Thermosphere and satellite drag

Sean Bruinsma, Thierry Dudok de Wit, Tim Fuller-Rowell, Katherine Garcia-Sage, Piyush Mehta, Fabian Schiemenz, Yuri Y Shprits, Ruggero Vasile, Jia Yue, Sean Elvidge

Available online at www.sciencedirect.com

ScienceDirect

Advances in Space Research xxx (xxxx) xxx

ADVANCES IN SPACE RESEARCH (a COSPAR publication)

www.elsevier.com/locate/asr

Thermosphere and satellite drag

Sean Bruinsma [a,*], Thierry Dudok de Wit [b,c], Tim Fuller-Rowell [d], Katherine Garcia-Sage [e], Piyush Mehta [f], Fabian Schiemenz [g], Yuri Y. Shprits [h,i], Ruggero Vasile [h], Jia Yue [j], Sean Elvidge [k]

[a] GET/CNES, Space Geodesy Office, 18 Avenue E. Belin, 31401 Toulouse cedex 4, France
[b] LPC2E, University of Orléans, 3A avenue de la Recherche Scientifique, 45071 Orléans cedex 2, France
[c] International Space Science Institute (ISSI), Hallerstrasse 6, 3012 Bern, Switzerland
[d] CIRES University of Colorado and NOAA Space Weather Prediction Center, 325 Broadway, Boulder, CO 80302, USA
[e] NASA GSFC, 8800 Greenbelt Rd, Greenbelt, MD 20771, USA
[f] West Virginia University, Department of Mechanical and Aerospace Engineering, Morgantown, WV 26506-6106, USA
[g] Airbus Defence and Space GmbH, Claude-Dornier-Straße, 88090 Immenstaad, Germany
[h] GFZ German Research Centre For Geosciences, Potsdam, Germany
[i] UCLA, Earth Planetary and Space Sciences, Los Angeles, CA 90095-1565, USA
[j] NASA GSFC, Greenbelt, MD 20771, and Catholic University of America, 620 Michigan Ave., N.E., Washington, DC 20064, USA
[k] Space Environment Research (SERENE), University of Birmingham, Edgbaston, Birmingham B15 2TT, UK

Received 23 January 2023; received in revised form 2 May 2023; accepted 5 May 2023

Abstract

Accurate forecasts of thermosphere densities, realistic calculation of aerodynamic drag, and propagation of the uncertainty on the predicted orbit positions are required for conjunction analysis and collision avoidance decision making. The main focus of the Committee on Space Research (COSPAR) International Space Weather Action Teams (ISWAT) involved in atmosphere variability studies is satellite drag, and this paper reviews our current capabilities and lists recommendations.

The uncertainty in the density of thermosphere models is due to the combined effect of employing simplified or incomplete algorithms, inconsistent and sparse density data, incomplete drivers for upper atmosphere heating processes (proxies for solar and geomagnetic activity), and forecast error of said drivers. When calculating drag, the uncertainty is amplified due to the satellite shape and aerodynamic model. The sources of uncertainty are reviewed in this paper, and possible and promising ways forward are proposed. Data assimilation models/approaches have demonstrated superior skill in reproducing the thermosphere's state and are the most promising way forward. However, data to drive the models is generally lacking, and they require significant computational resources.

Substantial progress can only be made by means of setting up a full-blown observing system, including not only density and composition measurements, but equally the necessary model drivers.

Keywords: Satellite drag; Thermosphere model; Thermosphere density data; Thermosphere model drivers

1. Introduction

Currently, the number of objects in space is growing exponentially, notably due to the construction of mega constellations like Starlink, and the number of probable collisions increases even faster (Boley and Byers, 2021).

* Corresponding author.
E-mail address: sean.bruinsma@cnes.fr (S. Bruinsma).

https://doi.org/10.1016/j.asr.2023.05.011
0273-1177/

Please cite this article as: S. Bruinsma, T. Dudok de Wit, T. Fuller-Rowell et al., Thermosphere and satellite drag, Advances in Space Research, https://doi.org/10.1016/j.asr.2023.05.011

More accurate conjunction analysis than currently accomplished is required to reduce the number of collision warnings, and assess the necessity to plan evasive maneuvers (Berger et al. 2020). The main uncertainty in aerodynamic drag calculation and forecasting of objects in low Earth orbit (LEO) is due to the highly variable, in both space and time, neutral upper atmosphere. The 3 February 2022 SpaceX launch of 49 Starlink satellites, of which 38 were lost shortly after due to a minor geomagnetic storm, is the most recent example of what the uncertainty can result in (Hapgood et al. 2022; Fang et al. 2022). The LEO altitude range is approximately from 100–1500 km. It is called the thermosphere where it is collisionally-dominated up to about 600 km, and the exosphere above; in the following we refer to the LEO altitude range as the thermosphere. The Committee on Space Research (COSPAR) International Space Weather Action Teams (ISWAT) has the objective of addressing key problems in space weather research and forecasting. The main focus of the atmosphere variability cluster is satellite drag, and this paper reviews our current modelling and forecasting capabilities, and provides recommendations.

The basic equation to calculate aerodynamic drag is:

$$a_{drag} = \frac{1}{2}\rho C_D \frac{A_{ram}}{m} v_r^2 \quad (1)$$

where a_{drag} is the drag acceleration, ρ is the total mass density predicted with a thermosphere model, v_r is the relative velocity of the satellite with respect to the co-rotating atmosphere (winds are in general neglected), A_{ram} is the projected area in the ram direction, m is the mass of the spacecraft, and C_D is the aerodynamic coefficient. The components and data necessary for drag calculation in an orbit determination program are shown in Fig. 1.

The uncertainty in modeled thermosphere density has several origins, and the drag calculation and forecast total error is the sum of errors or inconsistencies in the:

1. thermosphere model estimates of total mass density and composition;
2. solar and geomagnetic activity proxies and indices, and their forecasts, which are the drivers of the thermosphere model;
3. uncertainty in the drag coefficient and its variation with neutral atmosphere composition and temperature, and the spacecraft's inertial attitude;
4. unrealistic propagation of uncertainty, or none, in the orbit computation.

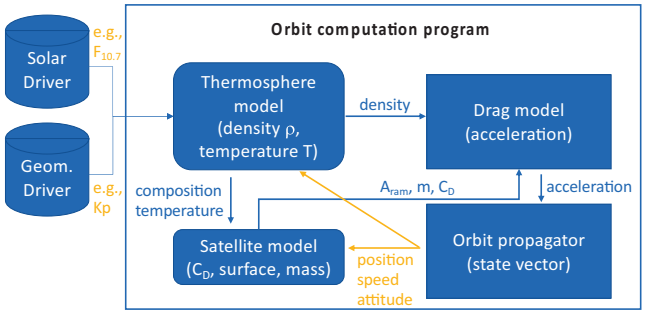

Fig. 1. Diagram showing the data and models of a drag calculation.

In current operational drag calculations, empirical, as opposed to physics-based models (aka general circulation models (GCM)), thermosphere models are employed because of their ease of use, and robust and fast algorithms. Such models provide climatological (i.e. average) predictions which have both low spatial and temporal resolution. Predicted parameters include temperature, total and partial densities of the main constituents as a function of location (altitude, latitude, longitude, local solar time), solar and geomagnetic activity, and season. Empirical thermosphere models are constructed by optimally fitting model coefficients to different combinations of density, temperature, and composition measurements. The densities are reproduced with varying precision as a function of e.g. altitude, solar and geomagnetic activity. Recently, two updates of CIRA (COSPAR International Reference Atmosphere) models have been released, NRLMSIS 2.0 (Emmert et al. 2020) and DTM2020 (Bruinsma and Boniface, 2021). To better fit new or reprocessed density data, both new models predict lower densities than JB2008 (Bowman et al. 2008), DTM2013 (Bruinsma, 2015) and NRLMSISE-00 (Picone et al. 2002). The lower densities are the result of fitting to different datasets than those used in the previous releases. It is important to note here that the bulk of thermosphere composition and temperature measurements were taken in the 1970s and 1980s with mass spectrometers on the Atmosphere Explorers (Pelz et al. 1973; Nier et al. 1973) and the Dynamics Explorer 2 (Carignan et al. 1981).

Typically, ground-based proxies such as the 10.7 cm radio flux (F10.7) and the 30 cm radio flux (F30) are used as model drivers of solar activity, and Kp or ap of geomagnetic activity, for thermosphere models. These proxies are used to represent the upper atmosphere heating by solar Extreme UV (EUV) and UV emissions as well as Joule heating and particle precipitation. EUV measurements from space are also used, notably the S10 index in JB2008, but calibration issues complicate their use (Vourlidas and Bruinsma, 2018). First principles models, or GCMs, often use the solar energy deposition scheme described by Solomon and Qian (2005), but that model actually also relies on F10.7 as the driving index. The most used driver for geomagnetic activity in empirical models is the planetary index Kp or ap, but the Dst (disturbance storm-time) index is also used specifically for storms, e.g., in JB2008. A newly developed index is the Hpo or apo family, which has higher cadence than Kp and is no longer capped at 9 (Yamazaki et al. 2022). Fig. 2 shows the solar indices F10.7, F30 and S81c (top) starting 2000, and the geomagnetic indices ap, ap60 and Dst for the Halloween 2003 storm (bottom). The S81c solar index is used in the

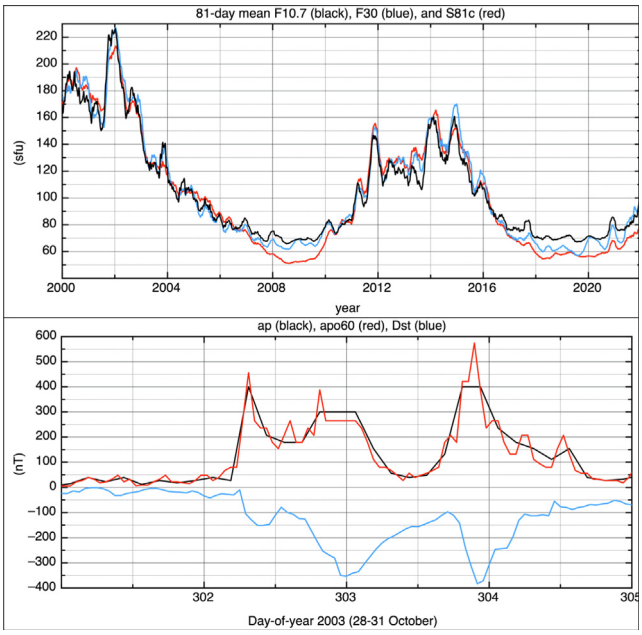

Fig. 2. The solar indices F10.7, F30, and S81c from JB2008 (top), and the geomagnetic indices ap, the new open-ended hourly index apo60, and Dst for the 2003 Halloween storm (bottom).

