those from NRLMSISE-00, C/DA-NRLMSISE-00, and the difference between them during February 8th, 2015 (a day with high solar activity, $F_{10.7} = 153$). The maps on the right columns indicate that C/DA changes the magnitude of TNDs, not only affect the simulation at the GRACE altitude of 410 km (30.7% on average) but also changes the TNDs of other altitudes especially around the longitude 0° associated with the local noon, (i.e., 300 km: 22.86%, 350 km: 26.47%, 400 km: 29.96%, 450 km: 33.54%, and 500 km: 37.11%). It should be considered that the C/DA procedure is implemented using GRACE-TNDs that belong to the medium altitude (e.g., 400–500 km), thus, it might be difficult to interpret the TNDs of much higher or lower altitudes, or where the composition of neutral mass density considerably changes. This is due to the fact that the temperature and density profiles (Fig. 7), as well as the global TND maps on different altitudes (Fig. 8) are computed based on the vertical assumption of NRLMSISE-00 and the C/DA techniques does not violate the physical concept that is implemented in the model.

Comparing the equatorial thermospheric neutral density anomalies. In this section, we explore the impact of C/DA on representing the characteristics of the equatorial thermosphere anomaly (ETA)⁷¹. Analo-

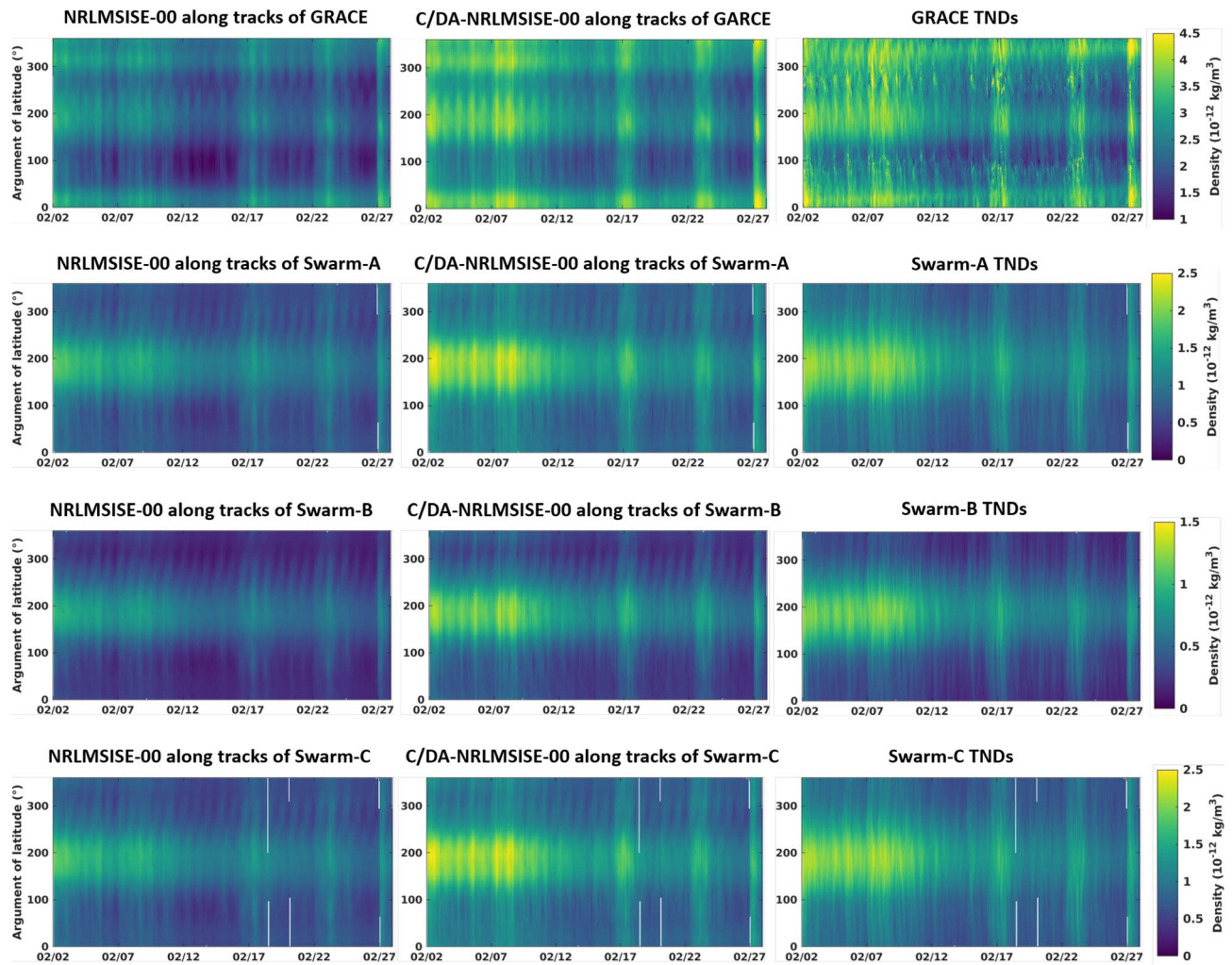


Figure 5. Maps of TND estimates retrieved from NRLMSISE-00 and C/DA-NRLMSISE-00 in the forecasting mode, as well as TND estimates derived along track of GRACE, Swarm-A, -B, and -C. The C/DA-NRLMSISE-00 model is implemented based on the description in the ‘Setup of the C/DA-NRLMSISE-00 Model’ section. The plots were generated using MATLAB (version R2021a, <https://www.mathworks.com/>).

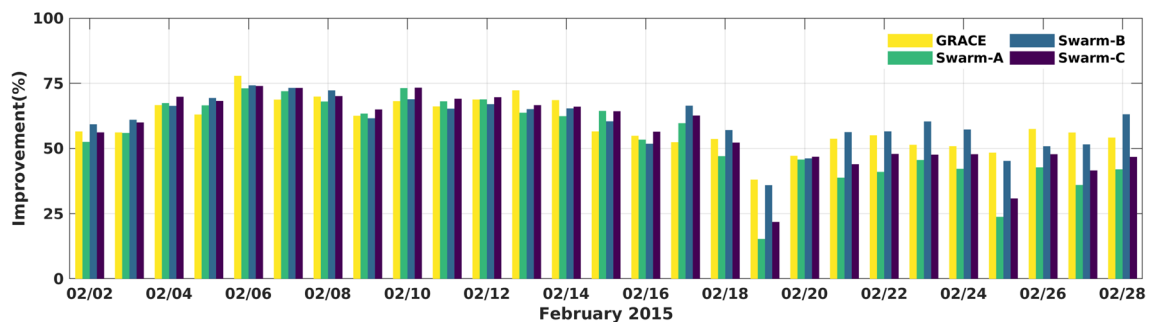


Figure 6. An overview of improvements in percent, which is used to assess the performance of the C/DA based on GRACE-TNDs as observation in the C/DA framework during February 2015. The figure was generated using MATLAB (version R2021a, <https://www.mathworks.com/>).

gous to the equatorial ionosphere anomaly (EIA)⁷², ETA is formed around the local noon. Various mechanisms are named as drivers of ETA namely heat transport due to zonal winds, chemical heating, and field-aligned ion drag⁷³. ETA is then represented as a function of local time, season, geomagnetic and solar activity index. To illustrate the effect of C/DA on representing ETA, the differences between are computed during different local times. The differences polar map for the altitude of 450 km are depicted in Fig. 9 on February 8th, 2015 with relatively high solar activity. The patterns of difference between TNDs from 11:00 to 19:00 LT indicate that the peak values

Mission	Model					
	RMSE(kg/m ³)		NRMSE		AAPD (%)	
	NRLMSISE-00	C/DA-NRLMSISE-00	NRLMSISE-00	C/DA-NRLMSISE-00	NRLMSISE-00	C/DA-NRLMSISE-00
GRACE	7.10×10^{-13}	2.79×10^{-13}	-0.54	0.39	22.11	7.36
Swarm A	2.68×10^{-13}	1.09×10^{-13}	-0.09	0.49	21.95	8.89
Swarm B	1.73×10^{-13}	6.60×10^{-14}	0.16	0.66	28.40	11.75
Swarm C	2.78×10^{-13}	1.07×10^{-13}	-0.06	0.54	22.48	8.62

Table 2. A summary of statistical measures between the original and C/DA-NRLMSISE-00 compared to the TNDs derived from GRACE and Swarm measurements during February 2015 in the forecasting mode.

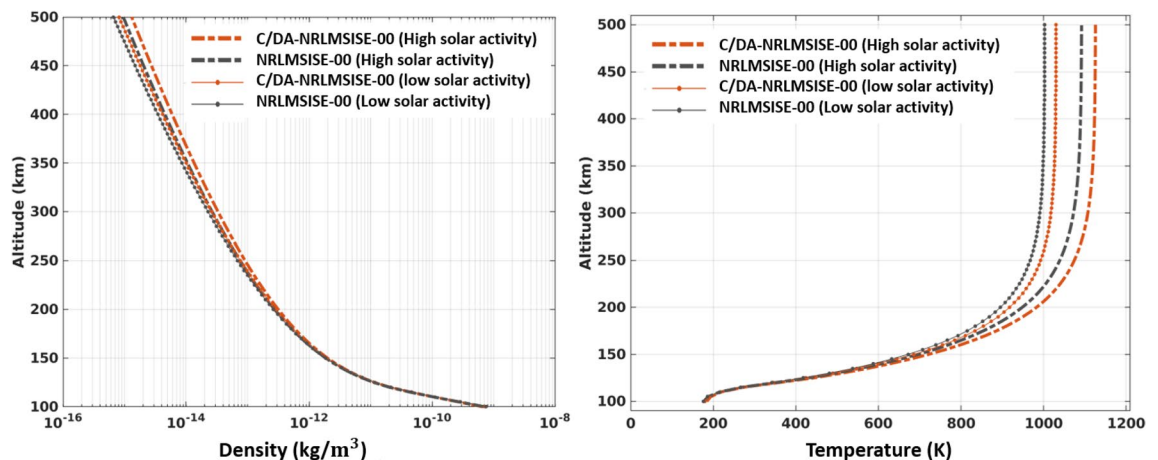


Figure 7. Temperature and density profiles during low solar activity (February 26th, 2015 at 12:00 UT, $F_{10.7} = 110$, gray) and high solar activity (February 8th, 2015 at 12:00 UT, $F_{10.7} = 153$, red) conditions. The solid and dashed lines are the corresponding NRLMSISE-00 and C/DA model, respectively. The vertical profile is estimated at latitude 0° and longitude 0° . The plots were generated using MATLAB (version R2021a, <https://www.mathworks.com/>).

locate around the ETA. In addition, the ETA features become more prominent at 15:00 LT due to its daytime phenomenon and tight coupling with the EIA, which disappear rapidly after sunset. These results indicate that the C/DA increases the magnitude of TND changes due to ETA up to about 20–30%. Therefore, we conclude that the under-estimation of TNDs from the original model can be eliminated by tuning it through the C/DA and using the accelerometer derived TND as observations.

Exploring the spatial and temporal impacts of C/DA through the principal component analysis (PCA). To investigate the dominant spatial and temporal changes caused by applying the C/DA method, the principal component analysis (PCA)^{34,74} method is applied on the hourly differences between NRLMSISE-00 and C/DA-NRLMSISE-00 during one day (i.e., February 8th, 2015) and entire February 2015. The mathematical details of PCA are described in the Supplementary file (Section S7 online). The global differences at the altitude of 400 km during February 2015 are presented in Fig. 10. The spatial patterns known as empirical orthogonal functions (EOFs that are anomaly maps in terms of density kg/m³) and their associated uncorrelated temporal patterns (principal components, or PCs that are unit-less) represent the orthogonal modes and are plotted beside each other.

The first dominant mode corresponds to 75.55% and 52.34% of the total variance of monthly and daily TND differences, respectively. The first mode of monthly differences in Fig. 10 (the global anomaly map and its temporal PC on top) indicates that the average difference in terms of TNDs is found to be 7.13×10^{-13} kg/m³ (in the range of $-1.06 \times 10^{-13} - 2.20 \times 10^{-12}$ kg/m³). The results also show that the magnitude of differences during February 1st–10th is higher than the rest. PC1 is found to be correlated with the solar activity index with the correlation coefficient of 50.04% (see Fig. 1) shows the standardized solar index along-side of PC1. This confirms the previous results of e.g., Forootan et al.³² who reported that empirical models do not reflect recent neutral mass density changes caused by solar activity. The first mode of daily TND differences Fig. 10 (the global anomaly map and its temporal PC on the bottom of figure) indicates that the magnitude of differences reach up to 10^{-12} kg/m³, and it is dominated by the diurnal frequency. The impact of these differences on the drag acceleration estimation is equivalent with 17.79, 33.42, 50.87 and 68.16% on a satellite at 300, 400, 500 and 600 km altitude thus it cannot be ignored in the precise orbit determination applications. It should be mentioned here that the atmospheric drag depends on the cross-sectional area perpendicular to the velocity vector of satellite (A), mass of the satellite (m),

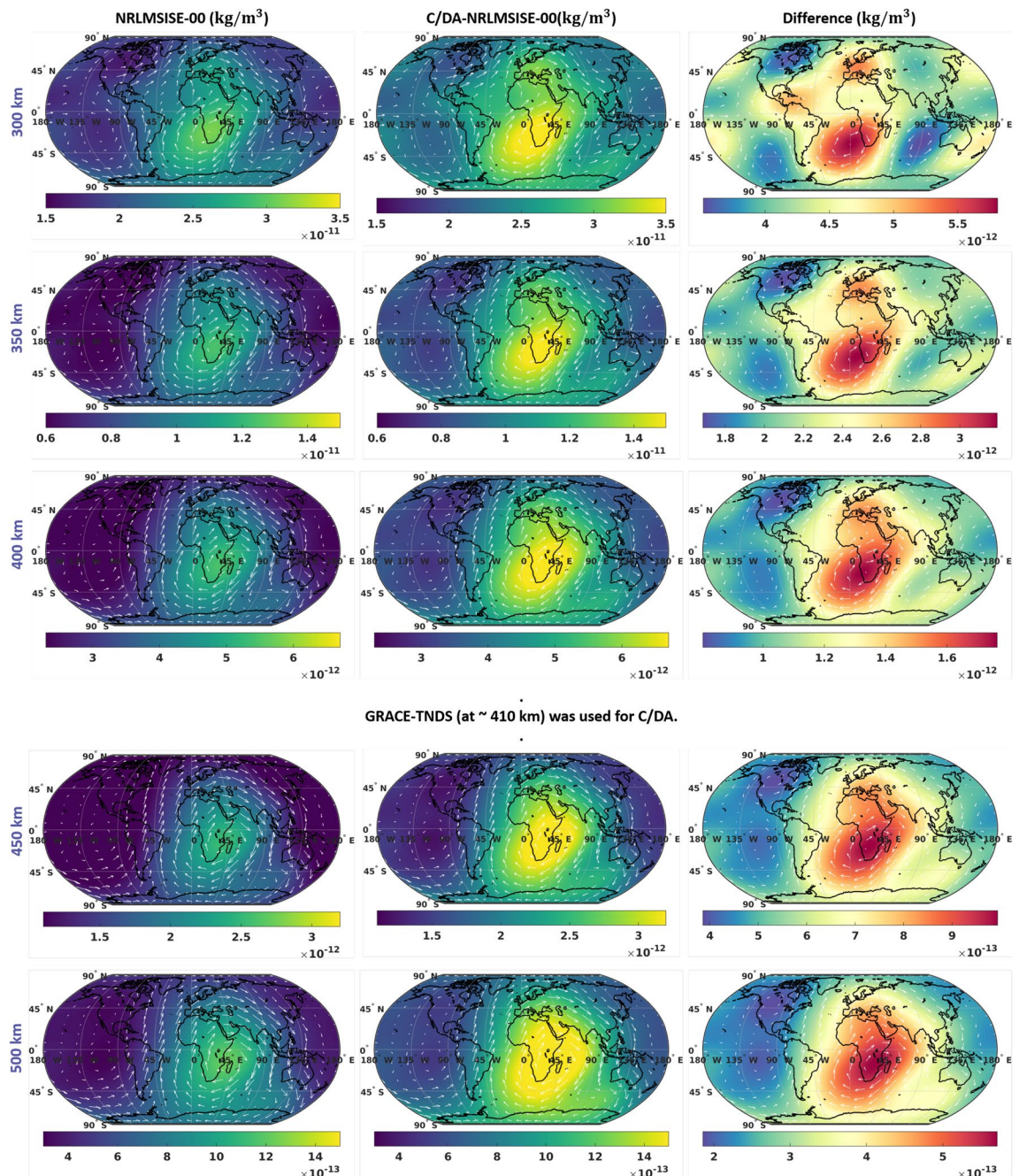


Figure 8. (left) Global maps of thermosphere neutral density (TND) from the original NRLMSISE-00, (middle) the forecasts of C/DA-NRLMSISE-00 with 1 h lead time, and (right) their difference at 12h UT on February 8th, 2015 for five different altitudes. The contour lines present the gradient vectors of density. These plots were generated using MATLAB (version R2021a, <https://www.mathworks.com/>).

drag coefficient (C_D), density (ρ), the norm of the relative velocity vector (v_{rel}^2), and the direction of the velocity relative to the co-rotating atmosphere (\hat{v}_{rel}). The values presented above are determined for spherical satellites ($a_{drag} = \frac{1}{2} \frac{A}{m} C_D \rho v_{rel}^2 \hat{v}_{rel}$) with the area of 1.00135 m², mass of 487.2 kg and velocity of 7600 m s⁻¹ on average.

Assessing the impact of updating TND estimates on forecasting ionospheric parameters. The main objective of this section is to demonstrate the impact of the thermospheric constituents from C/DA-NRLMSISE-00 on the forecasts of electron density. For this, the physics-based TIE-GCM model is chosen to simulate ionospheric variables. For this experiment the atomic oxygen (O), molecular oxygen (O₂), helium (He), and neutral temperature (TN) from the outputs of the C/DA-NRLMSISE-00 model are used as the primary history files of TIE-GCM. This is done by converting the species abundances from C/DA-NRLMSISE-00 to the mass mixing ratio γ with the estimation of conversion factor and this assumption $\gamma_{N_2} = 1 - (\gamma_O + \gamma_{O_2} + \gamma_{He})$ ⁵⁹, where γ_{N_2} , γ_O , γ_{O_2} , and γ_{He} represent mass mixing ratios of molecular nitrogen, atomic oxygen, molecular

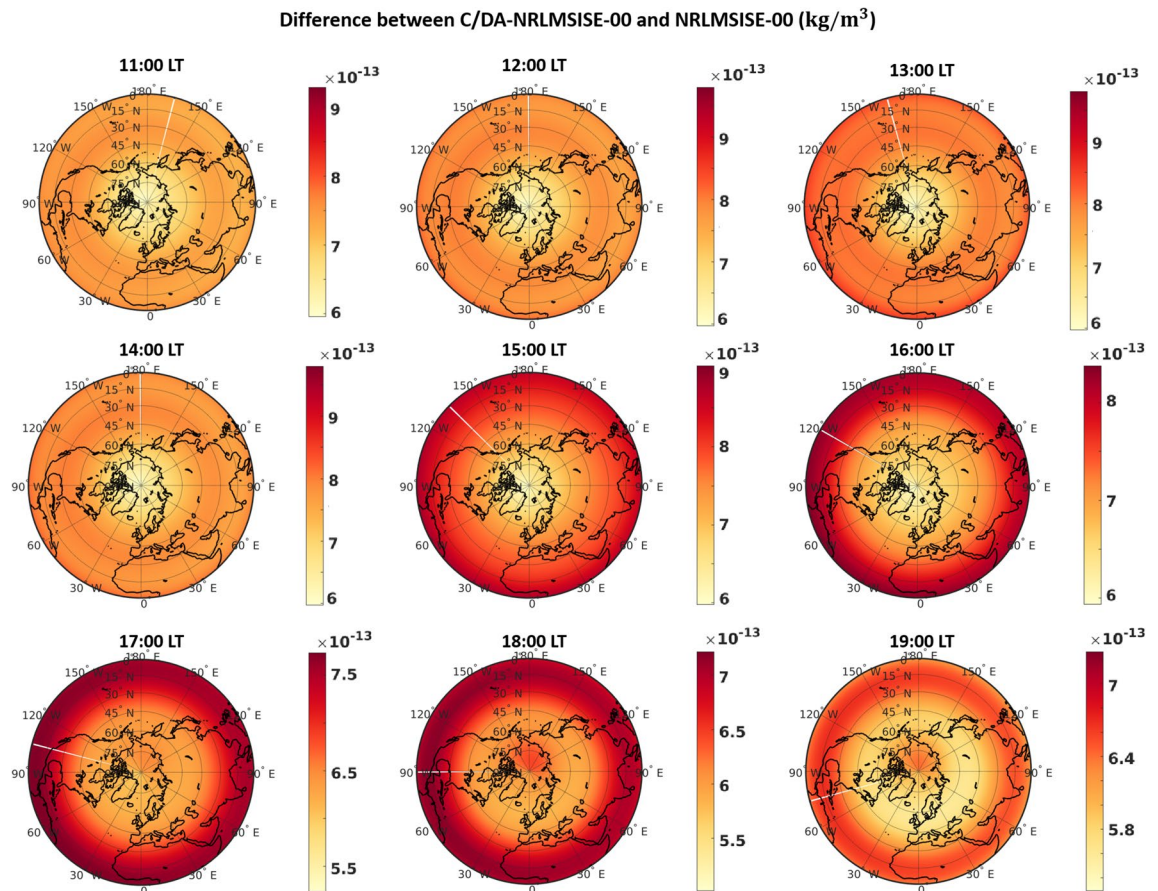


Figure 9. Thermosphere neutral density (TND) maps at the altitude of 450 km derived from the original NRLMSISE-00 (plots on left) and C/DA-NRLMSISE-00 (plots on middle) during 12 to 18 h local time (LT) on February 8th, 2015. The plots on the right correspond to the relative error (RE) maps that illustrate the impact of C/DA on representing the equatorial thermosphere anomaly (ETA). MATLAB (version R2021a, <https://www.mathworks.com/>) was used to generate these plots.

oxygen, and helium, respectively. The conversion factor C , which relates the number of densities to the mass mixing ratio, is assumed to be $C = \frac{P_{ti} \times m}{B \times T_n}$, where P_{ti} represents the pressure of each midpoint level i , and m is the mean molecular mass (g/mol) estimated as $\bar{m} = \frac{1}{\frac{\gamma_{O_2}}{32} + \frac{\gamma_O}{16} + \frac{\gamma_{He}}{4} + \frac{\gamma_{N_2}}{28}}$. B is the Boltzman constant $B = 1.3805 \times 10^{-16}$, and T_n is the neutral temperature (K). Finally, the number of density cm^{-3} from C/DA-NRLMSISE-00 ($\rho_{C/DA-NRLMSISE-00}$) is converted to the mass mixing ratio (mmr) as $\rho_{mmr} = \frac{w \times \rho_{C/DA-NRLMSISE-00}}{C}$, where w is the molecular weight of each species (i.e., $w_{O_2} = 32$, $w_O = 16$, $w_{He} = 4$, and $w_{N_2} = 28$).

To form TIE-GCM-I, which is the same as the original TIE-GCM with the improved primary history files based on the C/DA-NRLMSISE-00, and forecasting ionospheric variables the following steps are taken: (1) γ_{O_2} , γ_O , γ_{He} and γ_{N_2} and neutral temperature are estimated through the C/DA-NRLMSISE-00 in the forecasting mode; (2) the estimated parameters are converted to mmr and used to replace the default values in the primary history files; and (3) TIE-GCM-I is run for 1 h with recording the model output every 5 min. These three steps are repeated for 24 h of each day.

We chose February 8th, 2015 with high solar activity to test the results. The standalone TIE-GCM and TIE-GCM-I are compared with TEC from IGS products, and the RO-derived electron densities. Note that the model has an upper boundary of around 500–800 km altitude, while TEC estimates of IONEX represent $\sim 20,200$ km altitude. To reduce this inconsistency, the TEC estimates above the upper boundary of models are added using the simulation of the NeQuick ionosphere model⁷⁵.

In Fig. 11, we present nine vertical profiles of the electron density from models and RO during 24 h that correspond to high, mid, and low latitudes. To compare the results in terms of electron density, the Bi-linear and cubic spline interpolation scheme⁷⁶ was used to map the electron densities of TIE-GCM and TIE-GCM-I along the RO observations. Through the Comparisons with 300 COSMIC profiles during February 8th, 2015, we found the mean fitting coefficients (i.e., NRMSE based on the Eq. (8)) and the correlation coefficient to be 0.15 and 84.4% between TIE-GCM-I and RO profiles, respectively. These values show a better performance since those between TIE-GCM and RO was found to be -0.39 and 74.7% . Considering the results of the entire month, the TIE-GCM-I considerably reduces the RMSE between the electron density forecasts and RO observations within the range of $0.2 - 90.6\%$ (on average 30.92%).

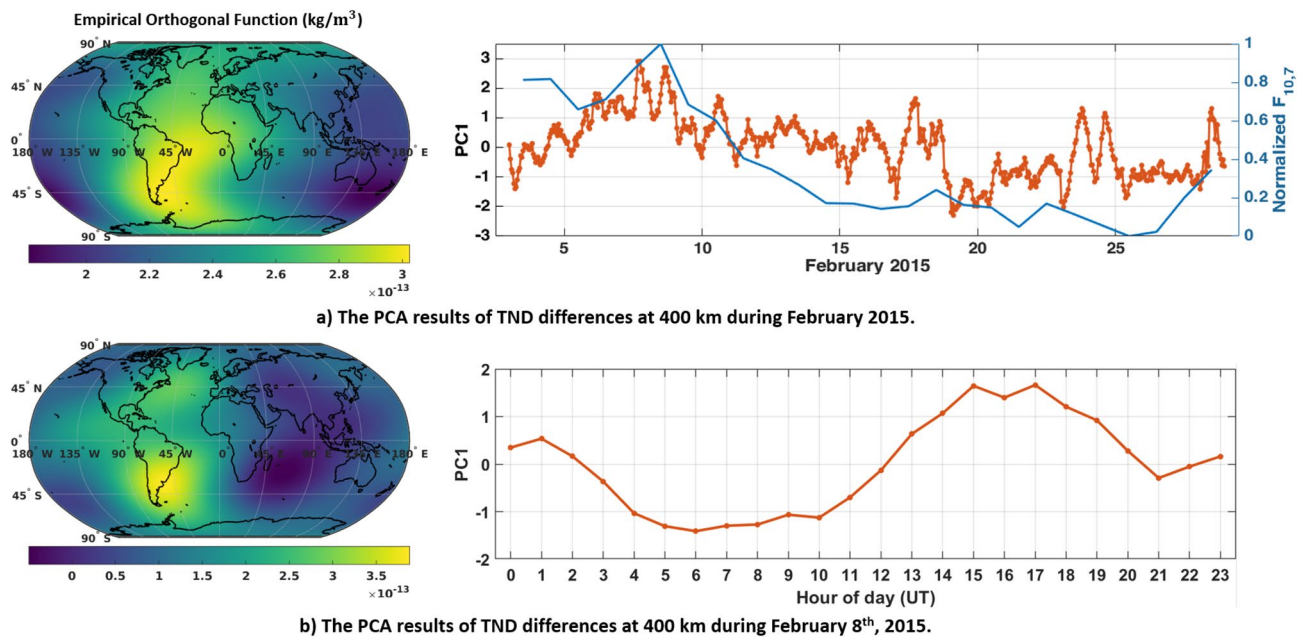


Figure 10. PCA of the TND differences between NRLMSISE-00 and C/DA-NRLMSISE-00 at 400 km during February 2015. The left plots are anomaly maps (EOFs) in terms of kg/m^3 , which can be multiplied by the unit less time series (PCs) on the right to derive orthogonal modes. The first mode of monthly differences (on top) represents 75.55% of the total variance of TND differences and the first mode of daily differences (on bottom) mode indicates 52.34% of the variance. These plots were generated using MATLAB (version R2021a, <https://www.mathworks.com/>).

The effect of TIE-GCM-I in forecasting of TEC is shown here as an example during February 8th, 2015 (see Fig. 12), where from left to right, we present: (1) four snapshots of TEC from IGS IONEX, (2) TIE-GCM, (3) TIE-GCM-I. Comparison between (2) and (3) indicates that the forecasts of TEC estimates in TIE-GCM-I agree better with those of IGS (i.e., RMSE of 13.4, 13.8, 14.6 and 14.4 TECU for (2), while 11.3, 10.6, 8.1 and 10.9 TECU for (3)). The average reduction in the errors of TEC forecasts was found to be 26.6%. We also observe that the dynamics of TEC changes are well captured by TIE-GCM-I no matter the default patterns are homogeneous, see e.g., Fig. 12(1) at 06:00 or containing local peaks, see e.g., Fig. 12(1) at 14:00. Besides, the capability of TIE-GCM-I to influence the equatorial ionosphere anomaly (EIA)⁷⁷ is enhanced by increasing the magnitude of TEC changes due to EIA about 22.5% on average around $\pm 30^\circ$ latitude at global constant local time at 12:00 LT. Note that the results for altitudes below 200 km must be interpreted with caution as the uncertainty of RO estimates of this altitude range may be considerably larger than those at higher altitudes. Besides, the C/DA results of GRACE TND are effective (at a range of 100 to 200 km) around the altitude of the mission.

TIE-GCM-I is run to predict TEC for the entire February 2015, for which the TEC estimates from TIE-GCM-I are found to be better fitted to the IONEX final TEC products than the original TIE-GCM, see Fig. 13. Our results indicate that TIE-GCM-I can decrease the monthly average RMSE from 14.6 to 9.8 TECU and the average improvement is found to be 31.01% (in the range of 15.8–44.3%). This is, in fact, a considerable improvement that is achieved by replacing the input thermosphere-related history files only, without assimilating ionosphere-related measurements.

Conclusions

In this study, thermospheric neutral density (TND) estimates from the Gravity Recovery and Climate Experiment (GRACE) are used to re-calibrate the NRLMSISE-00's parameters. For this, the simultaneous calibration and data assimilation (C/DA) technique^{46,47} based on the Ensemble Kalman Filter (EnKF)⁷⁸ is implemented. This new calibrated model, C/DA-NRLMSISE-00, is then used to forecast multi-level global TND maps as well as individual neutral mass density and thermospheric temperature for the next 1 h to 21 h. The TND estimates from the European Space Agency (ESA)'s Swarm¹⁴ mission are used for validating the TND results. To understand the impact of the thermospheric TNDs on forecasting ionospheric variables, the output of C/DA-NRLMSISE-00 are used to replace the history files of TIE-GCM, i.e., forming an improved model TIE-GCM-I. The electron density and vertical total electron content (VTEC, i.e., simply called TEC in the entire manuscript) of TIE-GCM and TIE-GCM-I are evaluated with those from RO and GNSS measurements. Therefore, by this evaluation, one might say that the C/DA of thermosphere-related variables is validated against ionosphere-related measurements.

The main findings of this study can be summarized as:

- C/DA-NRLMSISE-00 provides TND and individual mass density estimates with smaller uncertainty than the original NRLMSISE-00 model. The TND measurements of GRACE and Swarm in both validation and forecasting steps were found to be closer to the C/DA estimates. Besides, the calibrated parameters provide

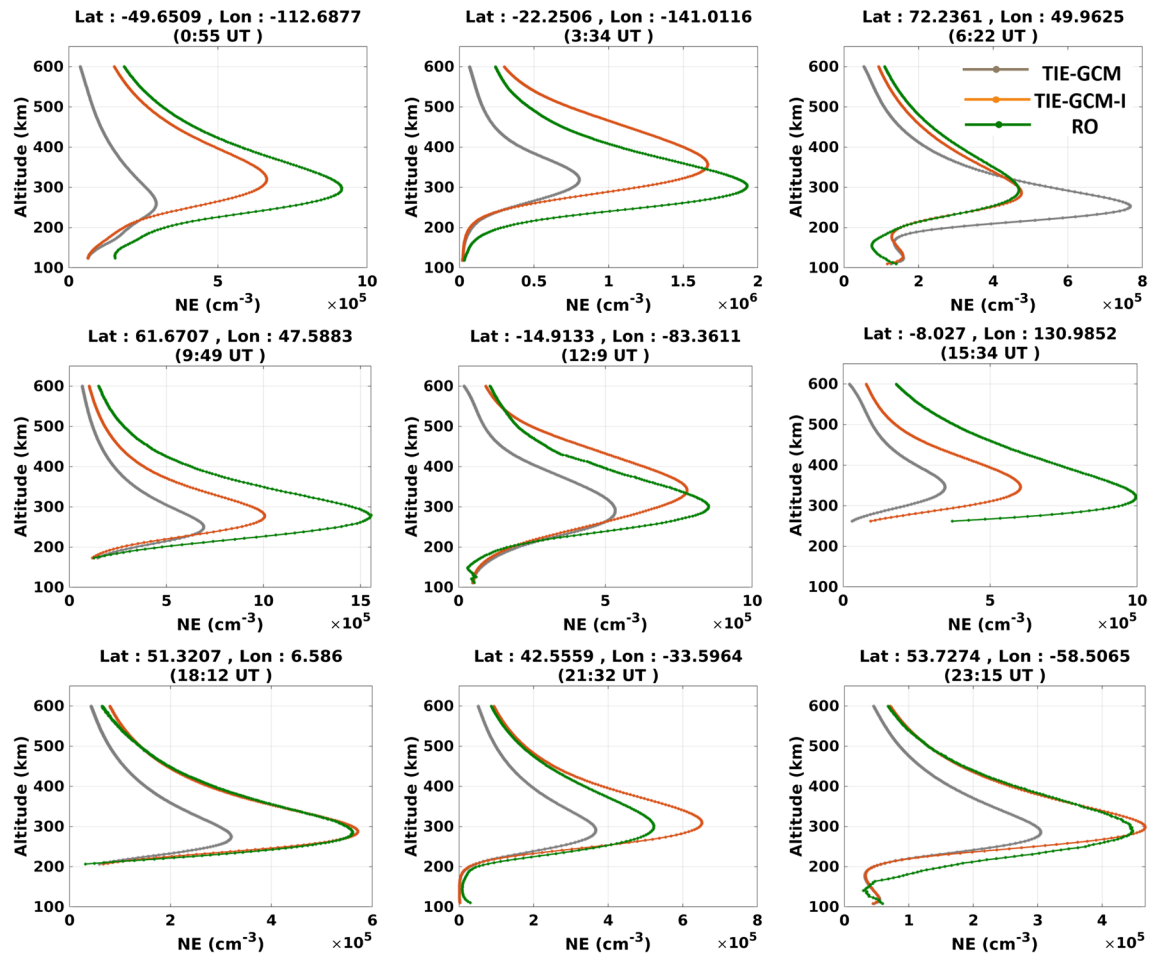


Figure 11. The altitude-dependent profiles of Ne retrieved from the COSMIC radio occultation data, as well as the original TIE-GCM, and TIEGCM-I during 24 h of February 8th, 2015. The plots were generated using MATLAB (version R2021a, <https://www.mathworks.com/>).

the opportunity to extend the updates globally and in different altitudes. From our assessments, we found considerable changes of TND estimates for the range of 300 to 500 km. In addition, C/DA-NRLMSISE-00 is found to be more effective than the original NRLMSISE-00 to represent the equatorial thermosphere anomaly (ETA). The differences between the original and C/DA-NRLMSISE-00 indicate that the maximum relative error is reached around ETA and the increase in its magnitude is found to be around 20–30%.

- Exploring the spatial and temporal differences between the TNDs of NRLMSISE-00 and C/DA-NRLMSISE-00 using the principal component analysis (PCA) technique indicates that the daily difference between these two models in TND simulations reaches up to 10^{-12} kg/m³ and the effect of this difference on the atmospheric drag is determined about 34% (on average) along typical GRACE-like satellite orbits at 400 km altitude.
- The impact of the updated mass mixing ratio (mmr) values, from C/DA-NRLMSISE-00, to be used as the primary history files of TIE-GCM (shown by TIE-GCM-I) provides a very positive impact on the global forecast of electron density and TEC. Experiments during February 8th, 2015 (i.e., a day with high solar activity) shows that the proposed method is able to decrease the errors by up to 26.48% compared to the TEC estimates from IGS product (IONEX)⁴⁹. We also compared TIE-GCM and TIE-GCM-I with RO measurements, which showed that the vertical profiles of the forecasts can be positively improved 30.92% on average.
- Previous studies, dealing with the data assimilation of upper atmosphere, already demonstrated the success EnKF-based data assimilation while using TIE-GCM as base model^{41,79,80}. In these studies, however, electron density was used as main observation to tune the model's skills for estimating ionosphere related parameters. In fact, the assimilation of electron observations requires the application of localization techniques that limits their application for large scale or global case studies. The proposed C/DA instead is able to eliminate this requirement because the global input primary history files can be replaced by those of C/DA-NRLMSISE-00.
- The electron density is expected to reach its maximum at the equator because of the energy received from the Sun. The equatorial ionization anomaly (EIA)⁷⁷ is a measure that represents this fact, and can be around ± 30 degree. Comparing the estimates of TEC from TIE-GCM and TIE-GCM-I for a typical day in February 2015 showed stronger anomalies after introducing the C/DA-NRLMSISE-00 input files. The differences were around 22.5% of the magnitude of EIA itself, which indicates that the underestimation of neutral density mass in the original model limits its ability to represent the latitudinal electron propagation.

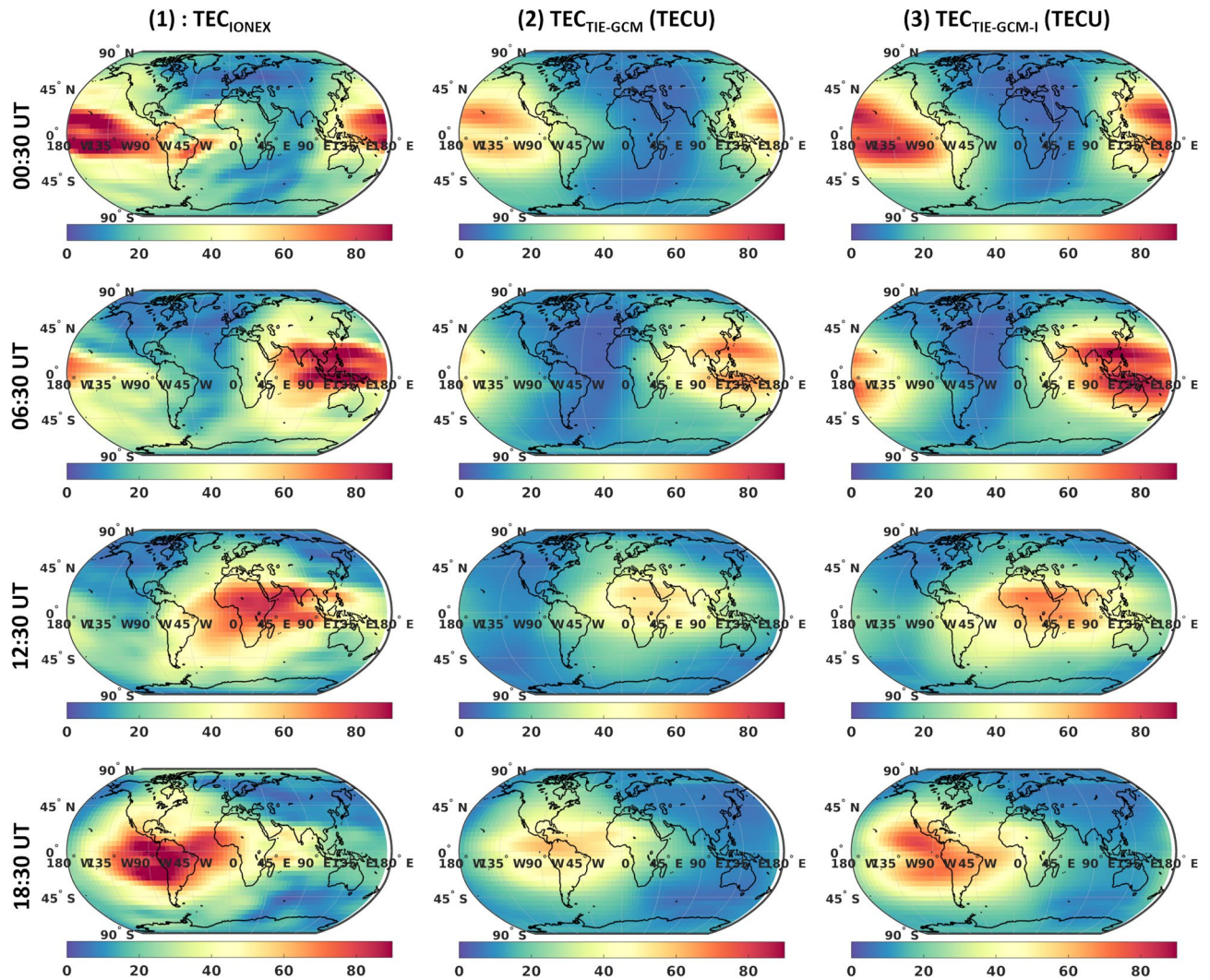


Figure 12. The 1st, 2nd, and 3rd columns present the TEC estimated derived from the IONEX final products, TIE-GCM, and TIE-GCM-I, during February 8th, 2015, respectively. MATLAB (version R2021a, <https://www.mathworks.com/>) was used to generate these plots.

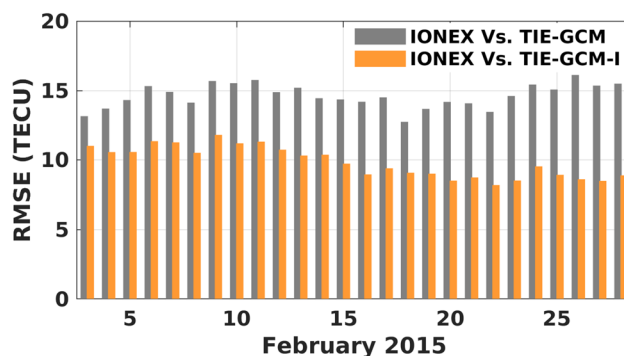


Figure 13. A comparison between the RMSE of TEC estimates during February 2015. Here, TEC estimates in the forecasting mode of the original TIE-GCM and TIE-GCM-I models are compared with the IONEX final TEC product from IGS. The plot was generated using MATLAB (version R2021a, <https://www.mathworks.com/>).

- Results of this study, in particular the demonstrated opportunity by implementing the complementary simultaneous data assimilation and model calibration framework, have important implications for maximizing the effective utilization of TND data in thermospheric models. Implementing the C/DA on an empirical model is computationally very efficient and fast (~ 111 s using 75 ensembles, 4 cores, and 10 GB RAM) thanks to the light-computational costs of NRLMSISE-00. This provides an opportunity to replace the initial values of coupled models such as TIE-GCM to reach global and multi-level thermosphere and ionosphere predictions. A similar 3-h data assimilation (using DART⁸¹) of TNDs that works directly for TIE-GCM takes (~ 4800 s for 75 ensembles, using 4 cores and 10 GB RAM). Besides, this latter implementation does not produce globally distributed and multi-level updated forecasts.

This work can be extended by performing multi-mission TND C/DA, as well as testing this technique on other types of empirical models. The advantage of the proposed method for near-real time orbit prediction and point positioning applications will be further explored.

Received: 16 November 2021; Accepted: 19 January 2022

Published online: 08 February 2022

References

- Berger, T. E., Holzinger, M. J., Sutton, E. K. & Thayer, J. P. Flying through uncertainty. *Space Weather* **18**, e2019SW002373. <https://doi.org/10.1029/2019SW002373> (2020).
- Yakovlev, M. The “IADC Space Debris Mitigation Guidelines” and supporting documents. In *4th European Conference on Space Debris*, vol. 587, 591–597 (2005).
- Farzaneh, S. & Forootan, E. Reconstructing regional ionospheric electron density: A combined spherical Slepian function and empirical orthogonal function approach. *Surv. Geophys.* **39**, 289–309. <https://doi.org/10.1007/s10712-017-9446-y> (2018).
- Gu, S. *et al.* Multi-GNSS PPP/INS tightly coupled integration with atmospheric augmentation and its application in urban vehicle navigation. *J. Geodesy* **95**, 1–15. <https://doi.org/10.1007/s00190-021-01514-8> (2021).
- Liu, Q., Hernández-Pajares, M., Lyu, H. & Goss, A. Influence of temporal resolution on the performance of global ionospheric maps. *J. Geodesy* **95**, 1–16. <https://doi.org/10.1007/s00190-021-01483-y> (2021).
- Salles, L. A., Vani, B. C., Moraes, A., Costa, E. & de Paula, E. R. Investigating ionospheric scintillation effects on multi-frequency GPS signals. *Surv. Geophys.* <https://doi.org/10.1007/s10712-021-09643-7> (2021).
- Ren, D. & Lei, J. A long-range forecasting model for the thermosphere based on the intelligent optimized particle filtering. *Sci. China Earth Sci.* <https://doi.org/10.1007/s11430-021-9847-9> (2021).
- Roble, R. G. & Ridley, E. C. A thermosphere-ionosphere-mesosphere-electrodynamics general circulation model TIME-GCM: Equinox solar cycle minimum simulations (30–500 km). *Geophys. Res. Lett.* **21**, 417–420. <https://doi.org/10.1029/93GL03391> (1994).
- Calabia, A. & Jin, S. New modes and mechanisms of thermospheric mass density variations from GRACE accelerometers. *J. Geophys. Res. Space Phys.* **121**, 11191–11212. <https://doi.org/10.1002/2016JA022594> (2016).
- Reigber, C., Lühr, H. & Schwintzer, P. CHAMP mission status. *Adv. Space Res.* **30**, 129–134. [https://doi.org/10.1016/S0273-1177\(02\)00276-4](https://doi.org/10.1016/S0273-1177(02)00276-4) (2002).
- Tapley, B. D., Bettadpur, S., Ries, J. C., Thompson, P. F. & Watkins, M. M. GRACE measurements of mass variability in the Earth system. *Science* **305**, 503–505. <https://doi.org/10.1126/science.1099192> (2004).
- Flechtner, F., Morton, P., Watkins, M. & Webb, F. Status of the GRACE Follow-On Mission. In *Gravity, Geoid and Height Systems. International Association of Geodesy Symposia*, Vol. 141 (eds Marti, U.) 117–121. https://doi.org/10.1007/978-3-319-10837-7_15 (Springer, Cham, 2014).
- Albertella, A., Migliaccio, F. & Sansó, F. GOCE: The earth gravity field by space gradiometry. *Celest. Mech. Dyn. Astron.* **83**, 1–15. <https://doi.org/10.1023/A:1020104624752> (2002).
- Visser, P., Doornbos, E., van den IJssel, J. & Da Encarnação, J. T. Thermospheric density and wind retrieval from Swarm observations. *Earth Planets Space* **65**, 1319–1331. <https://doi.org/10.5047/eps.2013.08.003> (2013).
- Delft Technical University. <ftp://thermosphere.tudelft.nl/>
- European Space Agency. <https://earth.esa.int>
- Mehta, P. M., Walker, A. C., Sutton, E. K. & Godinez, H. C. New density estimates derived using accelerometers on board the champ and grace satellites. *Space Weather* **15**, 558–576. <https://doi.org/10.1002/2016SW001562> (2017).
- Vielberg, K. *et al.* TND-IGG RL01: Thermospheric neutral density from accelerometer measurements of GRACE, CHAMP and Swarm. *PANGAEA* <https://doi.org/10.1594/PANGAEA.931347> (2021).
- Belehaki, A., Stanislawski, I. & Liliensten, J. An overview of ionosphere—Thermosphere models available for space weather purposes. *Space Sci. Rev.* **147**, 271–313. <https://doi.org/10.1007/s11214-009-9510-0> (2009).
- Roble, R. G. & Ridley, E. C. A thermosphere-ionosphere-mesosphere-electrodynamics general circulation model (TIME-GCM): Equinox solar cycle minimum simulations (30–500 km). *Geophys. Res. Lett.* **21**, 417–420. <https://doi.org/10.1029/93GL03391> (1994).
- Fuller-Rowell, T. J. & Rees, D. A three-dimensional time-dependent global model of the thermosphere. *J. Atmos. Sci.* **37**, 2545–2567. [https://doi.org/10.1175/1520-0469\(1980\)037<2545:ATDTDG>2.0.CO;2](https://doi.org/10.1175/1520-0469(1980)037<2545:ATDTDG>2.0.CO;2) (1980).
- Picone, J., Hedin, A., Drob, D. P. & Aikin, A. NRLMSISE-00 empirical model of the atmosphere: Statistical comparisons and scientific issues. *J. Geophys. Res. Space Phys.* **107**, 1468. <https://doi.org/10.1029/2002JA009430> (2002).
- Bowman, B. R., Tobiska, W. K., Marcos, F. A. & Valladares, C. The JB2006 empirical thermospheric density model. *J. Atmos. Solar-Terr. Phys.* **70**, 774–793. <https://doi.org/10.1016/j.jastp.2007.10.002> (2008).
- Bruinsma, Sean. The DTM-2013 thermosphere model. *J. Space Weather Space Clim.* **5**, A1. <https://doi.org/10.1051/swsc/2015001> (2015).
- Emmert, J. Thermospheric mass density: A review. *Adv. Space Res.* **56**, 773–824. <https://doi.org/10.1016/j.asr.2015.05.038> (2015).
- Palmroth, M. *et al.* Lower-thermosphere-ionosphere (LTI) quantities: Current status of measuring techniques and models. *Annales Geophysicae* **39**, 189–237. <https://doi.org/10.5194/angeo-39-189-2021> (2021).
- Doornbos, E., Klinkrad, H. & Visser, P. Atmospheric density calibration using satellite drag observations. *Adv. Space Res.* **36**, 515–521. <https://doi.org/10.1016/j.asr.2005.02.009> (2005).
- Doornbos, E., Klinkrad, H. & Visser, P. Use of two-line element data for thermosphere neutral density model calibration. *Adv. Space Res.* **41**, 1115–1122. <https://doi.org/10.1016/j.asr.2006.12.025> (2008).
- Pérez, D., Wohlberg, B., Alan Lovell, T., Shoemaker, M. & Bevilacqua, R. Orbit-centered atmospheric density prediction using artificial neural networks. *Acta Astronautica* **98**, 9–23. <https://doi.org/10.1016/j.actaastro.2014.01.007> (2014).

30. Shi, C., Li, W., Li, M., Zhao, Q. & Sang, J. Calibrating the scale of the NRLMSISE00 model during solar maximum using the two line elements dataset. *Adv. Space Res.* **56**, 1–9. <https://doi.org/10.1016/j.asr.2015.03.024> (2015).
31. Ruan, H., Lei, J., Dou, X., Liu, S. & Aa, E. An exospheric temperature model based on champ observations and TIEGCM simulations. *Space Weather* **16**, 147–156. <https://doi.org/10.1002/2017SW001759> (2018).
32. Forootan, E., Farzaneh, S., Lück, C. & Vielberg, K. Estimating and predicting corrections for empirical thermospheric models. *Geophys. J. Int.* **218**, 479–493. <https://doi.org/10.1093/gji/ggz163> (2019).
33. Weng, L., Lei, J., Sutton, E., Dou, X. & Fang, H. An exospheric temperature model from CHAMP thermospheric density. *Space Weather* **15**, 343–351. <https://doi.org/10.1002/2016SW001577> (2017).
34. Forootan, E. *Statistical Signal Decomposition Techniques for Analyzing Time-Variable Satellite Gravimetry Data*. Ph.D. thesis, University of Bonn (2014).
35. Kalman, R. E. A new approach to linear filtering and prediction problems. *J. Basic Eng.* **82**, 35–45. <https://doi.org/10.1115/1.3662552> (1960).
36. Matsuo, T., Richmond, A. D. & Nychka, D. W. Modes of high-latitude electric field variability derived from DE-2 measurements: Empirical Orthogonal Function (EOF) analysis. *Geophys. Res. Lett.* <https://doi.org/10.1029/2001GL014077> (2002).
37. Matsuo, T., Richmond, A. D. & Lu, G. Optimal interpolation analysis of high-latitude ionospheric electrodynamics using empirical orthogonal functions: estimation of dominant modes of variability and temporal scales of large-scale electric fields. *J. Geophys. Res. Space Phys.* **110**, A06301. <https://doi.org/10.1029/2004JA010531> (2005).
38. Matsuo, T. & Forbes, J. M. Principal modes of thermospheric density variability: Empirical orthogonal function analysis of CHAMP 2001–2008 data. *J. Geophys. Res. Space Phys.* **115**, A07309. <https://doi.org/10.1029/2009JA015109> (2010).
39. Mehta, P. M., Linares, R. & Sutton, E. K. A quasi-physical dynamic reduced order model for thermospheric mass density via Hermitian space-dynamic mode decomposition. *Space Weather* **16**, 569–588. <https://doi.org/10.1029/2018SW001840> (2018).
40. Mehta, P. M., Linares, R. & Sutton, E. K. Data-driven inference of thermosphere composition during solar minimum conditions. *Space Weather* **17**, 1364–1379. <https://doi.org/10.1029/2019SW002264> (2019).
41. Matsuo, T. Upper atmosphere data assimilation with an ensemble Kalman filter. In *Modeling the Ionosphere–Thermosphere System* (eds Huba, J., Schunk, R. & Khazanov, G.) 273–282. <https://doi.org/10.1002/9781118704417.ch22> (American Geophysical Union (AGU), 2014).
42. Godinez, H. C. *et al.* Specification of the ionosphere-thermosphere using the ensemble Kalman filter. In *International Conference on Dynamic Data-Driven Environmental Systems Science* 274–283. https://doi.org/10.1007/978-3-319-25138-7_25 (Springer, 2014).
43. Morozov, A. V. *et al.* Data assimilation and driver estimation for the Global Ionosphere–Thermosphere Model using the ensemble adjustment Kalman filter. *J. Atmos. Solar-Terr. Phys.* **104**, 126–136. <https://doi.org/10.1016/j.jastp.2013.08.016> (2013).
44. Pilinski, M. D., Crowley, G., Sutton, E. & Codrescu, M. Improved orbit determination and forecasts with an assimilative tool for satellite drag specification. In *Advanced Maui Optical and Space Surveillance Technologies Conference*, vol. 104 (2016).
45. Codrescu, S., Codrescu, M. & Fedrizzi, M. An ensemble Kalman filter for the thermosphere–ionosphere. *Space Weather* **16**, 57–68. <https://doi.org/10.1002/2017SW001752> (2018).
46. Schumacher, M. *Methods for Assimilating Remotely-Sensed Water Storage Changes into Hydrological Models*. Ph.D. thesis, University of Bonn (2016).
47. Forootan, E., Farzaneh, S., Kosary, M., Schmidt, M. & Schumacher, M. A simultaneous calibration and data assimilation (C/DA) to improve NRLMSISE00 using thermospheric neutral density (TND) from space-borne accelerometer measurements. *Geophys. J. Int.* **224**, 1096–1115. <https://doi.org/10.1093/gji/ggaa507> (2020).
48. Constellation Observing System for Meteorology, Ionosphere and Climate. <http://www.cosmic.ucar.edu>
49. Schaer, S., Gurtner, W. & Feltens, J. Ionex: The ionosphere map exchange format version 1. In *Proceedings of the IGS AC Workshop, Darmstadt, Germany*, vol. 9 (1998).
50. Lei, J. *et al.* Comparison of COSMIC ionospheric measurements with ground-based observations and model predictions: Preliminary results. *J. Geophys. Res. Space Phys.* <https://doi.org/10.1029/2006JA012240> (2007).
51. Yang, K.-F., Chu, Y.-H., Su, C.-L., Ko, H.-T. & Wang, C.-Y. An examination of FORMOSAT-3/COSMIC ionospheric electron density profile: Data quality criteria and comparisons with the IRI model. *Terr. Atmos. Ocean. Sci.* **20**, 193. [https://doi.org/10.3319/TAO.2007.10.05.01\(F3C\)](https://doi.org/10.3319/TAO.2007.10.05.01(F3C)) (2009).
52. Hernández-Pajares, M. *et al.* The IGS VTEC maps: A reliable source of ionospheric information since 1998. *J. Geodesy* **83**, 263–275. <https://doi.org/10.1007/s00190-008-0266-1> (2009).
53. The Crustal Dynamics Data Information System (CDDIS). <ftp://cddis.gsfc.nasa.gov/pub/gps/products/ionex/>
54. Kodikara, T. *Physical Understanding and Forecasting of the Thermospheric Structure and Dynamics*. Ph.D. thesis, RMIT University (2019).
55. Tapping, K. F. The 10.7cm solar radio flux (F10.7). *Space Weather* **11**, 394–406. <https://doi.org/10.1002/swe.20064> (2013).
56. Ahluwalia, H. S. Ap time variations and interplanetary magnetic field intensity. *J. Geophys. Res. Space Phys.* **105**, 27481–27487. <https://doi.org/10.1029/2000JA900124> (2000).
57. Saltelli, A. Sensitivity analysis for importance assessment. *Risk Anal.* **22**, 579–590. <https://doi.org/10.1111/0272-4332.00040> (2002).
58. Maute, A. Thermosphere-ionosphere-electrodynamics general circulation model for the ionospheric connection explorer: TIEGCM-ICON. *Space Sci. Rev.* **212**, 523–551. <https://doi.org/10.1007/s11214-017-0330-3> (2017).
59. Qian, L. *et al.* The NCAR TIE-GCM: A community model of the coupled thermosphere/ionosphere system. *Model. Ionos. Thermosphere Syst.* **201**, 73–83. <https://doi.org/10.1002/9781118704417.ch7> (2014).
60. Richards, P., Fennelly, J. & Torr, D. EUVAC: A solar EUV flux model for aeronomic calculations. *J. Geophys. Res. Space Phys.* **99**, 8981–8992. <https://doi.org/10.1029/94JA00518> (1994).
61. Solomon, S. C. & Qian, L. Solar extreme-ultraviolet irradiance for general circulation models. *J. Geophys. Res. Space Phys.* <https://doi.org/10.1029/2005JA011160> (2005).
62. Webb, D. F. & Howard, R. A. The solar cycle variation of coronal mass ejections and the solar wind mass flux. *J. Geophys. Res. Space Phys.* **99**, 4201–4220. <https://doi.org/10.1029/93JA02742> (1994).
63. Heelis, R., Lowell, J. K. & Spiro, R. W. A model of the high-latitude ionospheric convection pattern. *J. Geophys. Res. Space Phys.* **87**, 6339–6345. <https://doi.org/10.1029/JA087iA08p06339> (1982).
64. Hagan, M. E., Roble, R. G. & Hackney, J. Migrating thermospheric tides. *J. Geophys. Res. Space Phys.* **106**, 12739–12752. <https://doi.org/10.1029/2000JA000344> (2001).
65. Evensen, G. The ensemble Kalman filter for combined state and parameter estimation. *IEEE Control Syst. Mag.* **29**, 83–104. <https://doi.org/10.1109/MCS.2009.932223> (2009).
66. Rideout, W. & Coster, A. Automated GPS processing for global total electron content data. *GPS Solut.* **10**, 219–228. <https://doi.org/10.1007/s10291-006-0029-5> (2006).
67. Liou, Y.-A. *et al.* FORMOSAT-3/COSMIC GPS radio occultation mission: Preliminary results. *IEEE Trans. Geosci. Remote Sens.* **45**, 3813–3826. <https://doi.org/10.1109/TGRS.2007.903365> (2007).
68. Lathuillière, C. & Menvielle, M. Comparison of the observed and modeled low- to mid-latitude thermosphere response to magnetic activity: Effects of solar cycle and disturbance time delay. *Adv. Space Res.* **45**, 1093–1100. <https://doi.org/10.1016/j.asr.2009.08.016> (2010) (Special Issue: Recent Advances in Space Weather Monitoring, Modelling, and Forecasting).
69. Lei, J., Thayer, J. P., Burns, A. G., Lu, G. & Deng, Y. Wind and temperature effects on thermosphere mass density response to the November 2004 geomagnetic storm. *J. Geophys. Res. Space Phys.* <https://doi.org/10.1029/2009JA014754> (2010).

70. Doornbos, E. *Thermospheric Density and Wind Determination from Satellite Dynamics* (Springer, 2012).
71. Hsu, V. W., Thayer, J. P., Lei, J. & Wang, W. Formation of the equatorial thermosphere anomaly trough: Local time and solar cycle variations. *J. Geophys. Res. Space Phys.* **119**, 10456–10473. <https://doi.org/10.1002/2014JA020416> (2014).
72. MacDougall, J. W. The equatorial ionospheric anomaly and the equatorial electrojet. *Radio Sci.* **4**, 805–810. <https://doi.org/10.1029/RS004i009p00805> (1969).
73. Lei, J. *et al.* Simulations of the equatorial thermosphere anomaly: Physical mechanisms for crest formation. *J. Geophys. Res. Space Phys.* <https://doi.org/10.1029/2012JA017613> (2012).
74. Preisendorfer, R. *Principal Component Analysis in Meteorology and Oceanography* (Elsevier, 1988).
75. Nava, B., Coisson, P. & Radicella, S. A new version of the NeQuick ionosphere electron density model. *J. Atmos. Solar-Terr. Phys.* **70**, 1856–1862. <https://doi.org/10.1016/j.jastp.2008.01.015> (2008).
76. De Boor, C. & De Boor, C. A *Practical Guide to Splines*, vol. 27 (Springer, 1978).
77. Appleton, E. V. Two anomalies in the ionosphere. *Nature* **157**, 691. <https://doi.org/10.1038/157691a0> (1946).
78. Evensen, G. The ensemble Kalman filter: Theoretical formulation and practical implementation. *Ocean Dyn.* **53**, 343–367. <https://doi.org/10.1007/s10236-003-0036-9> (2003).
79. Kodikara, T., Zhang, K., Pedatella, N. M. & Borries, C. The impact of solar activity on forecasting the upper atmosphere via assimilation of electron density data. *Space Weather* **19**, e2020SW002660. <https://doi.org/10.1029/2020SW002660> (2021).
80. Hsu, C.-T., Matsuo, T., Wang, W. & Liu, J.-Y. Effects of inferring unobserved thermospheric and ionospheric state variables by using an ensemble Kalman filter on global ionospheric specification and forecasting. *J. Geophys. Res. Space Phys.* **119**, 9256–9267. <https://doi.org/10.1002/2014JA020390> (2014).
81. Anderson, J. *et al.* The data assimilation research testbed: A community facility. *Bull. Am. Meteorol. Soc.* **90**, 1283–1296. <https://doi.org/10.1175/2009BAMS2618.1> (2009).

Acknowledgements

The authors are grateful for the comments provided by Professor Xueliang Guo and two anonymous reviewers, which helped us to improve the quality of this study. We acknowledge the use of GRACE <https://doi.pangaea.de/10.1594/PANGAEA.931347>, Swarm <ftp://thermosphere.tudelft.nl/>, TEC estimates from IGS product <https://cddis.nasa.gov/>, and RO data <http://www.cosmic.ucar.edu>. The source code for the simulation models used in this study, the NRLMSISE-00 and TIE-GCM, are freely available at <https://rbm9.gfz-potsdam.de/> and <https://www.hao.ucar.edu/modelling/tgcm/>, respectively. E. Forootan and M. Schumacher would like to thank the Department of Planning, Aalborg University for the financial supports and the German Research Foundation (DFG) for the Grant Number SCHM2433/13-1 (Project TIPOD: Development of High-precision Thermosphere Models for Improving POD of Low-Earth-Orbiting Satellites).

Author contributions

Conceptualization, E.F. and M.S.; methodology, E.F., S.F. and M.S.; software, M.K.; validation, M.K. and S.F.; formal analysis, E.F., M.K. and S.F.; investigation, T.K. and I.F.-G.; resources, E.F. and S.F.; data curation, K.V. and M.K.; writing—original draft preparation, E.F. and M.K.; writing—review and editing, E.F., T.K., C.B. and S.F.; visualization, M.K. and E.F.; supervision, E.F., S.F. and M.S.; project administration, E.F.; funding acquisition, E.F. All authors have read and agreed to the published version of the manuscript. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-05952-y>.

Correspondence and requests for materials should be addressed to E.F.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022