FULL PAPER

Improving the estimation of thermospheric neutral density via two-step assimilation of in situ neutral density into a numerical model
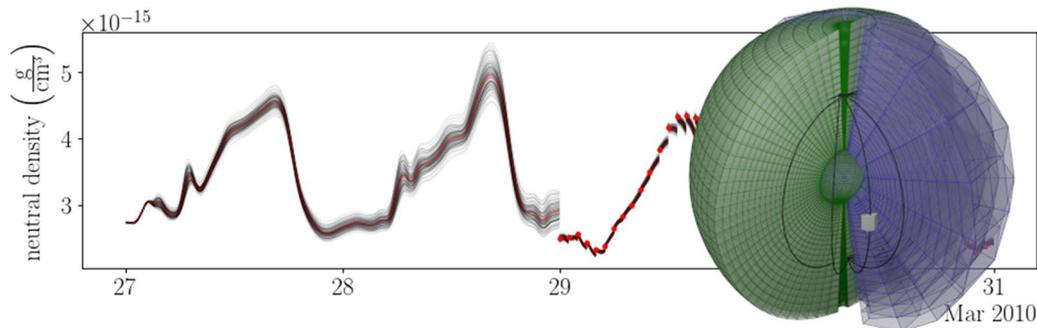
Armin Corbin[*] and Jürgen Kusche

Abstract

Neutral thermospheric density is an essential quantity required for precise orbit determination of satellites, collision avoidance of satellites, re-entry prediction of satellites or space debris, and satellite lifetime assessments. Empirical models of the thermosphere fail to provide sufficient estimates of neutral thermospheric density along the orbits of satellites by reason of approximations, assumptions and a limited temporal resolution. At high solar activity these estimates can be off by 70% when comparing to observations at 12-hourly averages. In recent decades, neutral density is regularly observed with satellite accelerometers on board of low Earth orbiting satellites like CHAMP, GOCE, GRACE, GRACE-FO, or Swarm. When assimilating such along-track information into global models of thermosphere–ionosphere dynamics, it has been often observed that only a very local sub-domain of the model grid around the satellite's position is updated. To extend the impact to the entire model domain we suggest a new two-step approach: we use accelerometer-derived neutral densities from the CHAMP mission in a first step to calibrate an empirical thermosphere density model (NRLMSIS 2.0). In a second step, we assimilate—for the first time—densities predicted for a regular three-dimensional grid into the TIE-GCM (Thermosphere Ionosphere Electrodynamics General Circulation Model). Data assimilation is performed using the Local Error-Subspace Transform Kalman Filter provided by the Parallel Data Assimilation Framework (PDAF). We test the new approach using a 2-week-long period containing the 5 April 2010 Geomagnetic storm. Accelerometer-derived neutral densities from the GRACE mission are used for additional evaluation. We demonstrate that the two-step approach globally improves the simulation of thermospheric density. We could significantly improve the density prediction for CHAMP and GRACE. In fact, the offset between the accelerometer-derived densities and the model prediction is reduced by 45% for CHAMP and 20% for GRACE when applying the two-step approach. The implication is that our approach allows one to much better 'transplant' the precise CHAMP thermospheric density measurements to satellites flying at a similar altitude.

Keywords: Thermosphere, Neutral mass density, Data assimilation, Geomagnetic storm

*Correspondence: corbin@geod.uni-bonn.de

Institute of Geodesy and Geoinformation, University of Bonn, Nussallee 17, Bonn 53115, Germany

Graphical Abstract

Background

Neutral thermospheric density plays an important role when computing the atmospheric drag acceleration acting on satellites (e.g., Vallado and Finkleman 2014). Applications in need for accurate drag estimates are precise orbit determination (e.g., Montenbruck and Gill 2005; Longuski et al. 2022), the forecasting of orbit decay or mission life time (e.g., Walterscheid 1989), and predicting the re-entry of satellites or space debris and identifying locations on Earth that might be endangered by this (e.g., Klinkrad et al. 2006). Especially, satellites flying below 1000 km are affected by atmospheric drag that is the largest non-gravitational acceleration at those altitudes.

One can determine the neutral density with different methods (Emmert 2015). In situ measurements are conducted with mass spectrometers or accelerometers mounted on satellites. However, there are only a few satellites equipped with such instruments. It is also possible to estimate time averaged neutral densities from observed satellite orbits. The advantage of this method is that it can be applied to any passive satellite, but it is less accurate when relying on two line elements (TLE) tracking. In case of active satellites—for example, satellites equipped with retro reflectors for satellite laser ranging—the time averaged neutral density can be determined within precise orbit determination with higher precision than using TLE.

Another approach is the use of thermospheric density models. In fact, several numerical and empirical models were developed over the last seven decades (e.g., Doornbos (2012, Table 2.1) and Vallado and Finkleman (2014, Figure 3)). Empirical density models are constructed from observations that are fitted to mathematical equations. This is in particular critical for the effect of solar and geomagnetic forcing on density, since the underlying physics including Joule heating, photochemistry, and particle precipitation is partly not understood owing to a lack of data, and partly too complex for these simple equations. Empirical density models represent the average state of the atmosphere (e.g., Emmert 2015). Examples for empirical models are the Jacchia–Bowman (JB, Bowman et al. 2008) model, Naval Research Laboratory Mass Spectrometer and Incoherent Scatter radar (NRLMSIS 2.0, Emmert 2021) model, and Drag Temperature Model (DTM, Bruinsma and Boniface 2021).

Numerical models propagate an initial state using physical laws and principles, for instance, heat and momentum balance, electromagnetism and chemical reactions. Typically, this is done by solving a set of (partial) differential equations on a grid. Examples for numerical models are the National Center for Atmospheric Research Thermosphere Ionosphere Electrodynamics General Circulation Model (NCAR TIE-GCM, Qian et al. 2014), NCAR Whole Atmosphere Community Climate Model (WACCM-X, Liu et al. 2018), and Global Ionosphere–Thermosphere Model (GITM, Ridley et al. 2006).

There are significant discrepancies between different models, and between models and observations (e.g., Gaposchkin and Coster 1990; Bruinsma et al. 2012, 2014; He et al. 2018; Panzetta et al. 2019). Hence, there is an ongoing effort in improving the models. In this paper, we present a new experimental approach for that purpose.

While numerical models provide physically consistent solutions, they do not exhibit improved skills in neutral density simulation when compared to empirical models (e.g., Emmert 2015). A common approach to nudge numerical model simulations closer to reality is merging them with observations via data assimilation.

In a coupled model, like the TIE-GCM, assimilating quantities associated with a compartment also affects the other compartments. One can assimilate observations of the electron density to improve the representation of the neutral mass density. For example, the total electron
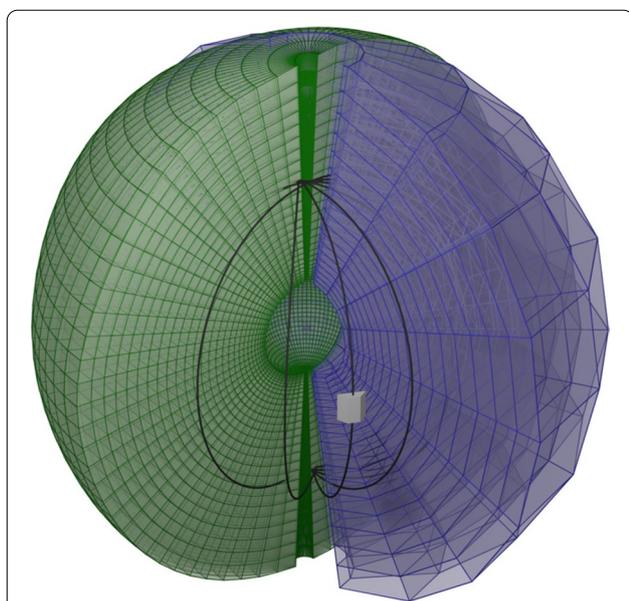
Fig. 1 The green lines illustrate the grid of the TIE-GCM at one epoch. We call it the *state grid*. The blue lines show the (constant) grid at which we evaluate the NRLMSIS 2.0. We call it the *data grid*. The black curve illustrates the orbit of the CHAMP satellite for some revolutions. In the two-step approach we first calibrate the NRLMSIS 2.0 with the accelerometer-derived density of the CHAMP satellite, and evaluate it on the *data grid*. In a second step we assimilated the densities located on the *data grid*—in data assimilation terminology this are the observations—into the TIE-GCM. The grids appear thicker as they are in reality, since the radius of Earth is not added. Actually both grids cover the entire Earth, but we show only a subset for a clearer illustration

content, an integrated measure of electrons along a path through the atmosphere, could be used for that. Although we only assimilate neutral mass densities here, our assimilative version of the TIE-GCM can be easily modified to assimilate other quantities.

The in situ measurements of a single satellite mission only intersect with a small subset of the model grid cells. This means, compared with models the in situ measurements of a single satellite are very sparse. The farther an observation is away from a grid cell, the less information is provided by it for that cell. For instance, a measurement of the neutral density on the day side provides little information about the density on the night side. Thus, assimilating such along-track data should only affect grid cells in the vicinity of the satellite's orbit. Matsuo et al. (2013) have assimilated neutral densities derived form the CHAMP (Reigber et al. 2002) accelerometer into the TIE-GCM. The model densities were only improved in the vicinity of the satellite's orbit. But they achieved global improvements by co estimation of model drivers.

Besides co estimating parameters for the model drivers, one could also assimilate data of many sources at different locations simultaneously for global model improvements. In this study we test another approach that consists of two steps and is illustrated in Fig. 1. We use the along-track densities derived from the CHAMP accelerometer to calibrate an empirical model (the NRLMSIS 2.0). The calibrated empirical model is viewed here as a combination of the data used to build the empirical model itself and the densities derived from the CHAMP accelerometer, with more weight given to the CHAMP observations. For the calibration we evaluate the empirical model along the CHAMP orbit and scale the observed densities by the modeled densities to derive scale factors. We also apply a low pass filter to the scale factors. The calibrated model is the output of the original model multiplied with the corresponding scale factor. In the first step, we evaluate the calibrated empirical model on a regular three-dimensional grid. We call it the *data grid*. In the second step, we assimilated the data located on the *data grid* into the TIE-GCM that is located on what we call the *state grid*.

Data assimilation of various observations types has already been applied to different models of the upper atmosphere in several studies: Solomentsev et al. (2012) have assimilated simulated GPS observations into a numerical model of the ionosphere using the Ensemble Square Root Filter. The aim was assessing the state estimation of the ionosphere and improving the estimation of model drivers. Observations from GPS Occultation have been assimilated by Lee et al. (2012) into the TIE-GCM using the Ensemble Kalman filter under geomagnetic quiet conditions. They aimed at improving the global ionospheric electron density specification. Matsuo et al. (2013) have assimilated CHAMP observations and GPS Occultation measurements into the TIE-GCM using the Ensemble Kalman filter. They found that assimilation of accelerometer-derived densities only improves the model densities in the vicinity of the satellite's orbit. But they also demonstrated that co estimating the F10.7 parameter together with accelerometer-derived densities impacts the global model. In the study of Morozov et al. (2013) the ensemble adjustment Kalman filter is used to assimilate CHAMP observations into the Global Ionosphere–Thermosphere Model during a geomagnetically calm period. They estimate the F10.7 index in a way it has a constant variance. They could reduce the model bias along the CHAMP and GRACE orbits. Codrescu et al. (2018) have assimilated neutral densities derived from accelerometers on board of the CHAMP mission into the CTIPe model during quiet conditions at solar minimum using the Ensemble Kalman Filter. The model results were improved when comparing to CHAMP and GRACE

observations. Forootan (2022) have applied a calibration and data assimilation technique to the empirical NRLM-SISE-00 model using observations from the GRACE accelerometer.