JB2008 model. Differences in the index amplitudes, notably F10.7 during solar minimum and ap60, or due to resolution are clearly visible. GCMs can also be used to estimate neutral mass density for drag but to date they have not been used in operational calculations. GCMs can employ more sophisticated drivers for geomagnetic activity, such as Weimer's electric field model (Weimer, 2005), Heelis' electric field model (Heelis et al. 1982) or AMIE (Assimilative Mapping of Ionospheric Electrodynamics; Richmond, 1992). AMIE is fit to a wide range of space- and ground-based electrodynamics observations, the Weimer model uses observed solar wind velocity and density as well as the interplanetary magnetic field (Bz and By components) at the L1 Lagrange point between the Earth and Sun, whereas the Heelis model uses 3-hourly Kp. The uncertainty in density prediction due to the drivers in empirical or physical models is two-fold. In addition to using proxies for the variable energy input received from the Sun, these errors grow with increased forecasting lead times (Emmert et al. 2017; Licata et al. 2020b). Currently, solar activity proxies can be predicted several days to approximately one week out, but longer forecast horizons require assimilating and propagating observations from beyond the visible limb of the Sun. In operational settings, the predictions often are typically just autoregression fits of the previous five solar rotations' data, which is not very accurate, as efforts that compare these predictions to issued data have shown (e.g., Vallado and Finkleman, 2014). Predicting the strength of geomagnetic storms accurately can only be done for four to six hours ahead because solar wind and Bz measurements at L1 gives very little lead time (Baumann and McCloskey, 2021).

The satellite shape, surface material characteristics, and mass can also contribute significantly to the uncertainty in the drag calculation. In many orbit computations, the satellite is often represented as a simple geometrical shape made of a small number of flat plates, a so-called macro-model, or even as a sphere with equivalent ram surface. Fig. 3 gives examples of the simplest to the most detailed satellite modeling. To compute drag, the ram surface with respect to the co-rotating atmosphere has to be calculated accurately, and therefore the attitude of the object in inertial space has to be known. This is not always possible, and specifically not in case of debris. Mass is also mostly unknown in case of debris. Another more serious error source is the drag coefficient model, which represents how the satellite surfaces interact with the ambient neutral atmosphere. The drag force can vary by more than 30% depending on the model selected, for which presently no standard is accepted (Mehta et al. 2022). In operational orbit computation and prediction often a single parameter is estimated instead, the ballistic coefficient (m/C_DA; sometimes its inverse), which absorbs the combined error in surface, C_D and mass. This is also true for active satellites, for which physical models and material properties are available, because it is often found to be more accurate simply to fit the ballistic coefficient as part of the orbit determination than to try to model the drag explicitly.

Finally, the uncertainty of the thermosphere densities is not presently taken into account in routine operations, despite the fact that various types of methods have been developed in recent years to incorporate such data. Significant effort is also being expended for providing uncertainty estimates with density predictions (Boniface and Bruinsma, 2021; Licata et al. 2022a, 2022b; Licata and Mehta 2022b). Fig. 4 shows that neglecting force model uncertainties in LEO has a large impact on the extrapolated orbital uncertainty – including the temporal validity of the Gaussian uncertainty assumption often found in routine operations.

In Section 2, the sources of uncertainty in the calculation and the forecast of drag are described and discussed, finishing with operational concerns. The conclusions and recommendations are given in Section 3.

2. Sources of uncertainty in drag calculation and forecasting

2.1. Upper atmosphere models

There are three categories of upper atmosphere thermosphere models, low (spatial and temporal) resolution empirical including machine-learning models (e.g., NRLMSIS-2.0, DTM2020, JB2008, HASDM-ML (Licata et al. 2022a)), high-resolution physics-based models (e.g., CTIPe, TIE-GCM, WAM, GITM, WACCM-X, GAIA), and those which use data assimilation (DA) to fuse observations with background models (e.g. HASDM (Storz et al. 2005) and AENeAS (Elvidge and Angling, 2019)). As a typical example, Fig. 5 presents model density maps

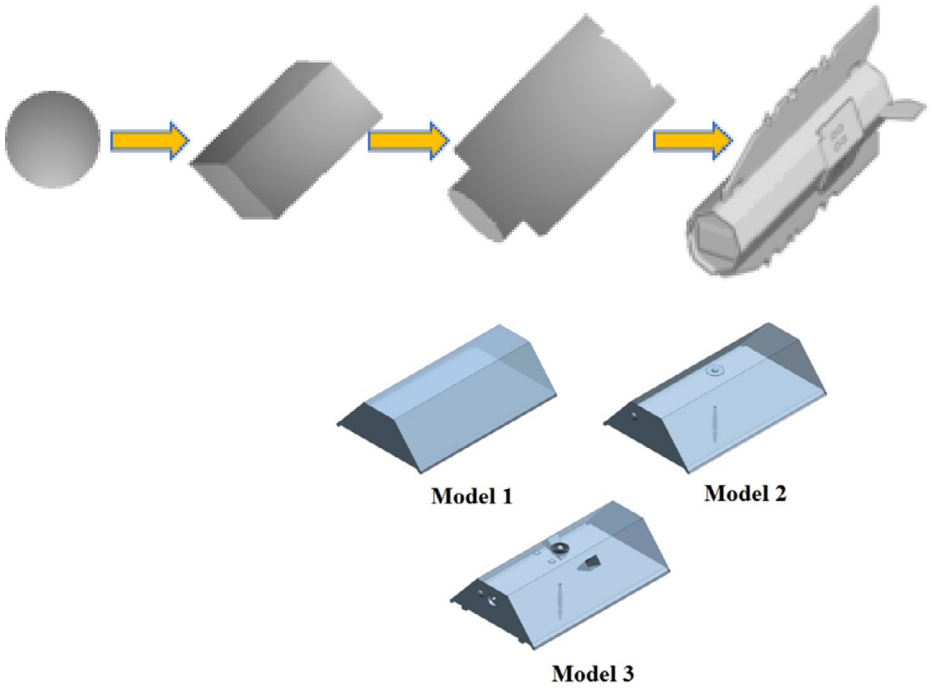

Fig. 3. Modeling of satellite shape, from sphere to detailed model (top). When inferring density, satellite shape differences are not so extreme, and realistic examples for GRACE are given (bottom).

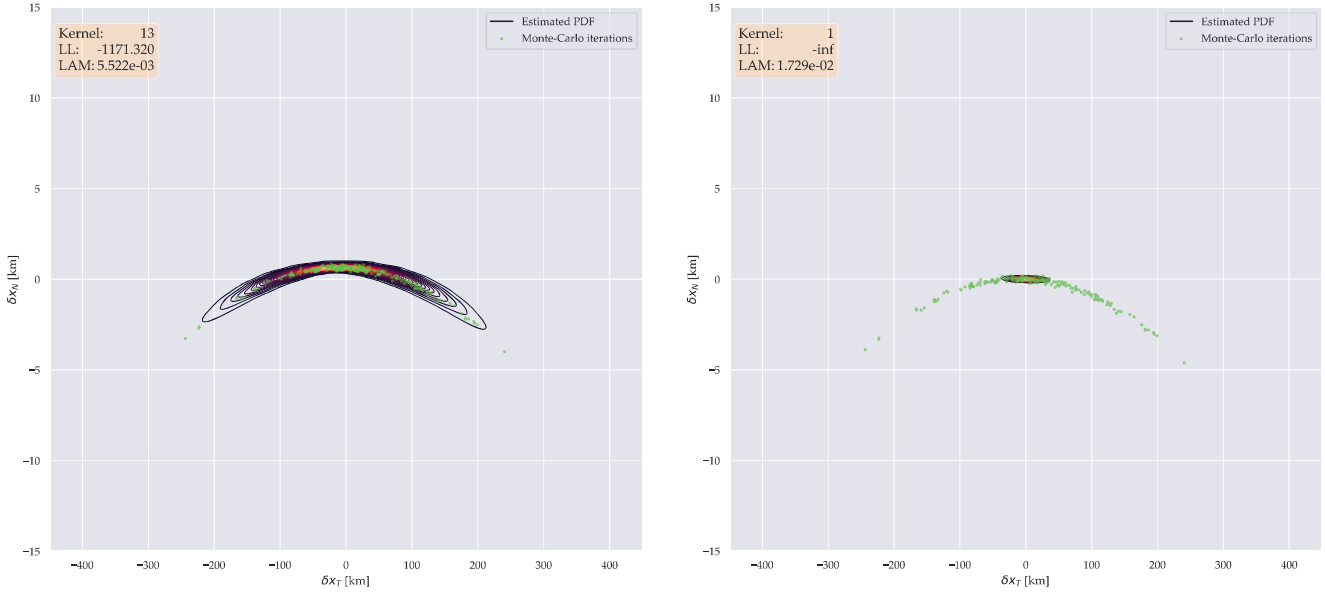

Fig. 4. Typical orbit error samples and corresponding probability density estimates after 10 days of propagation for a LEO-type spacecraft. Left: accounting for density uncertainty (input and model uncertainty), right: neglecting it. The force-model based consideration of the density uncertainty in the estimation process is able to produce probability density estimates which match well with the distribution of Monte-Carlo samples, whereas the classical approach of neglecting force-model uncertainties yields unrealistic and highly optimistic uncertainty estimates (). Source: Schiemenz et al. 2020c

at 250 km and densities just before and at the height of the 17 March 2013 storm. The physics-based models are able to follow response and recovery timescales in response to solar and geomagnetic activity and are similar in their precision for neutral density prediction. Some of the physics-based models now include the influence of variability from the lower atmosphere weather, but the impact on thermosphere density is small except below approximately 150 km altitude (Yue et al., 2022). Although the empirical models are the only ones currently used in operational application for orbit prediction and space traffic management, operational physics-based models could be used in the near future, if after assessment they prove to be more accurate in prediction than empirical models. For instance,

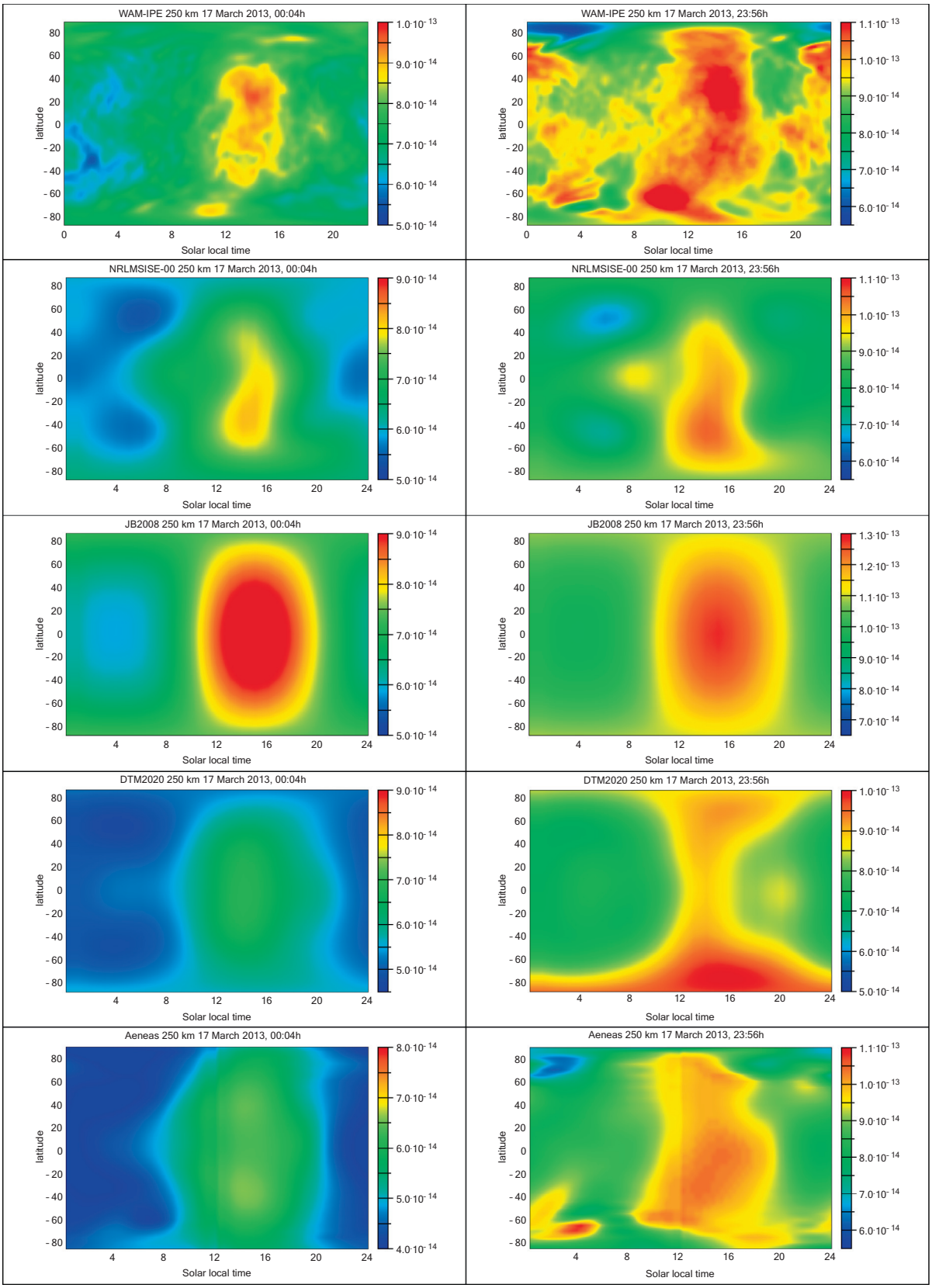

Fig. 5. Model density maps at 250 km just before (left frames) and at the height of the 13 March 2013 minor storm (right frames).

WAM (Akmaev et al. 2008; Fuller-Rowell et al. 2008) coupled to the IPE ionosphere (Maruyama et al. 2016), now runs operationally, in real time, at NOAA and the model neutral density estimates could be used in orbit prediction. The physical model does not need to be run for each satellite or debris object, but the same model 4D data cube (in space and time) can be sampled for each object. This was recently tested to be more computationally efficient than empirical model applications when orbits for thousands of objects were computed. However, practical problems must be addressed, such as dependable and secure physical model data access, algorithms must be used in order to raise the model top to at least 1000 km, the issue of model uncertainty (notably driver uncertainty) must be addressed, and then modification of orbit computation software is necessary.

The fidelity of thermosphere models is unequal under different solar weather conditions and is highest during quiet solar and geomagnetic conditions. Using proxies as model drivers for solar emissions in the EUV and geomagnetic activity leads to errors on time scales of hours to weeks. Due to the very local and sometimes intense variations at high latitude, errors are due to the coarse algorithm in case of empirical models. In case of first physics-based models, errors are due to a combination of (high but still) insufficient resolution and incomplete or inaccurate specification of the drivers (e.g. using planetary indices).

Models used in operational orbit determination software must be easy to implement and run. Therefore, empirical models, or empirical with DA (HASDM, but classified US access only) are used. They require the satellite position (altitude, latitude, longitude, local solar time), solar and geomagnetic activity in the form of indices, and season as input, and return low spatial (4500–9000 km) and temporal resolution climatological predictions of temperature, total and partial densities of the main constituents. Temperature and composition are used in some orbit determination software to compute the aerodynamic coefficient (c.f. Section 2.5). The total densities are reproduced with varying precision, which, due to the climatological nature of the models, improves over longer time scales. The largest errors are observed, unsurprisingly, during geomagnetic storms. Relative errors and uncertainties also increase with altitude, due to the Helium-Oxygen transition at 450–550 km under low solar activity, and the absence of high-resolution observations, and notably accelerometer-inferred, above that altitude. Presently, thermosphere models can only be assessed using orbit-mean or daily-mean densities above 500 km, which includes therefore the 700–800 km altitude range where many Earth Observation satellites operate. Empirical and physics-based model assessments are presented in Bruinsma et al. (2018), and specifically for geomagnetic storms in Bruinsma et al. (2021).

The uncertainties in physics-based models are similar, with larger uncertainties at the higher altitudes, at higher solar activity, and during strong geomagnetic activity. Drivers of the Joule heating in the physics-based models from geomagnetic activity typically either use solar wind and interplanetary magnetic field observations as input to empirical magnetospheric convection models (e.g., Weimer, 2005) and auroral precipitation models (e.g., Fuller-Rowell and Evans, 1987), or assimilation models such as AMIE (Richmond, 1992). The physics-based models also have an upper boundary at about 600 km where the fluid approximations start to break-down. The density at the top boundary can be extrapolated to higher altitudes (~1000 km), assuming diffusive equilibrium of the species and allowing for the decrease in the gravitational acceleration. Similar assumptions are used in the vertical structure in the empirical models at the higher altitudes. At the higher altitudes, Helium takes over as the dominant neutral species from atomic Oxygen, which must be accounted for in the physics-based model approaches. The transition height of atomic Oxygen to Helium varies greatly with solar and geomagnetic activity. Solar activity also typically is prescribed by F10.7 observations as the index to drive an empirical model of the solar EUV and UV spectrum (e.g., Qian and Solomon, 2012).

Presently, the necessary high-resolution measurements of Joule heating and particle precipitation to accurately drive the models is lacking. Another solution to tackle this problem is by means of a density data assimilation method, e.g., a Kalman filter, and estimate corrections to the drivers. High-quality and high-resolution density data at high latitudes is required in such a process. Several studies demonstrating the potential of this approach have been published on the subject (e.g., Codrescu et al. 2004; Morozov et al. 2013; Murray et al. 2015; Sutton 2018), but the lack of continuous, low latency density measurements impedes its operational application.

Finally, model assessment is currently not done according to a standard procedure and metric. Consequently, comparing the performance of models, or identifying limitations or weaknesses, is complicated. Such a procedure is desired by users in order to select the optimum model for their purposes, and essential for the developers of models, which is why it is being implemented at NASA/CCMC (Community Coordinated Modeling Center).

2.2. Upper atmosphere data

The greatest limitation to improving thermosphere models today is inconsistency between observation datasets, inconsistent quality, and above all the sparse distribution in space and time of upper atmosphere density and composition observations. The SET HASDM densities, which are pseudo-observations resulting from a model with data assimilation, are given every 3 h from 2000 to 2019 on equiangular grids extending from pole to pole [10° latitude × 1hr local time], and from 175 to 825 km altitude in steps of 25 km. This source of data is rather recent. Its precision and resolution are still under study, but several papers (e.g., Tobiska et al, 2021; Licata et al, 2021a,

2021b; Tobiska et al. 2022) have explored the uncertainties of these data that are now used for comparing global density models as well as in the development of reduced-order models. As stated in the previous section, data assimilation methods require a continuous stream of high-quality and high-resolution observations to adjust the model state, and this condition is presently far from being met. Since 2000, high resolution accelerometer-inferred densities from CHAMP (Doornbos, 2011), GRACE (Bruinsma 2015), and GOCE (Bruinsma et al. 2014) have greatly advanced thermosphere research and modeling, but in truth these are data of opportunity. Fig. 6 displays the distribution of the densities inferred from CHAMP, GRACE and GOCE, which shows the sparsity of data for high solar activity, in particular below 400 km (and no data above 550 km). Unfortunately, concurrent composition and temperature measurements were not made. In situ composition measurements have not been taken since the Dynamics Explorer 2 mission in the early 1980 s (Carignan et al. 1981). On the other hand, limb remote sensing measurements of thermosphere composition vertical profiles such as O, O_2 and N_2 were collected by GUVI on TIMED during 2002–2007 (Meier, 2015), followed by ICON (Immel et al., 2018) (not yet available). GOLD can measure O_2 profiles from 120 km to 240 km at two longitudes (33E and 128 W) using stellar occultation after 2018 (Lumpe et al., 2020). So, we are in the delicate situation of having most composition and temperature data, and accurate in situ total density data, but 18 years apart, i.e. almost two solar cycles.

Observations are also necessary for model assessment and validation, which often is impossible simply because of the absence (composition) or paucity of data, notably at low and high altitude, below about 250 km and above 550 km, respectively. The effects of severe and extreme geomagnetic storms on the thermosphere total mass density are not well-known because these are rare events on the one hand (only two storms with Kp = 9 after 2000), and because each storm was monitored by just one or two satellites, i.e. observations are restricted to two altitudes and local time planes.

Finally, a mix of density, temperature, and composition datasets are used to fit coefficients of empirical thermosphere models. Because of inconsistencies between datasets, which cannot always be detected and corrected for, or only partly corrected for, the estimation of model coefficients is done optimally in the least-squares sense. Large discrepancies are relatively easy to detect when several datasets are available, and the incriminated dataset can then either be corrected or rejected, but inconsistency at the level of 5% or less is hard to confidently identify. Moreover, many datasets are unique and can therefore only be screened very indirectly or not at all. Total mass densities since 2000 (i.e. when the modern accelerometer missions began) are described and compared in Bruinsma et al. (2022). Datasets inferred from different missions typically disagree 5–20%, but the numbers vary regarding altitude and solar activity. They are largest at 800 km, where they can reach 45%. Densities inferred from the *same* mission can also be quite different due to the processing, in which the adopted satellite model plays a major role (Mehta et al. 2022). This aspect will be discussed in Section 2.5.

2.3. Solar activity: Measurements and forecasts

One of the two primary drivers of upper atmosphere heating is the absorption of solar radiation in the Soft X-ray (XUV, 0.1–10 nm), Extreme-UV (EUV, 10–121 nm), and Far-UV (FUV, 122–200 nm) bands (Lilensten et al. 2008; Vourlidas and Bruinsma, 2018). These bands are defined in (ISO, 2007). A major issue with the spectrally-resolved solar irradiance is the lack of direct observations,

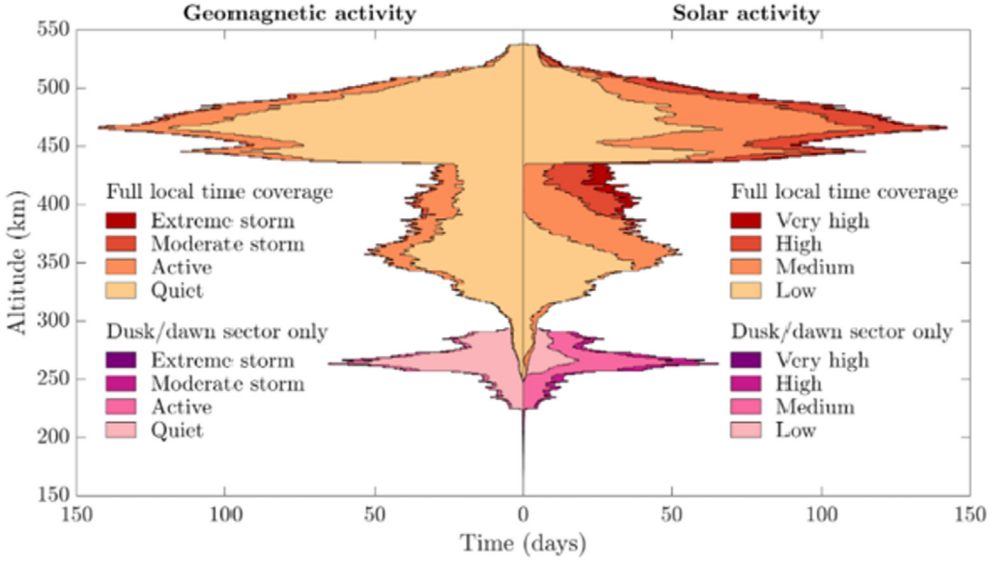

Fig. 6. Distribution of accelerometer density observations from CHAMP, GRACE and GOCE. (Figure used in EOS).

which suffer from a sparse coverage in time and/or in wavelength, and poor accuracy and stability. The radiometric quality of these measurements is affected by the rapid degradation of UV sensors in space, which requires complex mitigation strategies such as redundant sensors, cross-calibration with other missions, or the use of calibration rockets. Although attempts have been made to provide homogenous records of the observed spectral irradiance in the EUV and UV (Haberreiter et al. 2017, Deland et al. 2019), we are still lacking a database of spectral irradiances in the EUV and UV, whose accuracy is good enough to be used for scientific analyses and/or for operational space weather purposes. Empirical or semi-empirical models offer a better alternative (Coddington et al. 2019) although they do not describe flaring activity and most of them are not optimized to describe solar variability below 200 nm. There are some exceptions, however, such as SSPRING (Suess et al. 2016) for the EUV band and FISM for flare irradiance (Chamberlin et al. 2008). Each of these, however, relies on UV observations that come with stability issues.

This lack of direct observations of the XUV-EUV-FUV bands and their lack of radiometric quality very early on favored the use of solar proxies such as the radio flux at 10.7 cm (the F10.7 index) (Schmidtke, 1976). Several more have been developed since, such as S10, M10 and Y10 (Tobiska et al, 2008). While none of these proxies fully describes the complex variability of the XUV-EUV-FUV bands, they offer considerable advantages in terms of continuity, latency, and stability. Today, the three proxies that are most widely used in orbit computation are the F10.7 index (available since 1947; Tapping, 2013), the MgII index (available since 1978; Snow et al. 2019) and the F30 index (available since 1957; Dudok de Wit et al., 2014). The MgII index is derived from an ensemble of space-borne instruments while the radio fluxes are measured from the ground.

While these proxies are provided on a daily basis, they could potentially come with a higher cadence. For satellite operators, a cadence of 6 to 12 h would be more appropriate as this would allow to better track the rapid evolution of complex active regions that also pose a greater space weather risk. For radio observations, this would require a close coordination between several observatories that are distributed in longitude. Although different observatories are presently monitoring the Sun at centimetric wavelengths there is no such common and homogeneous data product. For the MgII index, the major obstacle is the latency, not the cadence.

A recurrent question is whether such solar inputs should include flare-driven irradiance changes, which are highly intermittent but can exceed the background level by orders of magnitude. However, such changes occur on time scales of one hour or less and are integrated in time by the thermosphere, so their impact on the neutral density is often negligible except for large (X-class) flares (Qian et al., 2011; Sutton et al. 2006).

In recent years, much effort has gone into the prediction of these solar indices days to years ahead. Their variability essentially has three contributions. The main one comes from the rotational variability of the Sun and the slow evolution of plages and faculae; this contribution is highly predictable, and even simple recurrence models (today's level is the same as one solar rotation ago) perform well. The second contribution comes from flares, which are unpredictable (Leka et al. 2019) on time scales that are relevant for satellite operators (i.e. up to a month ahead) but whose impact on the thermosphere can be neglected except for X-class flares. The real challenge lies in the third contribution, which is associated with active regions that may potentially emerge and evolve within days. Most prediction models here fall within two categories:

1. Time series models rely only on the past history of the index to predict future values, using linear models (e.g. Warren et al. 2017) or various types of nonlinear models, e.g. (Chatterjee et al. 2001, Yaya et al. 2017, Luo et al. 2020). However, even the most sophisticated models lead to improvements that remain modest with respect to simple linear or recurrence models.
2. The second category of models relies on additional inputs such as the EUV flux on the east limb of the Sun (Lean et al. 2009), or on flux transport model to predict the surface magnetic field, and from this the radiative output of the Sun, e.g. the ADAPT model (Henney et al. 2012; Henney et al. 2015).

Among all approaches, the one based on the surface magnetic field offers the best prospects for predicting solar forcing of the thermosphere up to a month ahead. What we are missing today is a framework for comparing the performance of all these models, similarly to what has been done with geomagnetic indices (Liemohn et al. 2018). More importantly for drag forecast, we are missing a realistic quantification of different sources of uncertainty of all these solar observables.

One concern with all these solar inputs is their long-term availability without interruptions. The MgII and F10.7 indices have so many users that there is a high pressure to maintain their production without interruptions. The production of the F30 index may temporarily stop in a near future as some of the instruments at the Nobeyama facility are facing shutdown, while a new ESA antenna in Poland is not yet operational. S10, M10, and Y10 continue to be produced by Space Environment Technologies since 1997 and are used by several models, including operational ones (Tobiska, 2008). The spectral irradiance in the EUV and UV very much depends on the lifetime of space missions, with presently no coverage anymore in large parts of the EUV spectrum. Finally, magnetograms that serve for models such as ADAPT are presently routinely produced for space weather research, with no guarantee for continuity either. These constraints partly explain why the MgII and

F10.7 indices remain so popular in spite of their limitations.

A different issue is the prediction of solar activity on time scales of years, which is important for mission planning and certain disposal requirements (IADC, 2021) for low-Earth orbit satellites. Many studies have been devoted to the specific problem of predicting the amplitude of future solar cycles, using the sunspot number (Pesnell, 2012; Petrovay, 2020). The prediction of the present cycle, for example, is regularly updated at https://www.swpc.noaa.gov/products/solar-cycle-progression.

Three categories of methods are used here: precursor methods (Kane, 2007) rely on some empirically-derived indicators to estimate the amplitude of the next cycle while extrapolation methods consider the problem from a time-series point of view. The most advanced methods are model-based and rely on dynamo models and surface flux transport models (Karak and Nandy, 2012). The comparison of over 100 predictions of the amplitude of solar cycle 24 (Pesnell, 2016) has revealed a considerable scatter between different approaches, with a large relative standard deviation of about 30%. However, as physics-driven models continuously improve, their predictions are becoming increasingly accurate, with the prospect to reduce the prediction error. An even more challenging problem is the prediction of the level of solar activity more than one cycle ahead. Here too, however, recent progress in dynamo models offer some prospect to bracket the value of the cycle amplitude of up to two cycles ahead (Charbonneau, 2020).

2.4. Geomagnetic activity: Measurements and forecasts

The Kp Index (Bartels, 1949; Mayaud, 1980) is an imperfect but reliable measure of global geomagnetic activity used to drive many upper atmosphere models (c.f. 2.2). It has been recorded since 1932 with three-hour cadence. Kp is calculated by combining measurements of Earth's magnetic field obtained from 13 specific Kp observatories (Rostoker, 1972). The *definitive* Kp index is not available in real time since its calculation requires knowledge of geomagnetic disturbances at stations before and after the actual time (https://www.gfz-potsdam.de/en/kp-index/). Due to its importance in real-time applications though, nowcast and forecast Kp are provided (https://isdc.gfz-potsdam.de/kp-index/). Trivial forecast methods consist in using an average value of Kp ('average Kp'), or are based on the most recent measured Kp ('persistence Kp'). A more physical approach considers Kp of previous solar Carrington rotations ('recurrence Kp'), since these measurements carry information about current conditions and partly account for incoming fast solar wind streams. Especially for short horizon times, direct measurements of solar wind from spacecraft at L1 (SW model) in forecasting Kp are beneficial. Combining these sources of data results in a model that carries information on Kp measurements and SW measurements (Full Model). Data driven models based on similar input data have also been developed (Boberg et al., 2000; Balikhin et al., 2001; Boaghe et al., 2001; Wintoft et al., 2017; Tan et al., 2018; Chakraborty and Morley, 2020). However, it remains unclear how these model predictions compare to each other, for instance in regard to different data sources or dependence on the forecast horizon. An objective approach to evaluate importance of different data inputs is provided in Shprits et al. (2019) using neural networks. The authors use a K-fold cross validation method to ensure that the derived model accuracies are independent of the interval of model validation. Data consists of 1-minute propagated measurements of solar wind and interplanetary magnetic field from the OMNI-Web data service. They performed an 11-fold cross-validation using data from 2005 to 2016 as described above, i.e., using one year of data for validation and the rest for training. The average of the validation errors over all the K-folds is used to estimate the average errors of the model over one solar cycle, which are shown in Fig. 7. The nowcast and short-term forecast performance based on recent solar wind measurements provides the most accurate predictions, but the RMSE error already reaches one after nine hours. Predictions more than 12 h ahead do not benefit from current solar wind measurements. In order to predict such horizons in the future, it is likely that historical Kp measurements and specifically values from the previous solar rotations (recurrence Kp) should be used. Models based on such inputs perform better than those based on current solar wind measurements for above two days ahead forecast (Shprits et al. 2019).

Thermosphere density increases greatly and rapidly under strong geomagnetic conditions. However, forecasting storm-time Kp is complicated due to the skewness of the Kp distribution towards lower values. Consequently, models better predict low geomagnetic conditions, tend to underestimate storms, and perform less accurately under those conditions as displayed in Fig. 8. To improve results for storm events, Shprits et al. (2019) applied data resampling techniques consisting in duplicating high Kp data

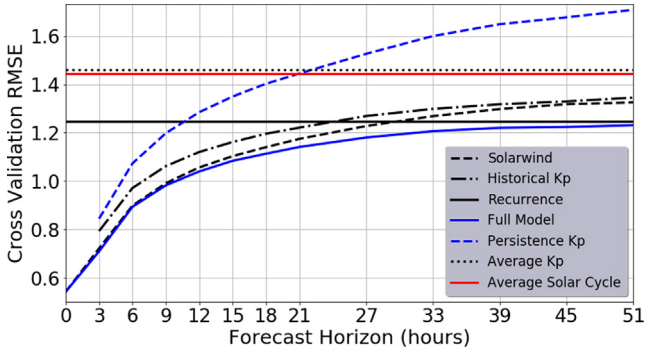

Fig. 7. K-fold Cross Validation root mean square error (RMSE) for the data models described as a function of the forecast horizon time. The solar wind-based model has the lower RMSE than Kp-based models, while the recurrence model outperforms the solar-wind based model for longer horizon times. The full model provides a smooth transition between the recurrence and solar-wind models (). adapted from Shprits et al. 2019

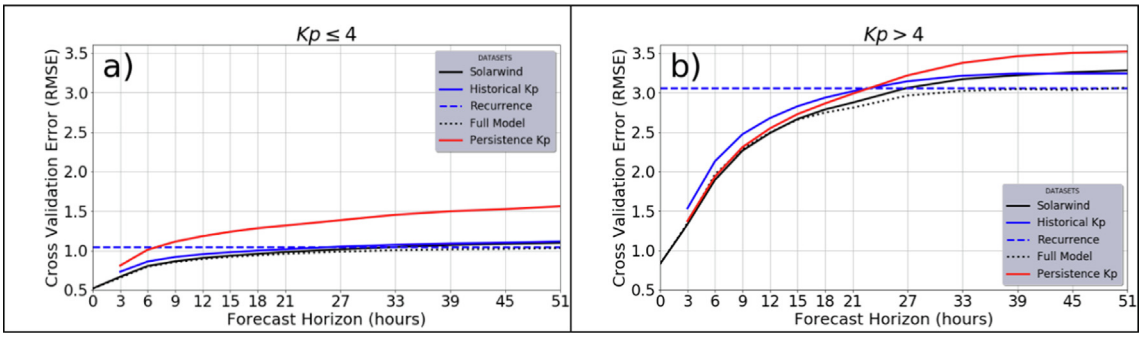

Fig. 8. Accuracy of different data models as a function of forecast horizon for (a) (Kp $\leq$ 4), and for (b) Kp greater than 4 (). adapted from Shprits et al. 2019

points, thus artificially flattening the Kp distribution. The effect is that a neural network trained on such rebalanced data sets will adjust equally to low and high values of Kp, leading to improved forecast for high Kp at the cost of less accuracy for low Kp values. However, Fig. 8b reveals that the RMSE error is already larger than one after three hours and larger than two after 9 h under enhanced geomagnetic conditions.

In Zhelavskaya et al. (2019), different machine learning models such as Linear Regression (LR), Gradient Boosting (GB) and Feed Forward Neural Networks (FNN) have been compared in performance. Moreover, the authors also investigated the possibility of reducing the dimensionality of the input data. This feature selection allows the reduction of dimensions of the training set by extracting the importance of each feature involved, and thereby selecting those which contribute more in the inference process. They applied different feature selection methods, such as Random Forest algorithm (Ho, 1995), a newly developed method based on the Fast Function Extraction (FFX) algorithm, as well as information based methods such as Maximum Relevance Minimum Redundancy (MRMR) (Ding and Peng, 2005; Peng et al., 2005), and Mutual Information Maximization (MIM) (Bollacker and Ghosh, 1996) (see Fig. 1 in Zhelavskaya et al. (2019)). The results shown in Fig. 3 of Zhelavskaya et al. (2019) show that all models performed similarly, when evaluated using two different metrics, the Root Mean Squared Error (RMSE) of the Pearson Correlation Coefficient (CC). These results show that most likely no other regression model can improve the prediction significantly. Improvements can only be made using different informative source data, such as those coming from solar wind predictions from global heliophysics model codes.

The three-hour cadence of the Kp index can be limiting when studying faster variations in geospace, e.g. storm onset. The Hpo index (Yamazaki et al. 2022) is a new Kp-like geomagnetic index showing very similar statistical properties but available at the higher cadence of 30 (Hp30) and 60 (Hp60) minutes, i.e. it is a higher-cadence substitute measurement. Hpo is also open-ended, and values above the intrinsic Kp limit of 9 enable more realistic characterization of severe storms. It is available also in a nowcast real time version at (https://www.gfz-potsdam.de/en/hpo-index/) from 1995 onwards. As for Kp, it is possible to train forecast models of Hpo based on solar wind measurements as well as on historical values. A preliminary operational forecast of Hpo, which achieves similar performance as the Kp forecast, is available at (https://spaceweather.gfz-potsdam.de/products-data/forecasts/hp-forecast).

Physics-based neutral density models typically use observed solar wind velocity and density, and the interplanetary magnetic field (IMF) at L1 to drive empirical magnetospheric convection models (e.g., Weimer, 2005) and auroral precipitation models (e.g., the TIROS-NOAA model of Fuller-Rowell and Evans, 1987). These empirical models are then used to calculate the spatial distribution and temporal variation of Joule and auroral particle heating. However, it is a challenge to forecast these model drivers. Relationships between a Kp forecast and the solar wind and IMF can and are used, for instance in the operational forecast from WAM-IPE at NOAA (Fang et al. 2018; 2022). In principle, the operational solar wind propagation model (e.g., WSA-ENLIL) can forecast the arrival at L1 of the solar wind conditions but predicting the orientation (i.e., southward or northward and serious perturbations in case of the former) of the IMF remains a significant challenge.

2.5. Satellite shape and aerodynamic coefficient modeling

The drag coefficient C_D is a component of a more general aerodynamic coefficient vector and characterizes the scaled drag acceleration (eq. (1) in the direction of satellite motion. In the context of spacecraft dynamics, the drag coefficient is generally characterized as either fixed, fitted, or physical. Fixed drag coefficients use a predetermined value that does not change. Fitted drag coefficients are derived using some form of a fitting or filtering process in the orbit determination process and they are typically updated over time (every few hours or orbits). The physical drag coefficient is computed by modeling the momentum

and energy exchange between the flow-field particles and the satellite. It is a function of various physical parameters: spacecraft shape and materials, attitude, atmospheric temperature and composition, spacecraft surface temperature, velocity of the spacecraft relative to the co-rotating atmosphere, and gas-surface interactions that describe the manner in which energy and momentum exchange takes place. Several studies have been conducted to investigate the sensitivity of the physical drag coefficient to the different parameters (Mehta et al. 2013, 2014a, 2014b;March et al. 2019a; 2019b 2021;Walker et al. 2014; Pilinski et al. 2016; Bernstein and Pilinski, 2022). The physical drag coefficient is most influenced by gas-surface interactions (GSI).

Historically, a fixed value of the drag coefficient, typically 2.2 (and still given for a 'typical spacecraft' in ISO 27852: Space Systems - estimation of orbit lifetime), was most used for both science and operations for its ease of implementation. However, clearly the assumption of constant drag coefficient breaks down very easily. In reality the drag coefficient varies both spatially (due to change in atmospheric composition with location – latitude, longitude, and altitude) and temporally (solar and geomagnetic energy input). As a result, mainstream adoption of fitted drag coefficients started with the development of the High Accuracy Satellite Drag Model (HASDM; Storz et al. 2005) which uses dynamic calibration of atmosphere (DCA) to constantly update the global thermosphere and ballistic coefficients using a set of calibration objects (~70–100) that are (close to) spherical in shape to neglect the effect of shape and attitude. For objects whose physical attributes (shape, mass, and attitude) are not known, as in the case of debris, the unknown parameters are lumped into a single ballistic coefficient ($=m/C_DA$) parameter to more easily consolidate the uncertainties.

HASDM is the operational system of the US Space Force and is not publicly available. With the limited knowledge about HASDM that can be gleaned from the publicly available literature, the forecasting for density for operations reverts to the use of Jacchia-Bowman-HASDM model (JBH09; Hejduk and Snow, 2018) that serves as the baseline for the DCA procedure in HASDM with a driver-based temperature correction, while the ballistic coefficient forecasts use a form of persistence. It is likely that actual operations involve approaches different than those available in the public literature. However, the process can result in unphysical values and there is no standard approach or baseline for estimating drag or ballistic coefficients. While this method may not be an obstacle for internal use, i.e., with the same software using the same thermosphere model, the increasingly dominating commercial space industry does not use the same model or software (HASDM) making adopting it in operations challenging. Additionally, much like the fixed drag coefficient, its use can easily be manipulated in the mission design and lifetime prediction. The 25-year re-entry requirements (IADC, 2021) or the more recent 5-year requirement of the FCC (The United States Federal Communications Commission; https://www.fcc.gov/document/fcc-adopts-new-5-year-rule-deorbiting-satellites-0) can be easily met by simply and freely changing the ballistic coefficient to the desired value.

In an ideal scenario, science and operations would use physically modeled aerodynamic coefficients and the importance of achieving this cannot be overstated. However, there are several limitations and challenges that need to be addressed before this can be achieved. Physically modeling the drag coefficient requires knowledge of several different space object and environmental parameters. The object parameters are velocity, mass, attitude and shape (also providing the cross-sectional area). While these properties maybe well known for (in)active satellites, these are not typically known for debris objects. Estimating the mass and shape of debris objects is an active area of research (e.g. Mehta et al. 2018; Friedman and Frueh, 2022; Baars and Hall 2022).

The space environment parameters are winds, atmospheric temperature and composition, all of which dictate the specifics of energy and momentum exchange between the atmosphere and the spacecraft, otherwise known as gas-surface interaction (GSI). The space environment is highly dynamic, spatially and temporally. The environmental properties and hence GSI can vary significantly during the solar cycle, in sunlight and on nightside, at the equator or near the poles, as well as with altitude. A well-known effect of these dynamics is change in the dominant species from atomic oxygen to helium during solar minimum and with altitude, otherwise known as oxygen-to-helium transition. Understanding and modeling of this transition is also an active area of research (e.g. Mehta et al. 2019; Bernstein and Pilinski, 2022), but more observations are required to fully understand the processes.

Even when the object and environment parameters are perfectly known, significant differences or errors can be imparted by the choice of gas-surface interactions models (up to 20%) and computational methods for estimating the drag coefficient (up to 15%) (Mehta et al. 2022). Options for the GSI model include different reflection kernels (e.g. diffuse or quasi-specular) and accommodation models (Pilinski et al. 2013; Walker et al. 2014). The choice of the computational methods includes analytical panel methods and numerical methods such as Direct Simulation Monte Carlo or Test Particle Monte Carlo (Mehta et al. 2014b; March et al. 2019a; Sinpetru et al., 2022). The choice of computational method is a trade-off between computational cost and fidelity. The response surface methodology (RSM) for modeling drag coefficient provides high fidelity, low computational cost and a user-friendly software (Sheridan et al. 2022). An example of the numerical values resulting from three GSI models for the same spacecraft, GRACE, is computed with RSM and shown in Fig. 9. The differences displayed here are worst cases, close to 500 km altitude and at low solar activity, and between diffuse (red and blue curves) and quasi-specular (green) GSI models with the Walker accommodation model (red and green) or fixed accommodation (blue).

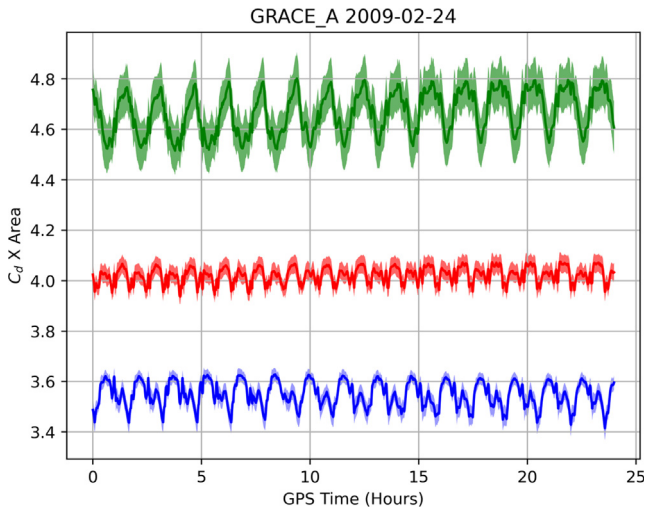

Fig. 9. The drag coefficient for the GRACE-A satellite on 24 February 2009. Shown are C_D using diffuse (red and blue curves) and quasi-specular (green) GSI models with the Walker accommodation model (red and green) or fixed accommodation of 0.82 (blue).

The largest differences of around 30% are between quasi-specular (green) and diffuse with fixed accommodation (blue). For a detailed description of the models we refer to Mehta et al. (2022).

For the last two decades, accelerometer-derived density datasets from the CHAMP/GRACE/GOCE satellites have been used as 'ground truth' for thermosphere science and space operations. Several different dataset versions for each satellite exist, each using a different combination of fidelity of spacecraft macro-model, GSI model, and computational method (e.g. ; Doornbos, 2011; Calabia and Jin, 2016; Mehta et al. 2017a, and March et al. 2019a). These datasets have been used to validate and improve thermosphere models; however, not using the improved models with a consistent methodology for drag coefficient modeling adds errors (biases) in operations.

2.6. Orbit extrapolation and uncertainty propagation

Uncertainty growth during orbit propagation is dominated by two factors: the initial semi-major axis (radial) uncertainty at the start of a propagation phase, and uncertainties in the dynamics. The former is naturally modeled via covariance or probability density propagation in a probabilistic uncertainty model, but usually force model uncertainties are presently neglected. Most space agencies and satellite operators hence employ some form of covariance scaling to mitigate overly optimistic covariance estimates or consider worst-case scenarios for collision probability estimates (Poore et al. 2016). CSpOC (Combined Space Operations Center) additionally employs a consider parameter, which is modeled as additive adjustment to the ballistic coefficient covariance and thus known to lack realism in modelling the true evolution of density uncertainty (Emmert et al. 2017).

The main sources of thermospheric density uncertainty are input uncertainty (sometimes also called driver-uncertainty) and grid-scale model uncertainty. Until recently, the treatment of force model uncertainty quantification (UQ) has received little attention, despite the general awareness that thermospheric density uncertainty is the main driver of orbital uncertainties in LEO. A method of kinematic state vector UQ, which accounts for the accumulating errors in thermospheric density, is needed to achieve orbital uncertainty realism. There are three general strategies to propagate input uncertainties, model uncertainties or combinations thereof to satellite orbits:

1. Analytic equations describe the propagation of density uncertainty to the in-plane mean motion uncertainty. The foundation for this type of research was laid by Emmert et al. (2014), and a rigorous derivation of the mean motion error equations caused by a temporal evolution of relative density errors was provided later in Emmert et al. (2017). The error equations in the latter study are general, whereas the provided variance approximations are valid only for white noise and Brownian motion relative density errors. The study also demonstrated that a Brownian motion process is representative for forecast input uncertainties encountered during orbit extrapolation (in-track variance growth with time t proportional to t^5), whereas a white noise error process can be considered as limiting case for the input uncertainty encountered during orbit determination (i.e. when published space weather proxy information is available; in-track variance growth proportional to t^3). No analytic relationships were provided to quantify the density uncertainty arising from uncertainties in the space weather inputs. First approximations of the relative density uncertainty caused by solar flux input uncertainty are provided in Schiemenz et al. (2019a). A significantly enhanced method to generically compute the relative density uncertainty due to input uncertainties is presented in Schiemenz et al. (2020a). When combined with Emmert et al. (2017), it allows a complete analytic propagation of proxy input uncertainties to the uncertainty in the orbital mean motion. The construction of complete density uncertainty covariance matrices including the derivation of the necessary correlation coefficients is shown in Schiemenz et al. (2019b). Due to the assumption of white noise or Brownian motion error process models, they are only eligible to study the impact of input uncertainties. Analytic equations for the propagation of grid-scale density model uncertainty are treated in Schiemenz et al. (2020b). Finally, Schiemenz et al. (2020c) make use of all these studies to formulate combined covariance estimates, which quantify the total and accumulated effect of input and grid-scale model uncertainty. These process noise estimates are subsequently used in a Gaussian mixture orbit determination algorithm capable of considering non-Gaussian uncertainties by estimating the complete

state vector error probability density function. It is shown that density uncertainty can become the dominant driver of the overall uncertainty growth. Another important finding is that, for long propagation times or when drag is enhanced, it is not correct to assume that kinematic state vector uncertainty and the increase of the orbit uncertainty due to the underlying density uncertainty remain Gaussian. The complete analytic density uncertainty approach can be found in (Schiemenz, 2021).

2. The orbital variability is analyzed by means of numerical orbit propagation and Monte Carlo sampling. Required inputs are probability distributions for driver and/or model uncertainties which are subsequently used for sample generation. Prominent works in this category are Bussy-Virat et al. (2018), Hejduk and Snow (2018) and Licata et al. (2020a). Sampling-based methods allow for probability density estimates free of assumptions with regard to the expected uncertainty volumes. However, they require significant computational resources and are only valid for the particular scenario under investigation. The method is therefore often used to study relationships/trends, or for sensitivity analyses. Real-time UQ, let alone catalogue uncertainty processing on CSpOC scales, is impossible on current hardware.
3. A third methodology, by Mehta and Linares (2017b), consists in the construction of Reduced Order Models (ROM) combined with simultaneous estimation of the kinematic state vector and one or more density parameters. Model order reduction makes use of techniques such as singular value decomposition to reduce the state vector dimension while retaining maximum information. Dynamic models that allow density forecasting are then estimated from historical density and space weather databases using the Dynamic Mode Decomposition with control (DMDc) algorithm or neural networks (Turner et al. 2020). The construction of the corresponding unscented Kalman filter is described in Gondelach and Linares (2020a,b). Gondelach and Linares (2020b) demonstrate that the method can also be used to quantify density uncertainty, as well as its impact on the kinematic state vector. Their analyses of the effect of grid-scale model density uncertainty on the probability of collision (Pc) with a ROM generated from JB2008 using DMDc confirm earlier results by Bussy-Virat et al. (2018) and Hejduk and Snow (2018), which demonstrate that the covariance inflation caused by density uncertainties can lead to both an increase or decrease in collision probability. Licata et al. (2020a) compared the in-track position error distributions due to forecast driver uncertainty with NRLMSISE-00, and two ROMs that were derived from their baseline models (TIE-GCM and HASDM) using machine learning techniques. The effect of driver uncertainty is added in Licata et al. (2021a). The HASDM-ML ROM was recently extended to provide density UQ based on Monte-Carlo dropout (Licata et al. 2021a, 2021b, Licata et al. 2022a). Together with DTM2020 and MSIS-UQ (Licata et al. 2022b) it is one of the three currently available models that provide an uncertainty estimate of the computed density.

2.7. Operational concerns

Operational nowcasts and forecasts of satellite orbit prediction and conjunction analysis require dependable, stable thermospheric density specifications, with reliable data streams as input. This includes not only solar and geomagnetic indices or drivers, but also any data that is assimilated into the density models. As discussed above, it is not only the accuracy of the satellite state predictions that is important, but also the uncertainty, in order to determine the probability of collision (Pc). Balancing these operational needs – accuracy, dependability, code stability, input data reliability, and quantified uncertainties – scientists and operators have come up with a variety of approaches, including assimilative and ensemble modeling approaches, which can allow operators to capitalize on the strengths of different models and combine those with prior knowledge of the system to improve model accuracy and quantify uncertainties.

The operational use of Pc and the relationship of neutral density errors and uncertainties on the calculation of Pc were discussed in detail by Hejduk and Snow (2018). Because the neutral density is the major source of uncertainty for satellite drag, density models that include those uncertainties are desirable for operations. Hejduk and Snow (2018) demonstrated that the inclusion of neutral density uncertainties in conjunction analysis may increase the number of false alarms (predictions of a close conjunction when one does not occur), but they decreased the number of more problematic missed detections of close conjunctions.

Multi-day predictions that include geomagnetic effects are needed in order to give the necessary 3–4 day lead time. This leaves solar and geomagnetic uncertainties as the major source of error for satellite drag operations. Accounting for solar and geomagnetic driver uncertainties is a particularly difficult problem because, as with density predictions, uncertainty quantification for operations should also be stable and consistent in quality. Solar and geomagnetic driver forecasting poses a difficulty across space weather subfields. Magnetospheric space weather studies have demonstrated the usefulness of ensemble modeling for dealing with these uncertainties and quantifying their effects, by perturbing the inputs to account for driver uncertainties (Morley et al. 2018).

Improving the accuracy of the nowcast typically requires a method to assimilate data or debias the density models. These methods require near real-time orbit determination data. The HASDM model and the tracking data of the satellites it relies on to debias the neutral density model (Casali and Barker, 2002; Storz et al. 2005) are

not publicly available, closing off this approach outside of the US Department of Defense and its partners. Ensemble Kalman filter (EnKF) and ensemble optimal interpolation (EnOI) assimilation into both physics-based and empirical thermosphere models has shown promising results (e.g. Matsuo et al. 2012; Murray et al. 2015), but in order for this approach to be operationally useful, a large amount of real-time, publicly- or commercially-available (e.g. orbit determination, mass spectrometer or accelerometer) data is required. The presence of commercial space situational awareness radar operators, such as LeoLabs, now make HASDM-quality tracking data available to non-DoD government and commercial entities.

Additional approaches may involve assimilation of related ionospheric data into physics-based models. These data may be more readily available in real time than the orbit determination data for direct density assimilation, but the assimilation of related parameters can constrain the model inputs or results in order to improve the neutral density specification. This type of approach may involve EnKF assimilation (Codrescu et al. 2004; Elvidge and Angling, 2019) or an iterative approach (Sutton, 2018).

Using multiple models can simultaneously improve accuracy and be used to quantify uncertainty. A weighted multi-model ensemble approach, as described by Elvidge et al. (2016) or Murray (2018), combines the model results using appropriate weighting based on performance metrics. Such a technique can be used to specify neutral density with the appropriate uncertainties. This multi-model ensemble approach can be combined with a data assimilation approach, as prototyped in the Dragster framework (Crowley and Pilinski, 2017; Pilinski et al. 2019), which combines an iterative assimilation approach with weighted multi-model ensembles.

Both iterative assimilation and ensemble approaches appear to be promising ways to specify neutral densities and quantify uncertainties. However, the computational cost of running model ensembles is high, especially if using physics-based models. While this computational cost may not preclude the operational use of such an approach, robust validation is needed to demonstrate the benefit of computationally-expensive approaches and thereby justify the computational costs. Empirical thermosphere models, however, typically have a low computational cost and would benefit greatly from a weighted multi-model ensemble approach.

In order to demonstrate the benefit of investing in a particular modeling framework or a particular real-time data stream, a series of Observing System Experiments (OSEs) and Observing System Simulation Experiments (OSSEs) should be carried out. OSEs demonstrate the effects of assimilating specific datasets into modeling frameworks, in order to determine the impact of a particular dataset on assimilative model results. OSSEs are similar but rely on models to produce synthetic data in order to predict the impact of potential future observing platforms. Conducting OSEs and OSSEs will allow us to ensure the investment in operational observations and models is worthwhile.

3. Conclusions and recommendations

The conclusions and recommendations are given below per topic presented in Section 2.

3.1. Upper atmosphere models

Empirical models of the thermosphere, with or without DA, are used in operational drag calculations. While developments and testing are ongoing, physical models will not replace them in the near future. We recommend:

1. Empirical model development should be continued, although the scarce supply of high-quality total density data in fact limits possible progress (c.f. 3.2).
2. Models based on machine learning, or some hybrid form, may potentially achieve higher precision than the current CIRA models and should be developed and tested.
3. Data assimilation models/approaches have demonstrated superior skill in modelling thermospheric densities in comparison to empirical and physics-based models and are the most promising way forward. Several have already been developed, but presently all excepting HASDM require further development and testing before possible operational use.
4. The use of indirect observations (e.g. electron densities) to update estimates of thermospheric density must be further investigated, and notably how to construct the required accurate covariance matrices.
5. Multi-model ensembles that combine the output of a range of ideally independent models, which can provide an intermediate operational "stopgap" whilst improvements in first-principle modelling, data assimilation techniques and high-rate data become available. Some thermosphere use cases have demonstrated their potential viability (Elvidge et al. 2016), but their use is in its infancy.
6. Implement a standard and objective model assessment procedure for comparing the performance of models and identify limitations.

3.2. Upper atmosphere data

Thermosphere model improvement and thorough assessment are hampered by the limited availability of high-quality total density data (c.f. 3.1), and the near absence of composition measurements. We recommend:

1. Consistency between existing, and future, density datasets must be guaranteed by means of accurate (re-) processing of satellite tracking or accelerometer data employing high fidelity satellite models (c.f. 2.5).

2. An observation system is needed to achieve significant progress in thermosphere modeling. Simultaneous high quality mass density (e.g. GOCE, CHAMP) and composition (O_2, N_2, O, He) data by means of mass spectrometers (e.g. DE2) or the Terahertz microwave technique (Wu et al. 2016) are necessary to revise the composition information in the models.
3. Uninterrupted high-resolution monitoring of geomagnetic storms is needed to augment the current small database of storm events, which severely penalizes both modeling and model assessment. At present, only GRACE-FO (data of opportunity, climate mission: ftp://thermosphere.tudelft.nl) provides high-resolution total mass density data.
4. As the density and composition variation in the thermosphere during storm-time is mainly driven by Joule heating and NO cooling, ideally temperature, composition, wind, and NO cooling measurements in the lower thermosphere from 100 to 200 km will complement the total density measurements and close the loop of physics (i.e., understanding). But observing in situ in that altitude region is very complicated due to the very high atmospheric drag experienced by a satellite, which imposes eccentric orbits and hence limited spatial coverage.
5. On the contrary, inferring densities for altitudes above 600 km is complicated due to the very low drag and consequently data is sparse there too. Satellites with simple shapes (e.g. sphere, cube), thus minimizing errors due to radiation forces (Bruinsma et al. 2022), and equipped with GNSS receivers could at least provide mean densities in support of Earth Observation satellites in the 700–800 km altitude range, and of course also at lower altitudes.

A good start to a future observing system will be NASA's GDC (Global Dynamics Constellation) mission with 6 satellites, which is planned to provide in-situ measurements of low-resolution total mass density, composition, temperature and wind at 300–400 km, along with magnetospheric drivers that contribute to thermospheric heating. (GDC STDT Report Final, https://science.nasa.gov/science-pink/s3fs-public/atoms/files/GDC%20STDT%20Report%20FINAL.pdf).

3.3. Solar activity: Measurements and forecasts

Our recommendations are very similar to those already listed in (Vourlidas and Bruinsma 2018):

1) Maintain the measurement of solar proxies that are widely used for operational use, in particular F10.7, F30, S10, MgII/M10 and Y10. The measurement of solar surface magnetism is equally important because it is a key input for physical models.
2) Provide a common framework allowing to compare the various uncertainties associated with these different solar inputs, what assumptions go into them, and for deriving community-consensus composites, similarly to what has been done with the Total Solar Irradiance (TSI).
3) Expand the transition from solar proxies to EUV irradiance observations for operational use in thermosphere models. Recommendations have already been made as how to reconstruct the EUV spectrum from the measurement of a few lines (Dudok de Wit et al. 2005) or a few spectral bands (Cessateur et al. 2011; Thiemann et al. 2019). The EUVS sensors on the GOES series (Viereck et al. 2007) all observe the same spectral bands for that reason. An international coordination should ensure that these bands will be covered by future missions. Sufficient overlap in time must be taken into account for consecutive missions to allow for cross-calibration, and in-flight calibration must be performed to enable correcting the measurements for instrument degradation. These missions and measurements would ideally also be part of the observing system recommended in 3.2.

Regarding the forecasting of solar inputs, our recommendations are:

1) Continue the uninterrupted measurement of the solar surface magnetic field, which is the primary input for flux surface and for dynamo models, which are respectively needed for short-term and long-term forecasts.
2) Provide a common framework for comparing the many different forecast models that exist: test them on a common time interval or better continuously, and define common metrics of performance.
3) Measuring the solar surface magnetic field in polar regions will dramatically improve the forecasting capacity of dynamo models, but at a price of a fleet of Sun-orbiting spacecraft.

3.4. Geomagnetic activity: Measurements and forecasts

Using solar wind measurements from L1 improves predictions up to 12 h ahead. The results show that most likely no other regression model could improve the prediction significantly. Improvements for longer forecast horizons can only be obtained using different informative source data, such as those coming from solar wind predictions from global heliophysics models. In particular, flux rope modeling in order to correctly propagate the magnetic structure of the CME from the Sun to Earth will provide Bz at Earth. Missions at other Lagrange points (L5, L4) may help. However, it will take years before one will have enough data for a machine learning model to work smoothly.

3.5. Satellite shape and aerodynamic coefficient modeling

We need to move towards making physical modeling of drag coefficient mainstream. There are a few consensus decisions to be made by industry, operators and researchers to achieve this. We recommend:

1. Operators need (to be motivated) to develop and share accurate and high-fidelity models (geometry and optical properties) for all new satellites and missions.
2. The community needs to come to an agreement on the use of gas-surface interaction models and arrive at consistent density data sets for improving and validating thermosphere models.
3. The community needs to identify a baseline method for modeling of physical drag coefficient (e.g., Panel method vs DSMC).
4. The adopted method must be able to provide realistic uncertainty estimates to make space operations more robust.

Until improvements to physical drag coefficient modeling can be achieved, operations are likely to continue using fitted drag or ballistic coefficients. However, care must be taken to make sure that consistency between ballistic coefficients and density models is maintained to avoid unintentional biases. We suggest and hope that efforts to add more physical knowledge through inference of space object characteristics for debris continue and expand.

In case of collision avoidance (CA), however, events are driven by the prediction error of the secondary objects in conjunctions, which are often non-active objects. So fitted ballistic coefficients will remain the norm in CA calculations, and small improvements to primary object prediction error are unlikely to manifest notable operational improvements.

3.6. Orbit extrapolation and uncertainty propagation

The analytic approach is based on error process models. For driver uncertainty propagation, currently only Brownian motion and white noise error process models have been developed. We recommend:

1. Non-Gaussian uncertainty models, e.g. via Gaussian sums, must be developed and tested to take geomagnetic index uncertainties under storm conditions into account. However, this adds considerable complexity to the orbit determination process.
2. The analytic driver uncertainty propagation also needs to be updated based on the new empirical models DTM2020 and NRLMSIS 2.0. Numerical orbit propagation and MC sampling is promising, but not viable in case of many objects on current hardware, and thus only used for specific studies. The ROM-model approach is capable of dynamic error correction on model-predicted density, which is desirable for operational application. The machine learning techniques permit the study of nonlinear dynamics models for density estimation (Turner et al. 2020). However, employing ROM models more than doubles the state vector dimension, leading to increased runtimes, but also e.g. risk of observability issues and large matrix inversion difficulties caused by large conditioning numbers.
3. More studies are necessary to evaluate the compatibility of ROM models with operations.
4. Non-Gaussian UQ at density and kinematic state vector levels must be developed, and orbit determination algorithms must then be updated accordingly, taking current and future tracking capability into account.
5. The available methods for performing orbit UQ due to density uncertainties must be evaluated in terms of accuracy, performance, and ease of implementation in order to encourage and facilitate implementation into those systems safeguarding current and future space assets.

3.7. Operational concerns

The United States shifts to a new approach to space situational awareness (Berger et al., 2020):

1. New systems should include comprehensive space weather modeling with comprehensive and consistent uncertainties, run in ensemble mode. This can enable a "best of both worlds" approach to thermospheric modeling, with highly stable and rapidly-running empirical models providing the backbone to the thermospheric predictions, but with physics-based models providing a crucial check on the empirical models, especially during extreme driving conditions, when statistics are sparse and empirical model uncertainties are highest.
2. Owner/Operator predicted ephemerides are a critical input to SSA, especially to orbital safety assessments, because they are the only durable method to include future planned maneuvers in predicted states; but the primitive (and sometimes non-existent) atmospheric density models used in most owner/operator orbit determination software work to dilute (and often negate) the value of these predicted ephemerides. A standardized, easy-to-integrate atmospheric density module that can be made freely available to satellite operators, perhaps requiring a modestly-priced subscription service for module inputs, is needed in order to make solutions to this rather difficult problem available to operators in an affordable and technically tractable way.
3. Observing System Experiments (OSE) and Observing System Simulation Experiments (OSSE) should be carried out to identify the most impactful observation types and suitable arrangement of observing platforms for

data assimilation into operational models. Current and future missions, including GDC especially, should be analyzed from this perspective, in order to identify observations that should be prioritized for investment.
4. Provide stable data streams for future operations.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

SB is supported by CNES APR grant METEOESP.

SE is supported by the UK SWIMMR programme, NE/V002708/1.

PMM is supported by NASA CAREER award #2140204

References

Akmaev, R.A., Fuller-Rowell, T.J., Wu, F., Forbes, J.M., Zhang, X., Anghel, A.F., Iredell, M.D., Moorthi, S., Juang, H.M., 2008. Tidal variability in the lower thermosphere: comparison of Whole Atmosphere Model (WAM) simulations with observations from TIMED. Geophys. Res. Lett. 35, L03810. https://doi.org/10.1029/2007GL032584.

Bartels, J., 1949. The standardized index *Ks* and the planetary index *Kp*, *IATME Bulletin 12b,97.*

Boaghe, O., Balikhin, M., Billings, S., & Alleyne, H. (2001). Identification of nonlinear processes in the magnetospheric dynamics and forecasting of DST index. *Journal of Geophysical Research*, 106(A12), 30,047–30,066.

Baars, L., Hall, D., 2022. Processing Space Fence Radar Cross-Section Data to Produce Size and Mass Estimates, AAS Astrodynamics Specialist Conference (Paper #22-586), Charlotte NC, August 2022.

Balikhin, M., Boaghe, O., Billings, S., Alleyne, H.C., 2001. Terrestrial magnetosphere as a nonlinear resonator. Geophysical Research Letters 28 (6), 1123–1126.

Baumann, C., McCloskey, A.E., 2021. Timing of the solar wind propagation delay between L1 and Earth based on machine learning. J. Space Weather Space Clim. 11, 41. https://doi.org/10.1051/swsc/2021026.

Berger, T.E., Holzinger, M.J., Sutton, E.K., Thayer, J.P., 2020. Flying through uncertainty e2019SW002373. Space Weather 18. https://doi.org/10.1029/2019SW002373.

Bernstein, V., Pilinski, M., 2022. Drag coefficient constraints for space weather observations in the upper thermosphere, Space. Weather 20, e2021SW002977.

Boberg, F, Wintoft, P, Lundstedt, H, 2000. Real time Kp predictions from solar wind data using neural networks. Phys Chem Earth C: Solar Terr Planet Sci 25, 275–280.

Boley, A.C., Byers, M., 2021. Satellite mega-constellations create risks in Low Earth Orbit, the atmosphere and on Earth. Nat. Sci. Rep. 11, 10642. https://doi.org/10.1038/s41598-021-89909-7.

Bollacker, K.D., Ghosh, J., 1996. Linear feature extractors based on mutual information. In: Proceedings of 13th International Conference on Pattern Recognition, 2, pp. 720-724.

Boniface, C., Bruinsma, S., 2021. Uncertainty quantification of the DTM2020 thermosphere model. J. Space Weather Space Clim. 11, 53. https://doi.org/10.1051/swsc/2021034.

Bowman, B.R., Tobiska, W.K., