Evaluation of Physics-Based Data Assimilation System Driven by Neutral Density Data From a Single Satellite

Dexin Ren1,2,3 and Jiuhou Lei1,2,3

1CAS Key Laboratory of Geospace Environment, School of Earth and Space Sciences, University of Science and Technology of China, Hefei, China, 2Mengcheng National Geophysical Observatory, University of Science and Technology of China, Hefei, China, 3CAS Center for Excellence in Comparative Planetology, University of Science and Technology of China, Hefei, China

Abstract Accurate forecast of the thermospheric density is critical to the space community. The data assimilation approach that is based on the self-consistent upper-atmosphere model may provide better predictive capability of the coupled thermosphere system. In this study, a physics-based assimilation system (hereafter referred to as PIDA) that is based on the Thermosphere-Ionosphere-Electrodynamics General Circulation Model was used to validate the capability of reproducing the evolution of the global thermosphere state. The effective solar and geophysical drivers were estimated by ingesting neutral density from a single satellite into the PIDA. It was found that the PIDA can reproduce the temporal variation of the global thermospheric density at around the altitude where the orbit density was ingested. Furthermore, the PIDA is also capable of capturing the temporal evolution of the thermospheric density at various altitudes. However, a systematic bias, depending on altitude, is seen in the modeled neutral density of the PIDA. Moreover, this systematic bias in the thermospheric density is likely ascribed to the overestimation of the density in the lower thermosphere. Consequently, the spatial and temporal evolutions of the lower thermosphere under various conditions should be considered carefully in the physics-based data assimilation system. Additionally, the assessments of the obtained results suggested that the observations of multiple parameters at different altitudes are required to be assimilated into the thermospheric model.

1. Introduction

The time-dependent evolution of the neutral thermosphere is critical to the orbit prediction of the low Earth orbit (LEO) satellite. In recent years, accurate specification and forecast of the thermospheric state has become increasingly important in the space physics community (Berger et al., 2020). Great scientific efforts were made to improve the accuracy of data assimilation approach for the thermospheric parameters, such as thermospheric temperature (e.g., Cantrall et al., 2019; Ruan et al., 2018), neutral density (e.g., Sutton, 2018; Weimer et al., 2020), neutral compositions (Codrescu et al., 2004; Mehta et al., 2019; Mlynczak et al., 2018), and wind patterns (Cierpik et al., 2003; Lomidze & Scherliess, 2015). The proper orthogonal decomposition and empirical orthogonal function approaches were usually used to dynamically adjust the thermospheric temperature or density (Matsuo et al., 2012; Matsuo & Forbes, 2010; Mehta & Linares, 2017; Ruan et al., 2018). Picone et al. (2005) and Gondelach and Linares (2020) proposed the two-line element data assimilation to estimate thermospheric density. Various ensemble data assimilation frameworks, such as ensemble Kalman filter (Codrescu et al., 2018; Matsuo et al., 2013), ensemble optimal interpolation (Murray et al., 2015), and ensemble square root filter (Cantrall et al., 2019), were also conducted to assimilate the thermospheric observations into the empirical or theoretical thermosphere model. Additionally, Choury et al. (2013) applied neural networks to predict exosphere temperature. These methods were successful in achieving better agreement of the considered parameters between the measurements and model of the parameters considered.