In contrast to the existing approaches, the two-step approach enables us to assimilate globally distributed neutral densities that are derived from satellite accelerometers and an empirical model, which is also build from many observations.

Methods

Data assimilation

Data assimilation (e.g., Lahoz et al. 2010) combines the state estimation of a model with observations taking into account the uncertainty of both to get an estimate of the state with higher accuracy. There are many different data assimilation approaches. For highly nonlinear problems, like atmosphere models, ensemble Kalman filters (e.g., Vetra-Carvalho et al. 2018) are frequently employed. By implicitly representing the variance–covariance matrix of the model state by an ensemble of states, ensemble filters are very efficient in terms of computational cost and computer memory requirements and can be easily integrated into the model code. These filters utilize a sequence of *forecast* and *analysis* steps. The forecast step uses the model to predict the observations. At the subsequent analysis step the state of the model is fitted so that it minimizes the distance to the forecasted observations and the actual observations w.r.t. the associated variance–covariance matrices.

The Parallel Data Assimilation Framework (Nerger et al. 2020) developed at the Alfred Wegener Institute Bremerhaven is open-source software for ensemble-based data assimilation. It is designed for large scale numerical models and allows the application of different filter algorithms. Moreover, it enables parallel computation of all ensemble members, which in turn can also be calculated in parallel if supported by the model. Both, the TIE-GCM and PDAF are written in Fortran which simplifies the implementation. In this study, we use the localized (Nerger et al. 2006) error-subspace transform Kalman filter (Nerger et al. 2012b, ESTKF).

The original ensemble Kalman filter (Evensen 1994, EnkF) formulation is typically expanded with perturbed observations to account for an underestimation of the analysis error covariance that leads to a too small ensemble spread after the analysis step (Burgers et al. 1998; Houtekamer and Mitchell 1998). The perturbed observations are an additional source for sampling errors (Whitaker and Hamill 2002), which are avoided by a class of filters using a deterministic transformation from the forecast ensemble to the analysis ensemble (e.g., Tippett et al. 2003). We choose The ESTKF since it belongs to

this class and is formulated in an efficient way (Nerger et al. 2012b).

Determining neutral densities from space-borne accelerometers

Given the atmospheric drag acting on a satellite, together with a model describing the shape and material of the satellite one can derive the neutral mass density at the satellite's position (e.g., Doornbos 2012, p.91). Accelerometers onboard satellites measure the superposition of all non-conservative accelerations acting on the satellite. To isolate the atmospheric drag acceleration from the measurements, one needs to carefully model all other non-conservative forces and remove their effect. In practice, one needs to simulate the acceleration caused by thermal re-radiation, solar radiation pressure and Earth radiation pressure, and remove it from the the measured accelerations. For this study, we use the neutral mass densities derived from the accelerometers on board of the CHAMP (Reigber et al. 2002) and GRACE (Tapley et al. 2007) mission using the approach described in Vielberg et al. (2018); Vielberg and Kusche (2020). Further information can be found at Vielberg (2021).

Neutral density models used in this study

The NCAR Thermosphere Ionosphere Electrodynamics General Circulation Model (TIE-GCM, Qian et al. 2014), represents a global, numerical model of the upper atmosphere. The TIE-GCM ranges from approximately 97 km to 500 km altitude. The upper boundary is not fixed in geometric height coordinates, since the TIE-GCM uses pressure levels as vertical coordinate. For this study we use the latest version of the TIE-GCM (version 2.0). An important proxy driver for the TIE-GCM is the F10.7 index (e.g., Tapping 2013) that is used to compute the extreme ultraviolet (EUV) radiation, based on the model of Richards et al. (1994). The TIE-GCM includes two alternative empirical high-latitude potential models computing ionospheric convection: the Heelis model (Heelis et al. 1982) and the Weimer model (Weimer 2005). The first requires the three hourly Kp index (e.g., Matzka et al. 2021), whereas the second uses solar wind and interplanetary magnetic field parameters provided by the OMNI Dataset with one minute temporal resolution (Papitashvili and King 2020).

The Naval Research Laboratory Mass Spectrometer Incoherent Scatter radar 2.0 (NRLMSIS 2.0, Emmert 2021) model is a global, empirical model of the atmosphere. It takes location, time, geomagnetic activity represented by the Kp index, and solar activity represented by the F10.7 index as input. It computes the neutral composition, density, and temperature. Since it is derived from various observations at different periods and locations
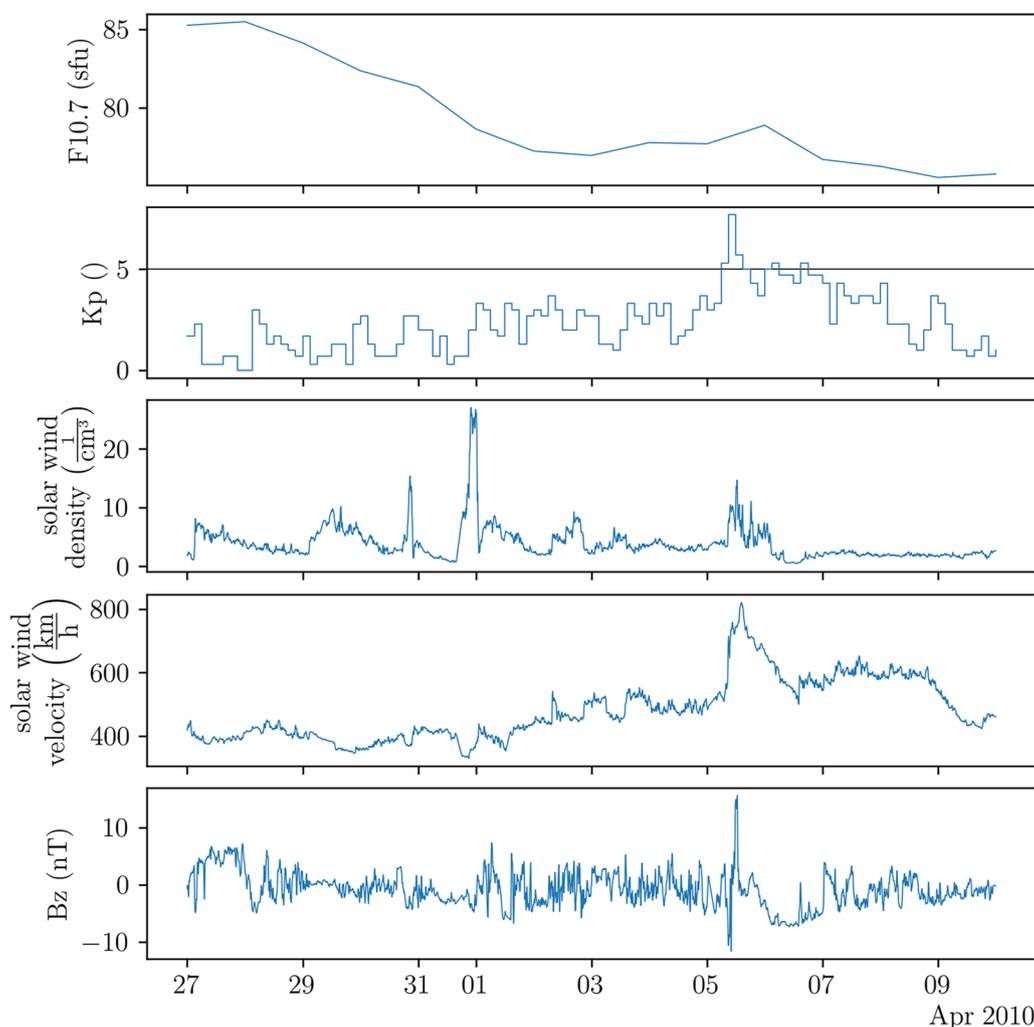
Fig. 2 External forcing time series during the period of the experiment required to run the NRLMSIS 2.0 and TIE-GCM. A Kp value greater than or equal to five is considered as storm. The main storm event is at 5th April from 6:00 to 18:00 UTC+0. The temporal resolution of the F10.7 index and Kp index is one day and three hours, respectively. The solar wind and the z component of the interplanetary magnetic field are taken form the OMNI data set and have a 1-min resolution

the results represent the averaged observed state of the atmosphere for the given inputs. The model extends from the ground to the exobase.

Period for experiments

The period of the assimilation experiment is restricted by three factors: first, it must contain measurements of the CHAMP and GRACE missions. That is, the period must be between 2002 and 2010. Second, the period must contain at least one strong storm (Kp ≥ 7), but also quiet condition, so we can evaluate the assimilation framework for different geomagnetic activity. Finally, the duration is restricted by the computing time. A 2-week-long period

is processed on our hardware (400 cores distributed over 25 Intel(R) Xeon(R) Gold 6130 processors) in about four hours and 20 min. This allows us to test different settings in a reasonable amount of time. We choose the period from 27 Mar 2010 00:00 UTC+0 till 10 April 2010 00:00 UTC+0, that satisfies all three conditions. In Fig. 2, the solar and geomagnetic activity during the experiments is illustrated. On April 5, 2010, an interplanetary coronal mass ejection reached Earth around 8:27h UTC+0 and triggered a geomagnetic storm (Lu et al. 2014; Sheng et al. 2017). The mean altitudes of the CHAMP and GRACE satellite during the experiment are 302 km and 474 km, respectively.
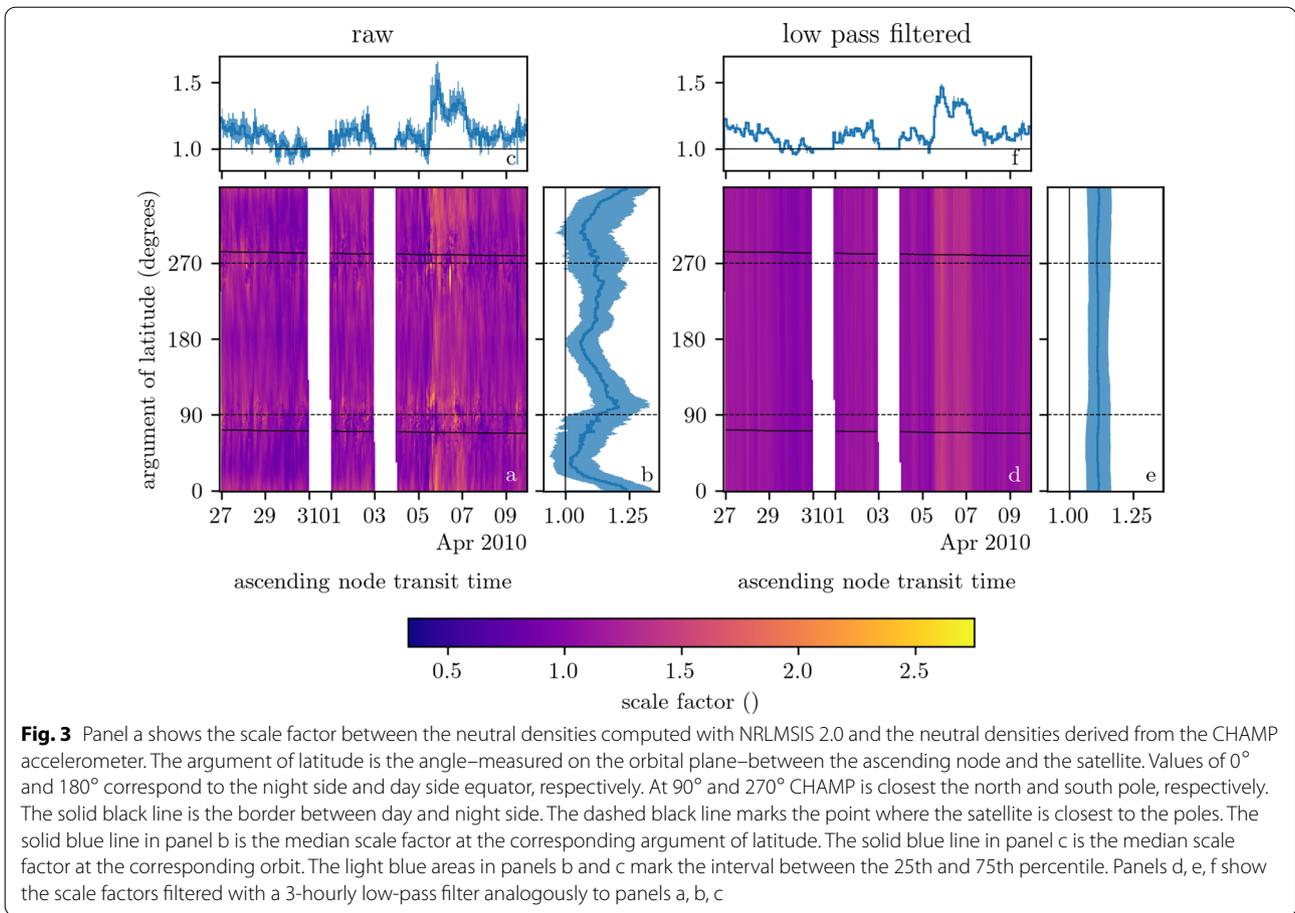
Fig. 3 Panel a shows the scale factor between the neutral densities computed with NRLMSIS 2.0 and the neutral densities derived from the CHAMP accelerometer. The argument of latitude is the angle–measured on the orbital plane–between the ascending node and the satellite. Values of 0° and 180° correspond to the night side and day side equator, respectively. At 90° and 270° CHAMP is closest the north and south pole, respectively. The solid black line is the border between day and night side. The dashed black line marks the point where the satellite is closest to the poles. The solid blue line in panel b is the median scale factor at the corresponding argument of latitude. The solid blue line in panel c is the median scale factor at the corresponding orbit. The light blue areas in panels b and c mark the interval between the 25th and 75th percentile. Panels d, e, f show the scale factors filtered with a 3-hourly low-pass filter analogously to panels a, b, c

Calibration of NRLMSIS 2.0

The majority of the observations used to estimate the NRLMSIS 2.0 parameters are between the ground and 105 km altitude (Emmert 2021, Table 1). Above this altitude only density observations derived from two line elements were used along with synthetic observations from the preceding model NRLMSISE-00. Since, the TIE-GCM starts at approximately 97 km altitude it does not intersect with most of the observations used for building the NRLMSIS 2.0. Moreover, empirical models return average states of the atmosphere. Thus, we calibrate the model to perform better at the altitudes covered by the TIE-GCM.

We calibrate the NRLMSIS 2.0 by scaling it with time dependent factors. The scale factor is thus defined as the quotient of the density derived from the CHAMP accelerometer ρ_{CHAMP} and the density $\rho_{\text{NRLMSIS 2.0}}$ predicted by the NRLMSIS 2.0 for the corresponding location and time (see also Zeitler et al. 2021, Eq. 10):

$$s(t) = \frac{\rho_{\text{CHAMP}}(t)}{\rho_{\text{NRLMSIS 2.0}}(t)}. \tag{1}$$

In Fig. 3, the scale factors for the duration of the experiment are plotted. During and after the strong storm (starting at 5th April), the scale factors are much larger compared to the rest of the period (Fig. 3 a and c). Within each orbit, the scale factors vary according to latitude and whether the satellite is on the day or night side (Fig. 3 b). For example, on the night side equator (argument of latitude is zero) the scale factors are systematically larger than on the day side equator (argument of latitude is 180°). The median scale factor on the night side is about 16% larger than the median on the day side.

As shown, the scale factor depends on the horizontal location. However, for each epoch, we can only derive the scale factor at one location if we use only one satellite. That is, the scale factor is expected to work best for locations near the CHAMP satellite.

Thus, we decided to filter out the orbital signal in the scale factor time series. We found that a cut-off frequency of three hours, which approximately corresponds to two revolutions of the CHAMP satellite, eliminates this signal (see Fig. 3 d and e). That is, the variability between day and night side within an orbit vanishes in the filtered scale factor time series. The three hourly filtered scale factor is an average value that can be applied to the whole model, at the cost of signal loss (compare Fig. 3 a and d). Consequently, the cut-off frequency also limits the duration between subsequent analysis steps.

As shown in Zeitler et al. (2021), half daily scale factors derived at different heights are highly correlated. For CHAMP and GRACE the correlation is 89%. Thus, when calibrating the NRLMSIS 2.0 with observations from the CHAMP mission, we expect that the calibrated model also fits better to the densities derived from the GRACE accelerometer.

Spatial resolution of the data grid

We choose the horizontal resolution of the *data grid*, so that it reflects the resolution of the NRLMSIS 2.0. The NRLMSIS 2.0 uses spherical harmonics up to degree six to expand the model parameters to the global atmosphere (Hedin (1987, Equation A22) and Emmert et al. (2021, Section2.4)). That is, the NRLMSIS 2.0 cannot resolve signals with wavelength smaller than $\frac{360°}{6} = 60°$

arc length. To sample these signals, one needs at least half the wavelength. Here, we choose a third with gives $20°$ for the horizontal resolution of the *data grid*. Using a finer grid would increase the number of observations and slow down the assimilation while the gain in information is limited. The *data grid* starts at 100 km and ends at 550 km covering most of the *state grid*. The vertical resolution is 25 km. This value corresponds approximately to the number density scale height at 200 km altitude.

Setup for TIE-GCM

We use the latest version of the TIE-GCM (version 2.0). The TIE-GCM either runs with $5.0°$ or $2.5°$ horizontal resolution. For this study, we use the coarser five degree resolution since it runs about ten times faster and requires one eighth of the disk space to store the results at the same temporal resolution. The step size—the time between subsequent model states—is 15 seconds. We found it difficult to use longer values as this caused some ensemble members to crash. We do not save every model step, since it would require too much memory: Saving the ensemble mean of the neutral density at the *state grid* at each model step for a 2-week-long period using single precision requires approximately 24 GB. But we also need to save the data at the *data grid* and the corresponding standard deviations. If one also stores the results of each

Table 1 The quantities that are perturbed to generate the ensemble are listed in the first column

	Units	Standard deviation	Truncation	Min	Max
External forcing					
F10.7	sfu	1	± 50	75.9	87.8
Hemispheric power	GW	5	± 9	15.3	155.0
Cross tail potential	kV	8	± 12	23.2	170.4
Solar wind density	$1/cm^3$	0.1	± 0.3	0.5	27.1
Solar wind velocity	km/h	45	± 135	329.3	820.8
Lower boundary conditions					
Neutral temperature	K	3	± 9	151.5	198.8
Neutral zonal wind	cm/s	500	± 1500	-5914.5	5841.5
Neutral meridional wind	cm/s	650	± 1950	-8073.3	8077.5
Geopotential height	m	70	± 210	97246.7	98652.7
Constants					
Characteristic Maxwellian energy of polar cusp electrons	keV	0.01	± 0.03	0.1	
Characteristic Maxwellian energy of drizzle electrons	keV	0.05	± 0.15	0.5	
Joule heating factor	–	0.15	± 0.45	1.5	

The solar wind parameters are only perturbed if the Weimer model is used. Whereas hemispheric power and cross tail potential are only perturbed in case of the Heelis model. The third and fourth column are the parameters used for sampling the truncated normal distributions. The mean is always zero. Note that correlations are not listed in this table. The truncation is symmetric. That means values larger than the positive truncation threshold as well as values less than the negative truncation threshold are rejected. The last two columns contain the minimal and maximal values of the unperturbed quantities during the experiment

ensemble member the storage requirements increase accordingly. Thus, we save the results every ten minutes.

The TIE-GCM requires lower boundary constraints for neutral temperature, horizontal neutral wind, and geopotential altitude. By default a 'flat' lower boundary is assumed. That is, there is no wind, the neutral temperature is 181 K, and the geopotential height is 96.4 km. Alternatively, one can use zonal monthly mean climatologies to specify the lower boundaries. This feature is based on the work of Jones et al. (2014). The TIE-GCM comes with a file containing zonal climatologies derived from the NRLMSISE-00 (Picone et al. 2002) and the Horizontal Wind Model (HWM07, Drob et al. 2008). To use it with the five degree grid one has to interpolate, since it is only given for the 2.5 degree grid. Regardless which method is chosen to specify the lower boundary, we additionally add tidal perturbations derived from the Global Scale Wave Model 2002 (GSWM-02 Hagan and Forbes 2002) to it to account for migrating diurnal and semidiurnal tides.

Ensemble generation

We create the ensemble by adding perturbations to the external forcing, lower boundary conditions and some constants to each ensemble member that are sampled from a truncated multivariate normal distribution with zero mean. The truncation of the normal distribution is necessary to prevent model crashes caused by extreme values sampled from the tails of the normal distribution that are invalid or unrealistic. We sample the perturbations once and use the same values for the entire duration of the experiment. For each lower boundary condition, only one value is sampled, which is added to all elements of the corresponding field. We do not consider errors within the field, but a global offset of the lower boundary condition. In Table 1, the parameters of the probability density function are listed.

Since we have no access to the true probability density function of the perturbed parameters, we have to determine it to a certain extend arbitrarily. The standard deviation of the constants that we perturb is assumed to be 10% of the constant itself. Following Tapping (2013) we assume that the standard deviation of the F10.7 index is one solar flux unit for values smaller than 100. When using the Heelis model we perturb hemispheric power and cross tail potential. Since both are computed within the TIE-GCM from the Kp index we introduce a correlation of 0.9. If the Weimer model is employed, we perturb the solar wind velocity and density instead. The lower boundary conditions of the TIE-GCM are perturbed by diurnal and semidiurnal tides computed by the global scale wave model

(GSWM). We decided to approximate the standard deviation of the lower boundary conditions from the GSWM tidal perturbations. We take 10% of the largest absolute tidal perturbation in March.

We use an ensemble with 100 members. The initial value for the assimilation are computed with an open-loop simulation, which starts 10 days before the assimilation experiment.

State vector and observation operator

The TIE-GCM approximates the state of the atmosphere at an epoch with multiple quantities located on the *state grid*, given for the current and previous model step. Since TIE-GCM computes time derivatives simply via finite differences from the previous and current steps, this can be viewed as equivalent to storing variables and their derivative in the state vector. The neutral mass density can be linked to these fields via the mass fractions w of the modeled species and the neutral temperature T. Thus, in our study the state vector is composed of

$$\boldsymbol{x} = [\boldsymbol{T}, \boldsymbol{T}', \boldsymbol{w}_{\mathrm{O}}, \boldsymbol{w}'_{\mathrm{O}}, \boldsymbol{w}_{\mathrm{O}_2}, \boldsymbol{w}_{\mathrm{O}_2}', \boldsymbol{w}_{\mathrm{He}}, \boldsymbol{w}'_{\mathrm{He}}]. \qquad (2)$$

O, O_2, and He denote atomic oxygen, molecular oxygen and atomic helium, respectively. The quantities evaluated at the previous step are denoted with a prime symbol.

The observation operator H is a composition of the function computing the neutral mass density $\rho(\boldsymbol{x})$ and an interpolation function $I()$ that computes the values at the *data grid* given values on the *state grid*:

$$\boldsymbol{y} = H(\boldsymbol{x}) = (I \circ \rho)(\boldsymbol{x}). \qquad (3)$$

For each cell of the *state grid* and each step, we can compute the neutral mass density assuming an ideal gas with

$$\rho = \frac{p\overline{M}(w_{\mathrm{O}}, w_{\mathrm{O}_2}, w_{\mathrm{He}})}{RT}. \qquad (4)$$

Here, p is the pressure, R denotes the gas constant and $\overline{M}$ is the mean molar mass

$$\overline{M}(w_{\mathrm{O}}, w_{\mathrm{O}_2}, w_{\mathrm{He}}) = \frac{1}{\dfrac{w_{\mathrm{O}}}{M_{\mathrm{O}}} + \dfrac{w_{\mathrm{O}_2}}{M_{\mathrm{O}_2}} + \dfrac{w_{\mathrm{He}}}{M_{\mathrm{He}}} + \dfrac{w_{\mathrm{N}_2}}{M_{\mathrm{N}_2}}}. \qquad (5)$$

The molar mass of a species is denoted with M, and the mass fraction of molecular Nitrogen is $w_{\mathrm{N}_2} = 1 - w_{\mathrm{O}} - w_{\mathrm{O}_2} - w_{\mathrm{He}}$.

Since the *state grid* is irregularly spaced in the vertical dimension, we first perform linear interpolation along this axis. Since the neutral mass density decreases almost exponentially with height, we apply the natural logarithm to the neutral densities before performing the vertical

interpolation. This ensures small interpolation and especially small extrapolation errors, since on a logarithmic scale the density profiles are almost linear. We then apply the exponential function to the vertically interpolated values to transform the interpolated values back to neutral mass densities. Two more linear interpolations along the horizontal coordinates axes are performed to compute the mass densities on the *data grid*. If the observations are given exactly at the time of the current step, another interpolation along the time axis is not necessary.

Observations
At each analysis step we assimilate the neutral mass densities from the calibrated NRLMSIS 2.0 located on the *data grid* into the TIE-GCM. These densities are not real observations, but we still use the term 'observation' here to be consistent with data assimilation terminology. The current implementation only supports uncorrelated observation. Thus, although the observations are highly correlated we cannot account for it by using the full variance–covariance matrix of the observations.

We approximate the standard deviation of the NRLMSIS 2.0 neutral mass density by multiplying the uncalibrated neutral mass density ρ with a factor f_h depending on the altitude. The height dependent factor is obtained from the standard deviations provided in (Emmert 2021, Data Set 5) for the epoch 2006-2013 at 250 km and 400 km altitude. We use linear interpolation to get the standard deviation for arbitrary heights:

$$f_h(h) = \begin{cases} 0.148 + \dfrac{h - 250}{1500} & h < 400\,km \\[2ex] 0.248 + \dfrac{h - 400}{3260} & h \geq 400\,km. \end{cases} \quad (6)$$

We introduce additional weights: $p_{\mathrm{ga}}(\mathrm{Kp})$, $p_{\mathrm{dist}}(d)$, and p_0 to account for the geomagnetic activity, the distance to the CHAMP satellite and a constant factor, respectively. The standard deviation is computed with

$$\sigma_\rho(\rho, h, \mathrm{Kp}, d) = f_h(h)\,\rho\,\frac{1}{p_{\mathrm{ga}}(\mathrm{Kp})}\frac{1}{p_{\mathrm{dist}}(d)}\frac{1}{p_0}. \quad (7)$$

We use the Kp index as indicator of the geomagnetic activity. For quiet periods ($\mathrm{Kp} < 4\frac{2}{3}$) the weight is one. For the maximal Kp value of 9, the factor is $\frac{1}{2}$. The values in between are linearly interpolated. This factor ensures that the NRLMSIS 2.0 observations have a lower weight during storms.

$$p_{\mathrm{ga}}(\mathrm{Kp}) = \begin{cases} 1 & \mathrm{Kp} < 4\frac{2}{3} \\[2ex] \dfrac{40 - 3\,\mathrm{Kp}}{26} & \mathrm{Kp} \geq 4\frac{2}{3}. \end{cases} \quad (8)$$

The standard deviation is weighted by the distance between the CHAMP satellite and the center of the corresponding *data grid* cell. We use two exponential decay functions depending on the spherical distance Δ_ϕ on the unit sphere and the vertical geocentric distance Δ_h:

$$p_{\mathrm{dist}}(\Delta_\phi, \Delta_h) = \exp\left(-\frac{\ln(2)}{\Lambda_\phi}\Delta_\phi\right)\exp\left(-\frac{\ln(2)}{\Lambda_h}|\Delta_h|\right). \quad (9)$$

The weighting is controlled by the half life parameters Λ_ϕ and Λ_h. For the experiments presented in this study Λ_ϕ is infinite. That is, the spherical distance has no impact on the weights.

Forecast duration
The forecast duration is the duration between two subsequent analysis steps. We refer to model runs with infinite forecast duration, i.e., runs in which no data are assimilated, as open-loop simulations. If the forecast duration is too long, it will resemble the open-loop simulation some time before the next analysis step. If it is too short, the model state is mainly constrained by the observations and not the model dynamics. In that case the assimilation is useless since one could use the observations directly. For this consideration also the standard deviation of the observations and the model forecast are important. If the observations have much larger uncertainties than the forecast, the result resembles the open-loop simulation and the other way round.

The lower bound of the forecast duration is given by the model step length. The forecast duration should not exceed the cut-off period of the low pass filter applied to the scale factors.

The model step length is 15 seconds and the NRLMSIS 2.0 has been filtered with a three hourly low pass filter. Thus, we choose to perform the analysis step hourly.

Localization
At each analysis step of a global filter, each element of the state vector is updated taking into account all observations, regardless of how far the observations are away from the location of the corresponding element of the state vector. This can be of advantage if there are long-range correlation in the system (e.g., the ocean or the atmosphere). For instance, a single in situ observation within the atmosphere could theoretically improve the state estimate of the whole atmosphere. However, this requires that significant long-range correlations exist and that they are correctly represented by the ensemble. Typically, the random errors in the representation of the covariances are larger than the actual signal for small

Table 2 Setup for open-loop simulations. Only settings that differ are listed

Experiment identifier	OLS 1	OLS 2	OLS 3
High-latitude potential model	Heelis	Heelis	Weimer
Lower boundary	flat	MSIS & HWM	MSIS & HWM

ensembles (Nerger et al. 2006, p.640). This can lead to spurious correlations (e.g., Hamill et al. 2001) and locally incredible estimates (Nerger et al. 2006, p.640). This problem is addressed by filtering out long-range correlations in the analysis step by applying localization (e.g., Nerger et al. 2012a).

For the ESTKF, domain localization (Nerger et al. 2006) together with an optional observation localization (Hunt et al. 2007) is implemented within PDAF: the *state grid* is subdivided into disjoint sub-domains. Only observations whose distance from the center of the corresponding sub-domain is smaller than a cut-off radius are used to update the elements of the state vector within the sub-domain.

The geometric horizontal extent of the *state grid* cells and *data grid* cells increases with altitude. An edge at 550 km altitude is 7% larger than at 100 km. Moreover, the vertices of each grid are located closer to each other at the poles. At the equator the distance between two neighboring vertices along a circle of latitude is almost six times larger than at 80° latitude. Additionally, the vertical extent of the *state grid* cells increases with altitude due to the use of pressure levels. To include about the same number of observations at each sub-domain, we compute the distance using the indices of the *state grid* cells as coordinates. To transfer the coordinates of the *data grid* to the index coordinates of the *state grid*, we use the geometric height of the ensemble mean.

We subdivided the *state grid* into sub-domains containing at most three cells in meridional, zonal and vertical direction. The distance is computed using the L2 norm. The cut-off radius is seven grid cells, which corresponds to a horizontal metric radius of about 4000 km at the equator and about 700 km at ±80° latitude.
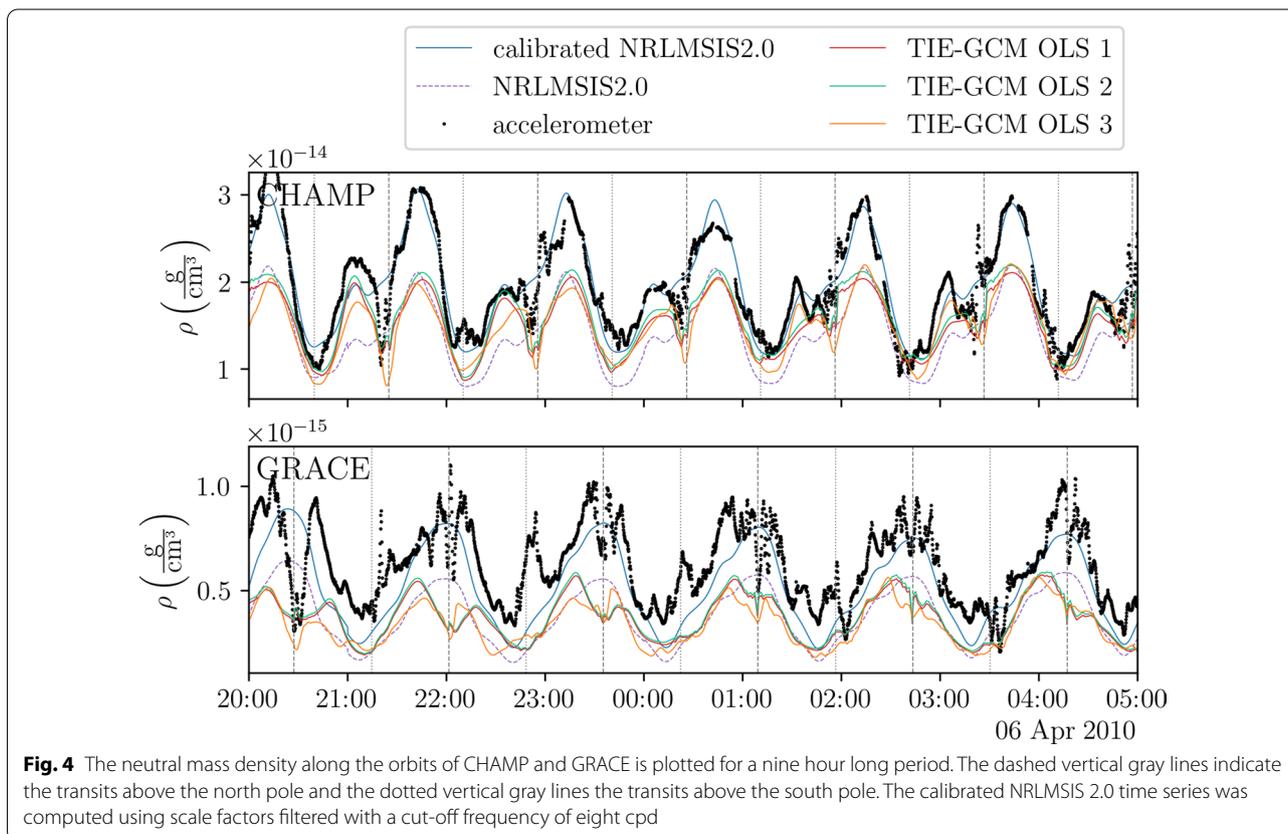

Fig. 4 The neutral mass density along the orbits of CHAMP and GRACE is plotted for a nine hour long period. The dashed vertical gray lines indicate the transits above the north pole and the dotted vertical gray lines the transits above the south pole. The calibrated NRLMSIS 2.0 time series was computed using scale factors filtered with a cut-off frequency of eight cpd
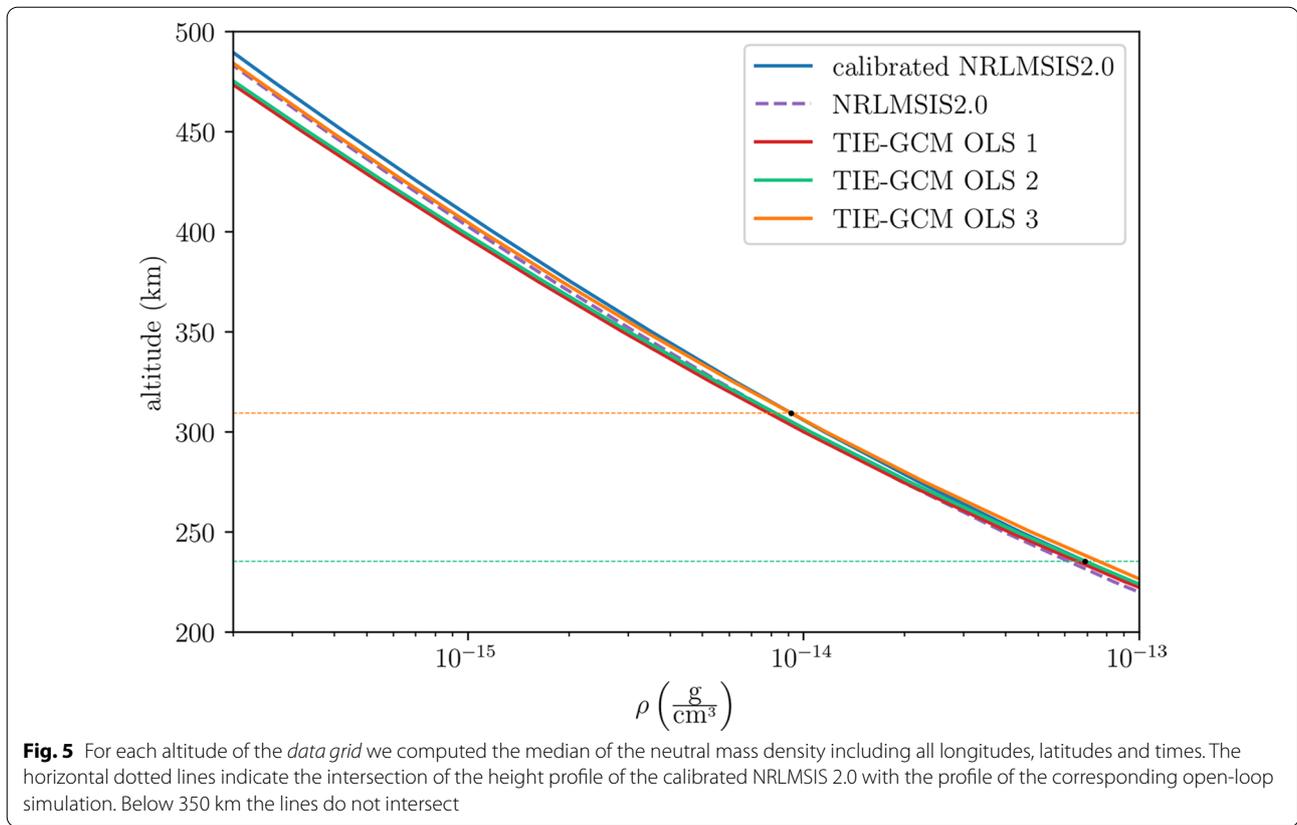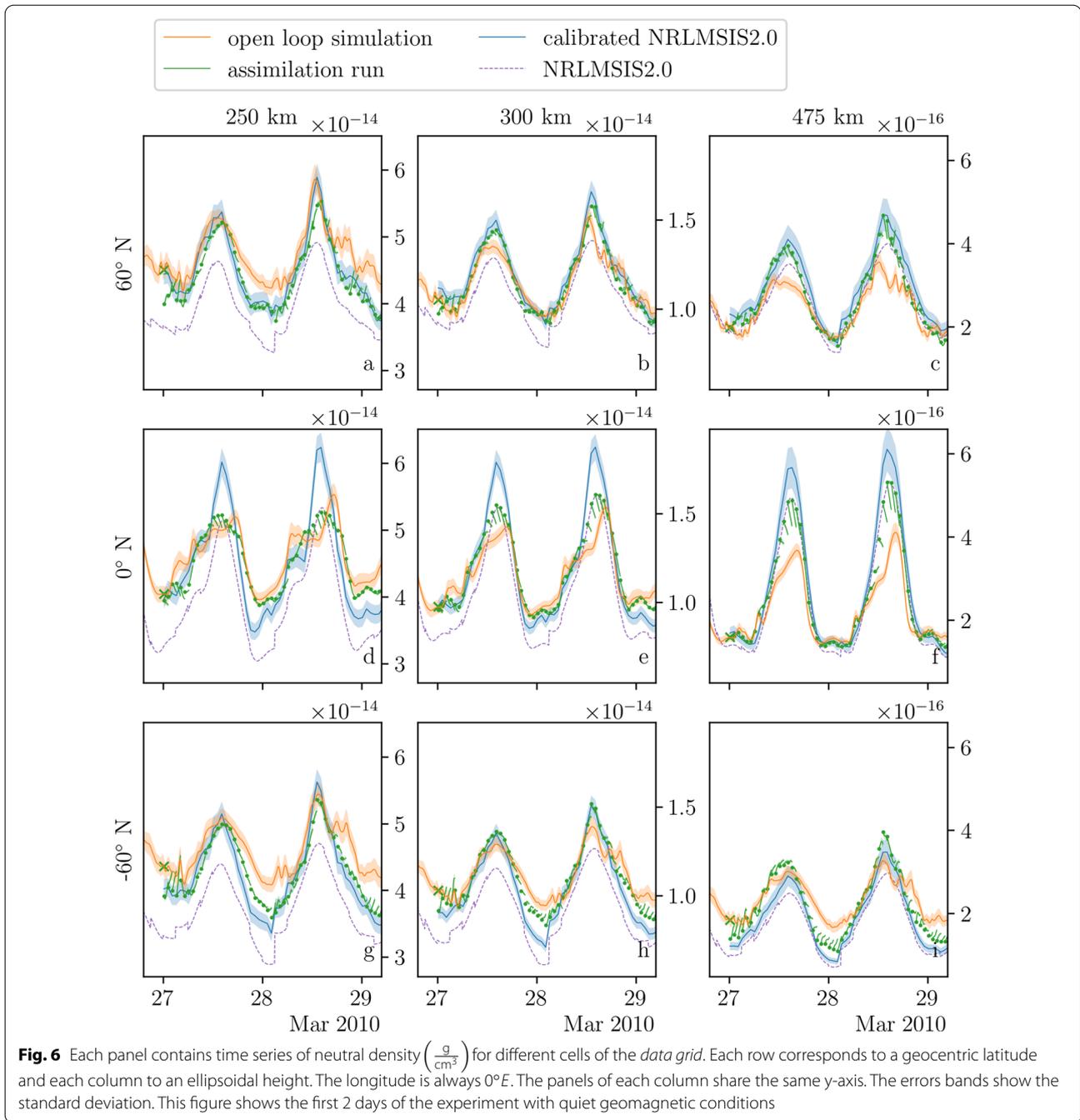
Fig. 5 For each altitude of the *data grid* we computed the median of the neutral mass density including all longitudes, latitudes and times. The horizontal dotted lines indicate the intersection of the height profile of the calibrated NRLMSIS 2.0 with the profile of the corresponding open-loop simulation. Below 350 km the lines do not intersect

Table 3 Setups for the data assimilation experiments

ID	completion (%)	Λ_h (km)	p_0	obs. loc.	median difference ($\frac{g}{cm^3}$)		RMSE ($\frac{g}{cm^3}$)	
					CHAMP	**GRACE**	**CHAMP**	**GRACE**
01	100	∞	1.0	No	$-9.6e{-}16$	$-7.1e{-}17$	$2.3e{-}15$	$1.2e{-}16$
02	100	∞	2.0	No	$-2.9e{-}16$	$-5.3e{-}17$	$1.8e{-}15$	$1.0e{-}16$
03	100	∞	3.0	No	$4.9e{-}17$	$-4.2e{-}17$	$1.7e{-}15$	$9.7e{-}17$
04	25	∞	4.0	No				
05	100	150	3.0	No	$-2.8e{-}16$	$-5.1e{-}17$	$1.8e{-}15$	$1.0e{-}16$
06	100	300	3.0	No	$-1.1e{-}16$	$-4.8e{-}17$	$1.8e{-}15$	$1.0e{-}16$
07	100	600	3.0	No	$-3.4e{-}17$	$-4.5e{-}17$	$1.7e{-}15$	$9.9e{-}17$
08	100	300	4.0	No	$9.7e{-}17$	$-4.1e{-}17$	$1.7e{-}15$	$9.6e{-}17$
09	53	150	6.0	No				
10	54	200	4.5	No				
11	100	∞	3.0	Yes	$-1.3e{-}15$	$-7.9e{-}17$	$2.6e{-}15$	$1.3e{-}16$
12	100	300	4.0	Yes	$-1.2e{-}15$	$-7.4e{-}17$	$2.5e{-}15$	$1.3e{-}16$
OLS 1	100				$-1.8e{-}15$	$-9.8e{-}17$	$3.1e{-}15$	$1.6e{-}16$
OLS 2	100				$-1.4e{-}15$	$-9.1e{-}17$	$2.7e{-}15$	$1.6e{-}16$
OLS 3	100				$-1.8e{-}16$	$-5.1e{-}17$	$2.4e{-}15$	$1.4e{-}16$

A 'completion' of 100% means that the experiment ran to the end without any errors. Values below 100% indicate that the TIE-GCM crashed after the corresponding percentage of the experiment's intended duration. Λ_h is the half life parameter for weighting the densities on the *data grid* based ob the vertical distance to CHAMP (see Eq. 9). p_0 is the constant weighting factor introduced in Eq. 7. We abbreviate observation localization with 'obs. loc.'. 'median difference' refers to the median of the differences between ensemble mean and accelerometer derived density of the corresponding satellite. The RMSE is computed using the difference between ensemble mean and accelerometer derived density. The statistics are only computed for experiments with 100% completion to ensure comparability
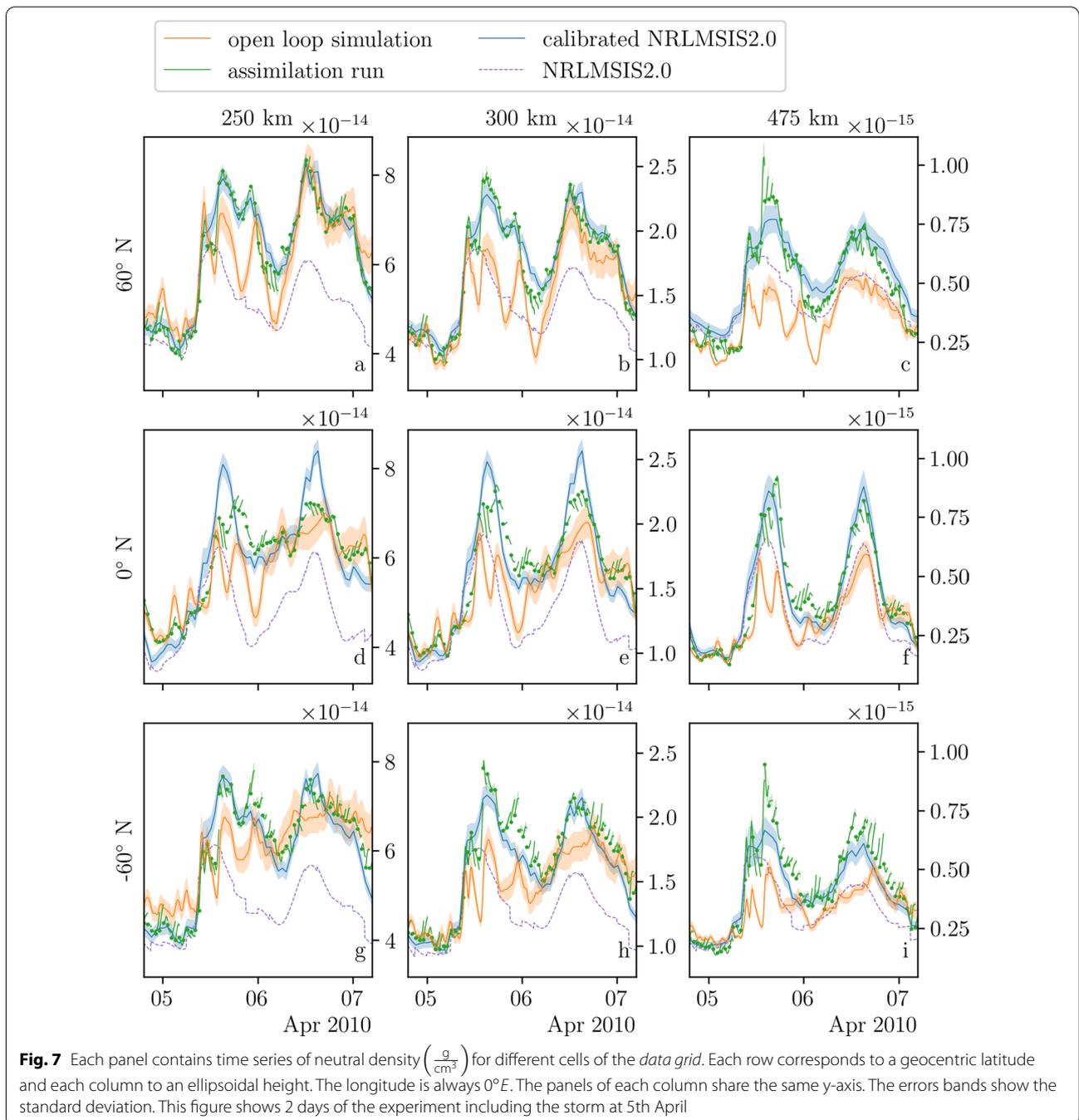
Fig. 6 Each panel contains time series of neutral density $\left(\frac{g}{cm^3}\right)$ for different cells of the *data grid*. Each row corresponds to a geocentric latitude and each column to an ellipsoidal height. The longitude is always 0°*E*. The panels of each column share the same y-axis. The errors bands show the standard deviation. This figure shows the first 2 days of the experiment with quiet geomagnetic conditions

Additionally, we apply observation localization in some experiments. For each sub-domain the associated observations are weighted based on the distance. PDAF computes the weights with a finite function that mimics a Gaussian, realized as polynomial of order five (Gaspari and Cohn 1999, Eq. 4.10). This weighting function monotonically decreases from one at zero distance to zero at a distance equal to the cut-off radius.

State constraints

After the analysis step the state vector may contain values that are physically impossible. Thus, we enforce the following constraints:

- The neutral temperatures given in Kelvin must be positive.

- The mass fraction of each species must be in the interval [0, 1].
- The mass fractions at each location must sum up to one.

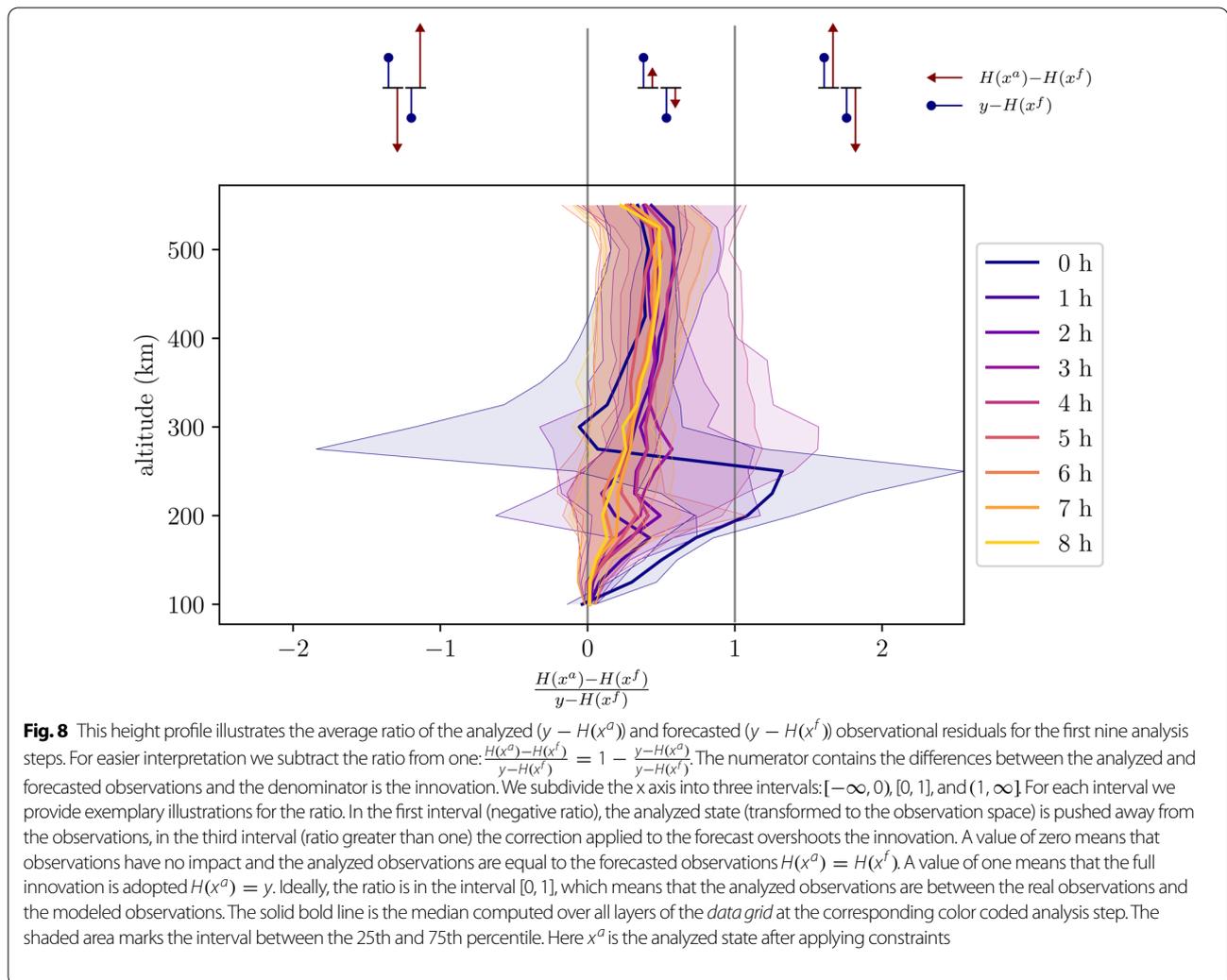**Fig. 7** Each panel contains time series of neutral density $\left(\frac{g}{cm^3}\right)$ for different cells of the *data grid*. Each row corresponds to a geocentric latitude and each column to an ellipsoidal height. The longitude is always $0°E$. The panels of each column share the same y-axis. The errors bands show the standard deviation. This figure shows 2 days of the experiment including the storm at 5th April

Results and discussion

Open-loop experiments

At first we try different setups for the TIE-GCM without assimilating any data. We investigate two setups for the lower boundary conditions: a flat lower boundary and a zonal mean climatology lower boundary derived from the MSIS and HWM. Additionally, we compare open-loop simulations using the Heelis and Weimer models. The Heelis model implies the use of the Kp index, whereas the Weimer model implies the use of solar wind and interplanetary magnetic field parameters. The F10.7 index is used in all experiments. The open-loop experiments are summarized in Table 2. The following discussion always refers to the ensemble mean of the open-loop experiments.

When using the zonal mean climatology instead of a flat lower boundary the median density (computed over the whole time period of the simulations) along the

CHAMP orbit increases by 5% (OLS 2 vs. OLS 1). When using the Weimer model instead of the Heelis model the median density along the CHAMP orbit increases by 13% (OLS 3 vs. OLS 2). The along-track densities of OLS 1 and OLS 2 show the same behavior and are basically separated by an offset, whereas the along-track densities of OLS 3 show other features (see Fig. 4), for example, a larger density drop at the north pole.

The height profiles showing the median neutral density in Fig. 5 associated with the OLS 2 and OLS 3 intersect with the profiles associated with the calibrated NRLM-SIS 2.0. The median density of OLS 2 and OLS 3 below the intersections is higher than the median density of the calibrated NRLMSIS 2.0. Accordingly, above the intersections the NRLMSIS 2.0 has higher densities. This means that at the analysis steps, the densities above the corresponding intersection must increase, while the densities below must decrease. In other words, the innovation (in
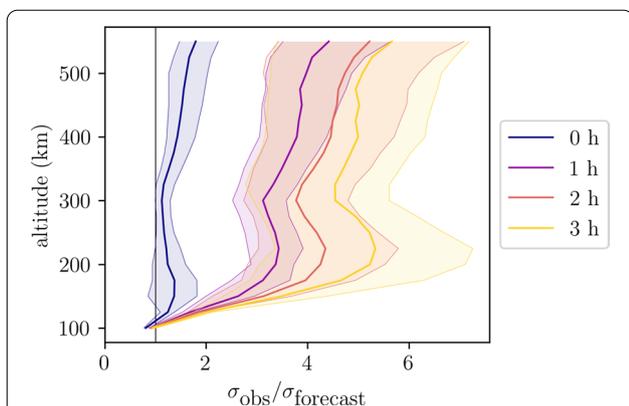

Fig. 8 This height profile illustrates the average ratio of the analyzed ($y - H(x^a)$) and forecasted ($y - H(x^f)$) observational residuals for the first nine analysis steps. For easier interpretation we subtract the ratio from one: $\frac{H(x^a)-H(x^f)}{y-H(x^f)} = 1 - \frac{y-H(x^a)}{y-H(x^f)}$. The numerator contains the differences between the analyzed and forecasted observations and the denominator is the innovation. We subdivide the x axis into three intervals: $[-\infty, 0)$, $[0, 1]$, and $(1, \infty]$. For each interval we provide exemplary illustrations for the ratio. In the first interval (negative ratio), the analyzed state (transformed to the observation space) is pushed away from the observations, in the third interval (ratio greater than one) the correction applied to the forecast overshoots the innovation. A value of zero means that observations have no impact and the analyzed observations are equal to the forecasted observations $H(x^a) = H(x^f)$. A value of one means that the full innovation is adopted $H(x^a) = y$. Ideally, the ratio is in the interval $[0, 1]$, which means that the analyzed observations are between the real observations and the modeled observations. The solid bold line is the median computed over all layers of the *data grid* at the corresponding color coded analysis step. The shaded area marks the interval between the 25th and 75th percentile. Here x^a is the analyzed state after applying constraints

Fig. 9 For the first four analysis steps, the average ratio of the standard deviation of the NRLMSIS 2.0, and the model forecast are plotted for each level of the *data grid*. A value of one means that both have the same uncertainty. Values greater than one indicate that the observation error is larger than the forecast error. The solid line is the median computed over all layers of the *data grid* at the corresponding color coded analysis step. The shaded area marks the interval between the 25th and 75th percentile

the language of data assimilation, it is generally the difference between observations and model forecasts) has a different sign depending on the altitude.

Data assimilation experiments

We tested different setups for the data assimilation that are summarized in Table 3. For the TIE-GCM we use the same setup as for the OLS 3, since the external forcing has the highest temporal resolution. We use the median of the differences between the accelerometer-derived neutral densities and the corresponding densities of the TIE-GCM as performance indicator for our experiments (see last column of Table 3). We show only the results of experiment 8 in more detail since it achieved the best improvement for GRACE and the second best improvements for CHAMP.

We present only the ensemble mean of the neutral densities, since it reduces the dimensionality of the ensemble, simplifying the illustration of the results.

Fig. 10 Each map corresponds to a pressure level of the *state grid*. The difference between the forecast mean and analysis mean of the neutral density $\left(\frac{g}{cm^3}\right)$ is indicated by the color bars on the right. Each row corresponds to a pressure level. each column to an analysis step. This figure shows the three first analysis steps. The horizontal position of CHAMP is marked with a black star, the position of the Sun with an orange circle, and the day–night border with an orange line

Fig. 11 Densities along the orbits of CHAMP and GRACE from various sources. The x axis is limited to a period where KP is always larger than or equal to 5 and contains the largest KP value of $7\frac{2}{3}$ in the experiment (9:00–12:00). The dotted vertical lines mark the analysis steps

We first look at the temporal evolution of the ensemble mean. Figure 6 shows the time series of the neutral mass densities at nine cells of the *data grid* for the first 2 days of the experiment. The columns contain the time series located at 250 km, 300 km, and 475 km, respectively. The average height during the experiment of the CHAMP and GRACE satellite are 302 km and 474 km, respectively. Thus, the second column is associated with the CHAMP mission and the third column with the GRACE satellite. The innovation is the difference between the densities of the calibrated NRLMSIS 2.0 and the densities forecasted by the model. At the first analysis step (27 March) and at 250 km (first column), the correction is larger than the innovation. This is especially visible at $60°N$ latitude (panel a). However, this *initial overshooting* does not affect the long term behavior of the assimilated time series. In fact the assimilated densities stays close to the calibrated NRLMSIS 2.0 densities after the second analysis step.

At 300 km and $60°N$ (panel b) one can observe the opposite: After the analysis step the density is pushed further away from the calibrated NRLMSIS 2.0. But, after

a few more analysis steps the assimilated time series follows the calibrated NRLMSIS 2.0 closely.

Figure 7 illustrates in the same way as Fig. 6 neutral density time series, but shows another period including the storm. The corrections are larger than in Fig. 6 showing quiet conditions. During the forecast phase the ensemble mean departs much faster from the analyzed state as under quiet conditions. Consequently, at the subsequent analysis steps larger corrections are necessary. The analyzed state does not fit to the model dynamics during the storm.

As suggested by Fig. 6, the overshooting after the first analysis step depends on the altitude. Figure 8 relates the innovation at each altitude of the *data grid* with the actual correction performed at the analysis step for the first nine analysis steps. Overshooting occurs mainly between 190 km and 260 km after the first analysis step (intersections of the dark blue line with right vertical line). The profiles of the analysis steps following the first one are similar: At 100 km the calibrated NRLMSIS 2.0 is basically ignored. With increasing altitude the influence of the calibrated NRLMSIS 2.0 gets larger. At

Fig. 12 The densities along the GRACE and CHAMP orbit were interpolated from the assimilated TIE-GCM and the open-loop simulation. Additionally, the (calibrated) NRLMSIS 2.0 was evaluated along the corresponding orbits. This figure shows the histograms of these three time series reduced by the accelerometer-derived densities. The histograms were calculated over the 14-day period of the experiment

500 km roughly half of the innovation is adopted at the analysis step. Above 500 km the influence of the calibrated NRLMSIS 2.0 decreases slightly.

In Fig. 9, we show the average ratio of the standard deviations of the calibrated NRLMSIS 2.0 (computed using Eq. 7) and the forecasted ensemble at the four first analysis steps. Before the first analysis step the median standard deviation of the calibrated NRLMSIS 2.0 is about one and a half times larger than the standard deviation of the ensemble. Only at 100 km altitude the calibrated NRLMSIS 2.0 has on average a slightly smaller standard deviation than the ensemble. After the first analysis step the ratio becomes larger, since the standard deviation of the ensemble is reduced: Above 200 km the standard deviation of the ensemble is about 4 times smaller. After the fourth step it is about 6 times smaller. The profiles have a local minimum at 300 km altitude that is caused by the weighting of the calibrated NRLMSIS 2.0 based on the vertical distance to CHAMP flying at that altitude.

The difference between the forecast and analysis ensemble mean is shown in Fig. 10 on world maps for different pressure levels and analysis steps. The position of CHAMP is marked with a star symbol. The model is updated globally, as expected. Above roughly 200 km negative values (density is increased at analysis step) mainly occur on the night side south pole. At the day side the density is mostly decreased at the analysis step (positive difference). Below 200 km the differences are almost only negative. A dependency on the location of the satellite is not visible.

To compare the neutral densities from the assimilation experiment with the densities derived from the accelerometers, we interpolate the densities on the *state grid* to the corresponding orbits using the same method that is used to interpolate the *state grid* to the *data grid*. In Fig. 11, the along-track neutral densities are plotted for a twelve hour long subset of the experiment including the storm. The assimilation run fits much better to the observations compared with the open-loop simulation, since it is not constantly too small. Indeed, it often resembles the

observed time series better than the calibrated NRLM-SIS 2.0 time series, for example, at 10:50, 14:05, 16:10. The same holds for the GRACE time series. Especially, from 15:00-18:00 the assimilation time series outperforms the others.

Figure 12 confirms the previous findings. Along the orbit of CHAMP the assimilation run is much closer to the accelerometer-derived densities than the open-loop simulation. Along the GRACE orbit the assimilation run also outperforms the open-loop simulation. The calibrated NRLMSIS 2.0 is closest to the accelerometer-derived densities and has the smallest spread for CHAMP and GRACE.

The major issue with our approach are large jumps in the ensemble mean time series especially during the storm. The model departs from the analyzed state much faster as during quiet conditions. We believe that those jumps can be reduced by co estimating suitable (internal) model parameters, that are not limited to corrections for external forcing, like the F10.7 proxy. For example, one could try corrections for hard coded parameters like chemical reaction rates, eddy diffusion, thermal diffusion, characteristic energy of particles, or solar proton flux, but at present we cannot yet provide a specific recommendation. Sheng et al. (2017) suggests that the cooling processes are not correctly represented in the TIE-GCM at the 5 April 2010 geomagnetic storm and Lu et al. (2014) suspects that the eddy diffusion coefficient is too high during the storm.

There are many options for the data assimilation setup: The ensemble generation, filter choice and related settings, and the calibration of the NRLMSIS 2.0 and the approximation of the standard deviation could be tuned in follow-on research. Within the limitations of this study, we have explored only a small subset.

A caveat is that our experiments were conducted in a period with small solar irradiance, and that for a period with high solar activity other settings might be necessary. Moreover, we suggest that further research could focus on adding, additional independent observations. An evident candidate are TEC values as provided via GNSS observations or—more representative for the low Earth orbit—from radar altimetry over the oceans.

Conclusions

We have implemented a new two-step approach for assimilating along-track observations into a physics-based model of the upper atmosphere. First an empirical model is calibrated via scale factors derived from the accelerometer aboard a LEO satellite. In a second step, the calibrated model is evaluated on a regular grid and the resulting neutral densities are assimilated into the numerical model. Here, we use densities derived from

the CHAMP accelerometer to calibrate the NRLM-SIS 2.0 and assimilate the densities into the TIE-GCM. We applied the approach to a 2-week-long period in 2010 including the 5 April 2010 geomagnetic storm.

We demonstrated that the assimilation approach has a global impact on the model and that the model prediction along the CHAMP and GRACE orbit fits well to the corresponding observations. When comparing the open-loop simulation with the assimilation run, the RMSE was reduced from 2.4×10^{-15} to $1.7 \times 10^{-15} \frac{g}{cm^3}$ for CHAMP and from 1.4×10^{-16} to $9.6 \times 10^{-17} \frac{g}{cm^3}$ for GRACE (experiment 8 in Table 3). The median difference between model and observations is 45% smaller for CHAMP and 20% smaller for GRACE. We believe this approach could be thus beneficial for 'transplanting' the high accuracy of in situ neutral density observations to other satellites at similar altitudes. It could also be used by modelers to improve the representation of processes and boundary conditions.

We found that large jumps in the density time series were introduced by the analysis steps. Here, the model moves away fast from the analyzed state, since the model parameters controlling the dynamics of the model are not updated at the analysis step. We suspect, that the jumps can be handled by adding additional entries to the state vector for estimating corrections for TIE-GCM parameters. For example, cooling rates, eddy diffusion, or reaction rates. Estimating the dynamics of the TIE-GCM in that way could produce solutions that outperform the calibrated NRLMSIS 2.0. Another idea for future research is applying the two-step approach to CHAMP and GRACE simultaneously.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s40623-022-01733-z.

Additional file 1. Accelerometer derived densities CHAMP March 2010.

Additional file 2. Accelerometer derived densities CHAMP April 2010.

Additional file 3. Accelerometer derived densities GRACE March 2010.

Additional file 4. Accelerometer derived densities GRACE April 2010.

Additional file 5. Calibrated NRLMSIS 2.0 GRACE.

Additional file 6. Assimilation experiment 8 GRACE.

Additional file 7. Open loop simulation 3 GRACE.

Additional file 8. Calibrated NRLMSIS 2.0 CHAMP.

Additional file 9. Assimilation experiment 8 CHAMP.

Additional file 10. Open loop simulation 3 CHAMP.

Declarations

References
Bowman B, et al (2008) A new empirical thermospheric density model JB2008 using new solar and geomagnetic indices. In: AIAA/AAS astrodynamics specialist conference and exhibit. https://doi.org/10.2514/6.2008-6438.

Bruinsma S, Boniface C (2021) The operational and research DTM-2020 thermosphere models. J Space Weather Space Clim 11:47. https://doi.org/10.1051/swsc/2021032 (**ISSN: 2115-7251**)

Bruinsma SL, Doornbos E, Bowman BR (2014) Validation of GOCE densities and evaluation of thermosphere models. Adv Space Res 54(4):576–585. https://doi.org/10.1016/j.asr.2014.04.008 (**ISSN: 0273-1177**)

Bruinsma SL et al (2012) Evaluation of the DTM-2009 thermosphere model for benchmarking purposes. J Space Weather Space Clim 2:A04. https://doi.org/10.1051/swsc/2012005

Burgers G, van Leeuwen PJ, Evensen G (1998) Analysis scheme in the ensemble Kalman filter. Mon Weather Rev 126(6): 1719–1724. https://doi.org/10.1175/1520-0493(1998)126<1719:ASITEK>2.0.CO;2. https://journals.ametsoc.org/view/journals/mwre/126/6/1520-0493_1998_126_1719_asitek_2.0.co_2.xml

Codrescu SM, Codrescu MV, Fedrizzi M (2018) An ensemble Kalman filter for the thermosphere-ionosphere. Space Weather 16(1):57–68. https://doi.org/10.1002/2017SW001752

Doornbos E (2012) Thermospheric density and wind determination from satellite dynamics. Springer Theses. Berlin, Heidelberg: Springer Berlin Heidelberg. ISSN: 978-3-642-25128-3 978-3-642-25129-0. https://doi.org/10.1007/978-3-642-25129-0

Drob DP et al (2008) An Empirical Model of the Earth's Horizontal Wind Fields: HWM07. J Geophys Res Space Phys. https://doi.org/10.1029/2008JA013668 (**ISSN: 2156-2202**)

Emmert JT (2015) Thermospheric mass density: a review. Adv Space Res 56(5):773–824. https://doi.org/10.1016/j.asr.2015.05.038 (**ISBN: 0273-1177.**)

Emmert JT et al (2021) NRLMSIS 2.0: a whole-atmosphere empirical model of temperature and neutral species densities. Earth Space Sci 8(3):e2020EA001321. https://doi.org/10.1029/2020EA001321 (**ISSN: 2333-5084**)

Evensen G (1994) Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. J Geophys Res Oceans 99(C5):10143–10162. https://doi.org/10.1029/94JC00572

Forootan Ehsan et al (2022) Forecasting global and multi-level thermospheric neutral density and ionospheric electron content by tuning models against satellite-based accelerometer measurements. Sci Rep 12:2095. https://doi.org/10.1038/s41598-022-05952-y

Gaposchkin EM, Coster AJ (1990) Evaluation of thermospheric models and the precipitation index for satellite drag. Adv Space Res 10(3):303–309. https://doi.org/10.1016/0273-1177(90)90360-C (**ISSN: 0273-1177**)

Gaspari Gregory, Cohn Stephen E (1999) Construction of correlation functions in two and three dimensions. Q J R Meteorol Soc 125(554):723–757. https://doi.org/10.1002/qj.49712555417 (**ISSN:00359009**)

Hagan ME, Forbes JM (2002) Migrating and nonmigrating diurnal tides in the middle and upper atmosphere excited by tropospheric latent heat release. J Geophys Res Atmos 107.D24:ACL6-1-ACL6-15. https://doi.org/10.1029/2001JD001236

Hamill Thomas M, Whitaker Jeffrey S, Snyder Chris (2001) Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter. Mon Weather Rev 129(11):2776–2790. https://doi.org/10.1175/1520-0493(2001)129<2776:DDFOBE>2.0.CO;2 (**ISSN: 1520-0493, 0027-0644**)

He C et al (2018) Review and comparison of empirical thermospheric mass density models. Prog Aerosp Sci 103:31–51. https://doi.org/10.1016/j.paerosci.2018.10.003 (**ISSN: 0376-0421**)

Hedin AE (1987) MSIS-86 thermospheric model. J Geophys Res Space Phys 92(A5):4649–4662. https://doi.org/10.1029/JA092iA05p04649 (**ISSN: 2156-2202**)

Heelis RA, Lowell JK, Spiro RW (1982) A model of the high-latitude ionospheric convection pattern. J Geophys Res Space Phys 87(A8):6339–6345. https://doi.org/10.1029/JA087iA08p06339 (**ISSN: 2156-2202.**)

Houtekamer PL, Mitchell Herschel L (1998) Data Assimilation Using an Ensemble Kalman Filter Technique. Mon Weather Rev 126(3): 796-811. https://doi.org/10.1175/1520-0493(1998)126<0796:DAUAEK>2.0.CO;2. https://journals.ametsoc.org/view/journals/mwre/126/3/1520-0493_1998_126_0796_dauaek_2.0.co_2.xml

Hunt BR, Kostelich EJ, Szunyogh I (2007) Efficient data assimilation for spatiotemporal chaos: a local ensemble transform Kalman filter. Phys D Nonlinear Phenom 230(1):112–126. https://doi.org/10.1016/j.physd.2006.11.008 (**ISSN: 0167-2789**)

Jones M Jr et al (2014) Impacts of vertically propagating tides on the mean state of the ionosphere-thermosphere system. J Geophys Res Space Phys 119(3):2197–2213. https://doi.org/10.1002/2013JA019744 (**ISSN: 2169-9402**)

Klinkrad H et al (2006) Re-entry prediction and on-ground risk estimation. In: Klinkrad H (ed) Space debris models risk analysis. Springer, Berlin, Heidelberg, pp 241–288. https://doi.org/10.1007/3-540-37674-7_9. ISBN:978-3-540-37674-3.

Lahoz W, Khattatov B, Menard R, eds. (2010) Data assimilation: making sense of observations. Springer, Berlin Heidelberg. https://doi.org/10.1007/978-3-540-74703-1. ISBN: 978-3-540-74702-4

Lee IT et al (2012) Assimilation of FORMOSAT-3/COSMIC electron density profiles into a coupled thermosphere/ionosphere model using ensemble Kalman filtering. J Geophys Res Space Phys. https://doi.org/10.1029/2012JA017700

Liu H-L et al (2018) Development and validation of the whole atmosphere community climate model with thermosphere and ionosphere extension (WACCM-X 2.0). J Adv Model Earth Syst 10(2):381–402. https://doi.org/10.1002/2017MS001232 (**ISSN: 1942-2466**)

Longuski JM, Hoots FR, Pollock GE. IV (2022) Perturbations due to atmospheric drag. In: Introduction to orbital perturbations. Springer International Publishing, Cham, pp. 277-301. https://doi.org/10.1007/978-3-030-89758-1_12. ISBN: 978-3-030-89758-1

Lu G et al (2014) Global ionospheric and thermospheric response to the 5 April 2010 geomagnetic storm: an integrated data-model investigation. J Geophys Res Space Phys 119(12):10358–10375. https://doi.org/10.1002/2014JA020555 (**ISSN: 2169-9402**)

Matsuo T, Lee IT, Anderson JL (2013) Thermospheric mass density specification using an ensemble Kalman filter. J Geophys Res Space Phys 118(3):1339–1350. https://doi.org/10.1002/jgra.50162

Matzka J et al (2021) The geomagnetic Kp index and derived indices of geomagnetic activity. Space Weather 19(5):e2020SW002641. https://doi.org/10.1029/2020SW002641 (**ISSN: 1542-7390**)

Montenbruck O, Gill E (2005) Satellite orbits: models, methods and applications. Corr., 3.print. Berlin u.a.: Springer. ISBN: 354067280X, 9783540672807. https://doi.org/10.1007/978-3-642-58351-3

Morozov AV et al (2013) Data assimilation and driver estimation for the global ionosphere-thermosphere model using the ensemble adjustment Kalman filter. J Atmos Solar Terres Phys 104:126–136. https://doi.org/10.1016/j.jastp.2013.08.016 (**ISSN: 1364-6826**)

Nerger L, Tang Q, Mu L (2020) Efficient ensemble data assimilation for coupled models with the parallel data assimilation framework: example of AWI-CM (AWI-CM-PDAF 1.0). Geosci Model Dev 13(9):4305–4321. https://doi.org/10.5194/gmd-13-4305-2020 (**ISSN: 1991-959X**)

Nerger L et al (2006) Using sea-level data to constrain a finite-element primitive-equation ocean model with a local SEIK filter. Ocean Dyn 56(5):634–649. https://doi.org/10.1007/s10236-006-0083-0 (**ISSN: 1616-7228**)

Nerger L et al (2012) A regulated localization scheme for ensemble-based Kalman filters. Q J R Meteorol Soc 138(664):802–812. https://doi.org/10.1002/qj.945 (**ISSN: 1477-870X**)

Nerger L et al (2012) A unification of ensemble square root Kalman filters. Mon Weather Rev 140(7):2335–2345. https://doi.org/10.1175/MWR-D-11-00102.1

Panzetta F et al (2019) Towards thermospheric density estimation from SLR observations of LEO satellites: a case study with ANDE-pollux satellite. J Geod 93(3):353–368. https://doi.org/10.1007/s00190-018-1165-8 (**ISSN: 1432-1394**)

Papitashvili NE, King JH (2020) OMNI 1-min data. https://doi.org/10.48322/45bb-8792.

Picone JM et al (2002) NRLMSISE-00 empirical model of the atmosphere: statistical comparisons and scientific issues. J Geophys Res Space Phys 107(A12):SIA-15-1-SIA—15-16. https://doi.org/10.1029/2002JA009430

Qian L, et al (2014) The NCAR TIE-GCM: a community model of the coupled thermosphere/ionosphere system. In: Huba J, Schunk R, George K, eds. Modeling the ionosphere-thermosphere system. First. Geophysical monograph series. American Geophysical Union (AGU), pp 73-83. https://doi.org/10.1002/9781118704417.

Reigber Ch, Lühr H, Schwintzer P (2002) CHAMP mission status. Adv Space Res 30(2):129–134. https://doi.org/10.1016/S0273-1177(02)00276-4 (**ISSN: 0273-1177**)

Richards PG, Fennelly JA, Torr DG (1994) EUVAC: a solar EUV flux model for aeronomic calculations. J Geophys Res Space Phys 99(A5):8981–8992. https://doi.org/10.1029/94JA00518 (**ISSN: 2156-2202**)

Ridley AJ, Deng Y, Tóth G (2006) The global ionosphere-thermosphere model. J Atmos Solar Terr Phys 68(8):839–864. https://doi.org/10.1016/j.jastp.2006.01.008 (. **ISSN: 1364-6826**)

Sheng C et al (2017) Thermospheric recovery during the 5 April 2010 geomagnetic storm. J Geophys Res Space Phys 122(4):4588–4599. https://doi.org/10.1002/2016JA023520 (**ISSN: 2169-9402**)

Solomentsev DV et al (2012) Ionosphere state and parameter estimation using the Ensemble Square Root Filter and the global three-dimensional first-principle model. Space Weather. https://doi.org/10.1029/2012SW000777

Tapley BD et al (2007) Neutral density measurements from the gravity recovery and climate experiment accelerometers. J Spacecrt Rockets 44(6):1220–1225. https://doi.org/10.2514/1.28843

Tapping KF (2013) The 10.7 cm solar radio ux (F10.7). Space Weather 11(7):394–406. https://doi.org/10.1002/swe.20064

Tippett MK, et al (2003) Ensemble Square Root Filters. Mon Weather Rev 131(7): 1485-1490. https://doi.org/10.1175/1520-0493(2003)131<1485:ESRF>2.0.CO;2. https://journals.ametsoc.org/view/journals/mwre/131/7/1520-0493_2003_131_1485_esrf_2.0.co_2.xml

Vallado DA, Finkleman D (2014) A critical assessment of satellite drag and atmospheric density modeling. Acta Astronaut 95:141–165. https://doi.org/10.1016/j.actaastro.2013.10.005 (**ISSN: 0094-5765**)

Vetra-Carvalho S et al (2018) State-of-the-art stochastic data assimilation methods for high-dimensional non-Gaussian problems. Tellus A Dyn Meteorol Oceanogr 70(1):1–43. https://doi.org/10.1080/16000870.2018.1445364

Vielberg K et al (2018) Comparison of accelerometer data calibration methods used in thermospheric neutral density estimation. Ann Geophys 36(3):761–779. https://doi.org/10.5194/angeo-36-761-2018

Vielberg K, Kusche J (2020) Extended forward and inverse modeling of radiation pressure accelerations for LEO satellites. J Geodesy 94(4):43. https://doi.org/10.1007/s00190-020-01368-6 (**ISSN: 1432-1394**)

Vielberg K, et al (2021) TND-IGG RL01: Thermospheric neutral density from accelerometer measurements of GRACE. CHAMP and Swarm data set. https://doi.org/10.1594/PANGAEA.931347

Walterscheid RL (1989) Solar cycle effects on the upper atmosphere—implications for satellite drag. J Spacecr Rockets 26(6):439–444. https://doi.org/10.2514/3.26089 (**ISSN: 0022-4650, 1533-6794**)

Weimer DR (2005) Improved Ionospheric Electrodynamic Models and Application to Calculating Joule Heating Rates. J Geophys Res Space Phys. https://doi.org/10.1029/2004JA010884

Whitaker JS, Hamill TM (2002) Ensemble data assimilation without perturbed observations. Mon Weather Rev 130(7):1913–1924. https://doi.org/10.1175/1520-0493(2002)130<1913:EDAWPO>2.0.CO;2

Zeitler L et al (2021) Scale factors of the thermospheric density: a comparison of satellite laser ranging and accelerometer solutions. J Geophys Res Space Phys 126(12):e2021JA029708. https://doi.org/10.1029/2021JA029708 (**ISSN: 2169-9402**)

Publisher's Note