It should be noted that the spatial and temporal evolutions of the thermosphere as a nonlinearly coupled complex system depend closely upon various parameters, such as temperature, wind, density, and neutral species (Shim et al., 2014). However, the thermospheric assimilation system that has the capability of...
modifying the various thermospheric parameters simultaneously is yet to be achieved. Since the physics-based general circulation models have been the most powerful medium in understanding the complex three-dimensional (3-D) time-dependent behavior of terrestrial atmosphere in different spatial and temporal scales (e.g., Emmert, 2015; Fesen et al., 2002; Huba et al., 2014; Qian & Solomon, 2011, and references therein), these physics-based general circulation models have been previously used in the assimilation system in order to provide a better predictive/forecast capability of the thermosphere (e.g., Mehta & Linares, 2017; Sutton, 2018). Recently, Sutton (2018) developed a new physics-based assimilation system (Iterative Reinitialization, Driver Estimation, and Assimilation [IRIDEA]) in which the LEO density was assimilated into the Thermosphere-Ionosphere-Electrodynamics General Circulation Model (TIEGCM). In this approach, a time history of effective solar and geophysical drivers was estimated by iteratively ingesting the neutral mass density from the Challenging Mini-Satellite Payload (CHAMP) accelerometer. The preliminary assimilation results showed that IRIDEA had good capabilities of reproducing the observed density from CHAMP and also from Gravity Recovery And Climate Experiment (GRACE) accelerometers. Further, Sutton (2018) suggested that IRIDEA may offer a great forecasting potential of the thermosphere. However, in Sutton (2018), the CHAMP satellite in 2003 sampled similar local times with those of the GRACE satellite. Moreover, besides the neutral density given from the IRIDEA system, other parameters, such as temperature and neutral species, were not analyzed in Sutton (2018). Note that the temporal evolution of the thermospheric state during the predicted period heavily depends on the accurate estimation of various parameters in the thermosphere at the current time step. That is, a data assimilation system, which can accurately reproduce the coupled thermospheric state, is critical to the thermospheric prediction. Therefore, a further evaluation of the physics-based assimilation system is desired to validate its capability of simulating the coupled thermosphere system.

This is a follow-up study of Sutton (2018). Hereafter, the assimilation system used in this study is referred to as PIDA (Physics-based Iterative Data Assimilation system). It should be noted that the local time of CHAMP and GRACE orbits in 2004 have about 4–6 hr difference, and as a result the CHAMP orbit density from day 80 to day 280 in 2004 is used to drive the PIDA. The corresponding GRACE orbit density is applied to evaluate the capability of the PIDA in reproducing the global thermospheric density. Furthermore, the retrieved altitude profiles of various parameters in the thermosphere from the daytime limb scans by Global Ultraviolet Imager (GUVI) on board the Thermosphere-Ionosphere-Mesosphere Energetics and Dynamics (TIMED) satellite are compared with the corresponding assimilation results from the PIDA. Moreover, future improvement of the physics-based assimilation system is discussed.

2. Models and Data

2.1. TIEGCM

The TIEGCM version 2.0 release is used in this study. The TIEGCM, as an open-source community model, is a global 3-D numerical model that simulates the coupled ionosphere-thermosphere system from ~97 to ~600 km altitude (Richmond et al., 1992). This model self-consistently solves the coupled nonlinear momentum, energy, and continuity equations for neutral and ion species. As the most widely used physics-based model in the upper atmospheric community, the TIEGCM has been validated since it is capable of capturing the most significant features in the ionosphere-thermosphere system (e.g., Chartier et al., 2013; Lei et al., 2007; Qian & Solomon, 2011; Ren et al., 2018, and references therein). The horizontal resolution in this study is 5° × 5° in longitude and latitude; the vertical resolution is half of a scale height. Migrating diurnal and semidiurnal tides at the lower boundary are specified using the global-scale wave model (Hagan & Forbes, 2002, 2003). Additionally, the magnetospheric convection at high latitudes is specified by the Heelis model as a function of the Kp index (Heelis et al., 1982). The scalar solar activity proxy \( F_{10,7} \) is used in this model to derive the strength of solar flux. The \( Kp \) and \( F_{10,7} \) indices during days 80–280 in 2004 that were used in the default simulation run are shown in Figures 1a and 1b.

2.2. Data

In this study, the neutral density from the CHAMP (Reigber et al., 2002) satellite is used to drive the PIDA, and the neutral density from the GRACE (Tapley et al., 2004) satellite is compared with model output. The CHAMP and GRACE neutral density data used here was derived by Mehta et al. (2017), who reestimated the drag coefficient using numerical simulations performed with the test particle Monte Carlo method. They
found that the newly derived density data sets showed an average bias of 14–18% for CHAMP and 10–24% for GRACE with respect to the original one. Note that in this study, we do not further apply the scaling factor to the CHAMP or GRACE densities. Since CHAMP and GRACE take about 130 and 160 days to cover all local time, the CHAMP orbit density from day 80 to day 280 in 2004 is used to drive the PIDA. Figures 1c and 1d display the local time and orbit-mean altitude of the CHAMP or GRACE satellite. The gray shadings represent the altitudinal ranges of the CHAMP and GRACE orbits. It should be noted that GRACE flew at about 90 km higher than CHAMP, and the local time of CHAMP and GRACE had about 4–6 h difference during the analyzed period.

The altitude profiles of multiple parameters, including neutral density, temperature, mass mixing ratio of atomic oxygen ($\Psi_O$), mass mixing ratio of molecular nitrogen ($\Psi_{N_2}$), and the ratio of the O number density to the $N_2$ number density ($O/N_2$) from the daytime limb scans by GUVI on board the TIMED satellite (Christensen et al., 2003), are further compared with model output. The altitude profiles of the number density of different neutral species (e.g., O, $N_2$, and $O_2$) and temperature between 110 and 667 km were inverted from the limb scans of far ultraviolet dayglow by using a forward model (Meier et al., 2015; Meier & Picone, 1994). The local time and latitude of the GUVI limb samples are shown in Figure 2. The local time of the CHAMP orbit is also shown by the blue lines in Figure 2a for comparison. The local time of the limb samples was generally limited during 6–20 LT. Additionally, the differences between the local time of the limb observations and the CHAMP orbit varied from 0 to 6 h depending on days in 2004. In order to avoid the auroral contamination, the limb observations at locations lower than 60° geographic latitude were derived (Emmert et al., 2006) and presented in Figure 2b.

### 2.3. PIDA

The PIDA used in this study is developed according to the IRIDEA system by Sutton (2018). In the IRIDEA and the PIDA, only the $Kp$ and $F_{10.7}$ indices, which represent the effective external forcing, are modified in order to match the neutral density measured by the LEO satellite. In this way, the state of the ionosphere-thermosphere system is updated self-consistently. Because the thermosphere needs time to respond to the changes in external forcing, the model is reinitialized to step back for a certain period to account for the response time of the thermosphere. For each iteration step in the IRIDEA and PIDA, the thermosphere-ionosphere model TIEGCM is reinitialized prior to 24 h by using the same initial state, and the neutral density from CHAMP during the latest 18 h is used to update the $Kp$ and $F_{10.7}$ indices. Note that, in the IRIDEA by Sutton (2018), the 1 day effective $F_{10.7}$ value and the latest 9 h effective $Kp$ index are
estimated repeatedly until the estimated results agree well enough with the observations. Under this situation, the bias in density between observation and the results from IRIDEA during 6–15 hr is attributed to the $F_{10.7}$ effect and thus may affect the estimation of $K_p$ during the period of the latest 9 hr. Therefore, unlike in the IRIDEA, the effective $K_p$ indices during 6–15 hr are also updated in the PIDA but with a constraint. As shown in Table 1, the maximum modification of the $K_p$ index during 6–12 hr is 0.33 based on the forecasted one from the last converged step, and the maximum change of the $K_p$ index during 12–15 hr is 0.67. Then, the Karush-Kuhn-Tucker conditions are applied to calculate the optimal estimation of effective $K_p$ and $F_{10.7}$. Additionally, the initial condition on day 80 is directly taken from the default run, which is driven by the actual $F_{10.7}$ and $K_p$ indices. Moreover, a 3 hr update window is applied in the PIDA after the iterative procedure converged.

3. Results
3.1. Comparison With Orbital Density

The variations of the orbit averaged neutral density from the CHAMP satellite are shown in Figure 3a. The gray line in Figure 3a denotes the observed results, and the blue and red lines represent the corresponding simulation results from the TIEGCM and PIDA, respectively. It is seen that the neutral density from the TIEGCM is greater than the corresponding observations during the entire analyzed period. In the meanwhile, the neutral density from the PIDA shows good agreement with the observations.

Figure 3b displays the scatterplots of the observed orbit averaged neutral density from CHAMP and the simulation results from the TIEGCM (the gray dots) and PIDA (the red dots). Obviously, the neutral density from the PIDA has a higher correlation with the observations than that from the TIEGCM. The correlation

| Table 1 | Constraint of the Effective Parameters That Represent the Effective External Forcing in Each Iteration Step of PIDA |
|---------|------------------------------------------------------------------------------------------------------------------|
| Parameter | $F_{10.7}$ | $K_p$ |
| $T$ (hr) | $0–24$ | $0–3$ | $3–6$ | $6–9$ | $9–12$ | $12–15$ | $15–18$ | $18–21$ | $21–24$ |
| Upper | $300$ | $K_p^f$ | $\min(K_p^f + 0.33, 9)$ | $\min(K_p^f + 0.67, 9)$ | $9$ |
| Bottom | $50$ | $K_p^f$ | $\max(K_p^f – 0.33, 0)$ | $\max(K_p^f – 0.67, 0)$ | $0$ |

Note. $K_p^f$ denotes the effective $K_p$ before updating.
The correlation coefficients between the orbit averaged neutral density from CHAMP and the simulation results are given in Table 2. The correlation coefficients increase from 0.68 for the TIEGCM to 0.99 for the PIDA. As displayed in Table 2, the averaged ratio of the TIEGCM simulation results to the observations is 1.48; i.e., the mean bias of the orbit averaged neutral density from the TIEGCM is 48%. For the PIDA, the mean bias decreases to less than 1%.

Figures 3c and 3d further display the scatterplots for the daytime and nighttime orbit mean density from the CHAMP satellite, respectively. It is seen that the correlation coefficients of daytime and nighttime density increase from 0.71 and 0.65 for the TIEGCM to 0.97 and 0.96 for the PIDA, respectively. Additionally, the daytime neutral density from the PIDA is 4% lower than the observations, and the nighttime one is 13% greater than the observation results, which is also seen from Sutton (2018). Overall, the comparison results suggest that the density from the CHAMP satellite has been effectively ingested in the PIDA.

The observation and simulation results of the orbit averaged neutral density from the GRACE satellite are shown in Figure 4a. The PIDA also has a higher accuracy than the TIEGCM when comparing with the observed density from GRACE. As shown by the scatterplot of the observed and simulated orbit averaged neutral density for the GRACE satellite in Figure 4b, the correlation between the observation result and the density from the PIDA is much higher than that from the TIEGCM. The correlation coefficients and mean ratios of the simulation results to observations for the GRACE satellite are also displayed in Table 2.

Table 2
The Correlation Coefficients and the Mean Ratio (With Standard Deviation) of the Simulation Results to Observations for the Orbit Averaged, Daytime Averaged, and Nighttime Averaged Density of the CHAMP and GRACE Satellites

| Satellite | Data        | Orbit average | Daytime     | Nighttime   |
|-----------|-------------|---------------|-------------|-------------|
|           |             | TIEGCM | PIDA | TIEGCM | PIDA | TIEGCM | PIDA |
| CHAMP     | Correlation coefficient | 0.68     | 0.99 | 0.71   | 0.97 | 0.65   | 0.96 |
|           | Mean ratio  | 1.48 ± 0.36 | 1.00 ± 0.04 | 1.39 ± 0.32 | 0.96 ± 0.07 | 1.70 ± 0.53 | 1.13 ± 0.15 |
| GRACE     | Correlation coefficient | 0.42     | 0.92 | 0.55   | 0.95 | 0.83   | 0.95 |
|           | Mean ratio  | 1.50 ± 0.57 | 0.94 ± 0.15 | 1.39 ± 0.59 | 0.86 ± 0.20 | 1.94 ± 0.61 | 1.33 ± 0.24 |

Figure 3. (a) Variation of the observed (gray line) and simulated (blue line for the TIEGCM; red line for PIDA) orbit averaged density for CHAMP samples during days 80–280 in 2004. (b) The scatterplot of observed orbit averaged density for CHAMP samples and the simulation results from the TIEGCM (gray dots) and PIDA (red dots). The scatterplots of the (c) daytime and (d) nighttime observed orbit averaged density for CHAMP samples and the simulation results from the TIEGCM (gray dots) and PIDA (red dots). The daytime orbit was selected for the orbit local time during 8–20 LT, the nighttime orbit was during 20–24 and 0–8 LT, and the data sets used here were sampled at locations lower than 60° latitude.
The correlation coefficient is 0.42 for the TIEGCM, and it increases to 0.93 for the PIDA. It is noted that the neutral density from the PIDA is approximately 6% lower than the observation results. The scatterplots of the daytime and nighttime results are shown in Figures 4c and 4d. As shown in Table 2, both the correlation coefficients for the daytime and nighttime data sets are higher than 0.94 for the PIDA. Additionally, the daytime neutral density from the PIDA is about 14% lower than the observations, and the nighttime one is 33% greater than the observations. The high correlation coefficient between the GRACE observations and the density from the PIDA suggests that the PIDA can well reproduce the temporal evolution of the thermospheric density. However, it should be noted that a systematic bias exists between the density from the PIDA and the GRACE orbit density.

3.2. Comparison With Limb Measurements

The variations of the limb density at 278, 388, and 428 km and the corresponding simulation results from the TIEGCM and PIDA are shown in Figure 5. The densities from the TIEGCM at these three altitudes are greater than the observation results. As shown in Figure 5b, the density from the PIDA agrees well with the limb density at 388 km, which is close to the altitude of the CHAMP orbit. With regard to the density variation at 428 km as shown in Figure 5a, the density from the PIDA is slightly lower than the observations. This situation is also observed from the orbit averaged and the daytime density from GRACE, which flies at about 90 km higher than CHAMP. As shown in Figure 5c, the PIDA improves the accuracy of the simulated density than that from the TIEGCM at 278 km. However, it is noteworthy that the density from the PIDA overestimates the observations at 278 km.

The scatterplots of the observed density and simulation results from the TIEGCM (gray dots) and PIDA (red dots) at 428, 388, and 278 km are shown in Figure 6. The corresponding correlation coefficients and mean ratios of simulation to observation results are displayed in Table 3. The correlation coefficients between the observed and simulated densities from the TIEGCM are about 0.7 at these three altitudes, and the correlation coefficients increase to more than 0.9 for the densities from the PIDA. The mean ratio of the PIDA modeled and observed densities is 0.97 at 388 km, which is close to the altitude of the CHAMP orbit. It is noted that the density at 428 km from the PIDA is about 11% lower than the observations and 27% greater than the observations at 278 km. This situation further suggests that the PIDA does a good job in reproducing the evolution of the thermospheric density at around the orbiting altitude of the satellite that...
is applied in driving the assimilation model. Moreover, the PIDA also has the capability of capturing the temporal evolution of the thermospheric density at different altitudes, though a systematic bias, depending on altitude, is seen in the density of the PIDA.

Note that the exponential decrease in the thermospheric density with altitude is generally determined by the scale height, which is closely related to the thermospheric temperature. That is, the thermospheric temperature is one of the most important parameters in determining the thermospheric structure. Figure 7 shows the variation of the observed and simulated orbit mean temperature of the limb samples at 388 km and the scatterplot of the observation and simulation results from the TIEGCM and the PIDA. The correlation coefficients and mean ratios of the modeled to the observed temperature are also displayed in Table 3. The correlation coefficient increases from 0.82 for the results from the TIEGCM to 0.86 for those from the PIDA. It should be noted that the mean ratio of the TIEGCM-modeled temperature to the observations is 0.96, and it decreases to 0.88 for the PIDA. This implies that, though the density at 388 km is well estimated in the PIDA, the temperature from the PIDA is lower than the observations by about 12%.

Figure 5. Variation of the observed (gray line) and simulated (blue line for the TIEGCM; red line for PIDA) orbit averaged density for GUVI limb samples at (a) 428 km, (b) 388 km, and (c) 278 km during days 80–280 in 2004.

Figure 6. The scatterplot of the observed orbit averaged density for GUVI limb samples and the simulation results from the TIEGCM (gray dots) and PIDA (red dots) at (a) 428 km, (b) 388 km, and (c) 278 km.
Figures 8a and 8c display the altitude profiles of the temporal and spatial averaged temperature and density from limb observations and the corresponding simulation results from the TIEGCM and PIDA, respectively. The mean ratios of the simulations to observations for temperature and density are shown in Figures 8b and 8d, respectively. The standard deviations are marked by the corresponding colored shadings in Figures 8b and 8d. Because the modeled range of the PIDA is sometime below than 519 km, the statistic results from the PIDA at 519 km is not shown in Figure 8. It can be seen that the thermospheric temperature from the TIEGCM is on the average about 5% lower than the observations at the altitude above 150 km, whereas it is about 22% lower for the PIDA as compared with the observations. The temperatures from the TIEGCM and PIDA at below than 130 km are greater than the observed. With regard to the simulated density by the TIEGCM, the density during the focused altitude (i.e., 110–520 km) is higher than the observations. The maximum bias of the density is about 1.8 at 140 km (i.e., the simulated density at 140 km is on the average about 80% greater than the observations). As shown by the red line in Figure 8d, the ratio of the PIDA modeled densities to limb observations at around 350 km decreases to nearly 1. That is, the PIDA has replicated the density measurements from the CHAMP satellite, and the density at near the CHAMP altitude is well estimated. Note that the adjusted $K_p$ and $F_{10.7}$ mainly affect the state of the upper thermosphere, and the results in the lower thermosphere (cf. below than 150 km) from the PIDA are not greatly altered by this procedure. Therefore, the thermospheric temperature in the upper atmosphere in the PIDA decreased in order to match the observed lower CHAMP density. This situation thus results in the thermospheric scale height being lower than the actual thermospheric conditions. Because temperature change affects density in a height-integrated way, the assimilated density from the PIDA at lower altitude is greater than the observations, and it would be lower than the observations at higher altitude. However, it should be noted that, in reality, the higher density at the CHAMP altitude from the PIDA could be alternately associated with the overestimation of the density in the lower thermosphere rather than of the thermospheric temperature.

Figures 9a and 9c further displays the altitude profiles of the averaged $\Psi_O$ and $\Psi_{N_2}$, respectively, from the limb observations, TIEGCM, and PIDA. The mean ratios of the simulated to observed $\Psi_O$ and $\Psi_{N_2}$ samples

| Data | $\rho$ at 428 km | $\rho$ at 388 km | $\rho$ at 278 km | $T_n$ at 388 km |
|------|-----------------|-----------------|-----------------|-----------------|
|      | TIEGCM          | PIDA            | TIEGCM          | PIDA            | TIEGCM          | PIDA            |
| Correlation coefficient | 0.72 | 0.93 | 0.72 | 0.94 | 0.70 | 0.95 | 0.70 | 0.95 |
| Mean ratio | 1.41 ± 0.48 | 0.89 ± 0.17 | 1.45 ± 0.43 | 0.97 ± 0.16 | 1.60 ± 0.29 | 1.27 ± 0.12 | 0.96 ± 0.06 | 0.88 ± 0.05 |

Figure 7. (a) Variation of the observed (gray line) and simulated (blue line for the TIEGCM; red line for PIDA) orbit averaged temperature for GUVI limb samples at 388 km during days 80–280 in 2004. (b) The scatterplot of the observed orbit averaged temperature for GUVI limb samples and the simulation results from the TIEGCM (gray dots) and PIDA (red dots) at 388 km.
are shown in Figures 9b and 9d, respectively. The $\Psi_O$ from the TIEGCM is lower than the observations from the lower thermosphere to about 300 km, and it is greater than the observations at higher altitudes. The $\Psi_{N_2}$ at altitudes lower than 300 km is about 10% greater than the observations, and the averaged ratio then decreases with altitude to about 0.7 at 520 km. It is noted that the PIDA also generally does not affect the abundance of the neutral species in the lower thermosphere. Due to the lower temperature, which causes the contraction of the thermosphere in the PIDA as compared with that in the TIEGCM, the averaged ratios of simulated to observed $\Psi_O$ and $\Psi_{N_2}$ from the PIDA increase and decrease accordingly. We also calculated the correlation coefficient between the observed and simulated $O/N_2$, which is proportional to the ratio of $\Psi_O$ to $\Psi_{N_2}$. For the results from the TIEGCM, the correlation coefficient for $O/N_2$ is 0.52 at 110 km, and it increases to 0.82 at above 300 km. However, the correlation coefficient for $O/N_2$ between the observations and the PIDA simulation results does not show a significant increase as compared with the results from the TIEGCM. Overall, the results suggest that the evolution of the abundance of the neutral species is not well estimated in the PIDA. However, future investigations are required to examine the uncertainties in the GUVI limb data retrieval and the possible association with the discrepancies between the limb data and the PIDA.

Figure 8. The altitude profiles of the temporal and spatial averaged (a) temperature and (c) density from limb observations (gray dotted lines) and the corresponding simulation results from the TIEGCM (blue dotted lines) and PIDA (red dotted lines). The mean ratio of the simulation results (blue line for the TIEGCM; red line for PIDA) to observations for (b) temperature and (d) density. The standard deviations were marked by the corresponding colored shading in (b) and (d).

Figure 9. The altitude profiles of the temporal and spatial averaged (a) $\Psi_O$ and (c) $\Psi_{N_2}$ from limb observations (gray dotted lines) and the corresponding simulation results from the TIEGCM (blue dotted lines) and PIDA (red dotted lines). The mean ratios of the simulation results (blue line for the TIEGCM; red line for PIDA) to observations for (b) $\Psi_O$ and (d) $\Psi_{N_2}$. The standard deviations were marked by the corresponding colored shading in (b) and (d).
4. Discussion and Summary

In this study, the CHAMP orbit density was used to drive the physics-based data assimilation system PIDA, in which the 1 day history effective parameters, $F_{10.7}$ and $K_p$, were estimated. The density from the PIDA agreed well with the CHAMP observations, which suggested that the CHAMP orbit density has been effectively ingested in the PIDA. Then, the GRACE density and limb density from GUVI were used to evaluate the capability of the PIDA in reproducing the thermospheric state. It was found that the PIDA could well reproduce the thermospheric density at around the altitude where the orbit density was ingested in the PIDA. Additionally, the correlation coefficients between the observed density and the corresponding simulation results from the PIDA were higher than 0.9 at different altitudes in the thermosphere. That is, the PIDA is also capable of capturing the temporal evolution of the thermospheric density at different altitudes.

However, a systematic bias in the lower thermosphere was observed in the PIDA in simulating the thermospheric density. Under the hydrostatic assumption in the neutral thermosphere (Nicolet & Mange, 1954; Rishbeth et al., 1987), the altitudinal profile of the thermospheric density depends closely on the neutral density in the lower thermosphere and the thermospheric scale height. It was found that the temperature and density in the TIEGCM are greater than the corresponding observations in the lower thermosphere. Although the parameters $K_p$ and $F_{10.7}$, which represent the effective external forcing, have significant effects on the upper thermosphere, they generally do not contribute greatly to modulating the state of the lower thermosphere. Therefore, the large bias still exists at different altitude of the thermosphere. The overestimation of the density in the lower thermosphere may be ascribed to the higher temperature in this region. It was observed that the cooling mechanisms due to the infrared radiation from CO$_2$ and NO, which maximizes between 100 and 200 km, played a significant role in controlling the thermospheric temperature (Roble, 1995), whereas the cooling rates in the TIEGCM have a large bias as compared with the observations during quiet and geomagnetic storm times (Chen & Lei, 2018; Li et al., 2018). Li et al. (2018) suggested that the tide specification in the TIEGCM lower boundary could be a possible contributor to the cause of this temperature overestimation in the lower thermosphere. That is, the tide speciation in the lower boundary of the TIEGCM should be an important issue in the further improvement of the physics-based thermosphere data assimilation system. Moreover, the assessments of the obtained results suggested that the spatial and temporal evolutions of the lower thermosphere state under various conditions should be considered in the data assimilation system in order to develop more accurate forecast systems for the thermosphere.

As a nonlinearly coupled complex system, thermospheric evolution depends heavily upon various parameters, such as temperature, wind, density, and neutral species (Shim et al., 2014). The abundance of the thermospheric neutral species and the O/N$_2$ ratio from limb observations were used to evaluate the physics-based data assimilation system PIDA. The thermospheric neutral species in the PIDA even showed larger bias than those from the TIEGCM at various altitudes. The PIDA, which solely applies the orbit density from a single satellite, cannot account for the information of the multiple parameters in the thermosphere. Consequently, the assimilated density and other parameters (i.e., thermospheric temperature and neutral species) would show a large bias. Our results also suggest that multiple parameters at different altitudes are required to be assimilated into the thermospheric model.

Data Availability Statement

Thermospheric orbital densities for CHAMP and GRACE are obtained from http://tinyurl.com/densitysets. The GUVI Limb data can be found from http://guvitimed.jhuapl.edu/data_products. The TIEGCM and PIDA simulation data, simulation codes, and analysis routines are available from http://iono.ustc.edu.cn/expdata/Publications/PIDA_datasets/.

References

Berger, T. E., Holzinger, M. J., Sutton, E. K., & Thayer, J. P. (2020). Flying through uncertainty. Space Weather, 18(1), e2019SW002373. https://doi.org/10.1029/2019SW002373
Cantrall, C. E., Matsuo, T., & Solomon, S. C. (2019). Upper atmosphere radiance data assimilation: A feasibility study for GOLD far ultraviolet observations. Journal of Geophysical Research: Space Physics, 124, 8134–8164. https://doi.org/10.1029/2019JA026910
Chartier, A. T., Jackson, D. R., & Mitchell, C. N. (2013). A comparison of the effects of initializing different thermosphere-ionosphere model fields on storm time plasma density forecasts. Journal of Geophysical Research: Space Physics, 118, 7329–7337. https://doi.org/10.1002/2013JA019034
Rishbeth, H., Fuller-Rowell, T. J., & Rees, D. (1987). Diffusive equilibrium and vertical motion in the thermosphere during a severe magnetic storm: A computational study. *Planetary and Space Science, 35*(9), 1157–1165. https://doi.org/10.1016/0032-0633(87)90022-5

Roble, R. G. (1995). Energetics of the mesosphere and thermosphere. In R. M. Johnson & T. L. Killeen (Eds.), *The upper mesosphere and lower thermosphere: A review of experiment and theory*. Geophysical Monograph Series (Vol. 87, pp. 1–21). Washington DC: AGU. https://doi.org/10.1029/GM087p0001

Ruan, H., Lei, J., Dou, X., Liu, S., & Aa, E. (2018). An exospheric temperature model based on CHAMP observations and TIEGCM simulations. *Space Weather, 16*, 147–156. https://doi.org/10.1029/2017SW001759

Shim, J. S., Kuznetsova, M., Rasttter, L., Bilitza, D., Butala, M., Codrescu, M., et al. (2014). Systematic evaluation of ionosphere/thermosphere (IT) models. *Modeling the ionosphere-thermosphere system* (pp. 145–160). Washington: John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118704417.ch13

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/10.1002/2017SW001785

Tapley, B. D., Bettadpur, S., Watkins, M., & Reigber, C. (2004). The gravity recovery and climate experiment: Mission overview and early results. *Geophysical Research Letters, 31*, L09607. https://doi.org/10.1029/2004GL019929

Weimer, D. R., Mehta, P. M., Tobiska, W. K., Dournhos, E., Mlynczak, M. G., Drob, D. P., & Emmert, J. T. (2020). Improving neutral density predictions using exospheric temperatures calculated on a geodesic, polyhedral grid. *Space Weather, 18*(1), e2019SW002355. https://doi.org/10.1029/2019SW002355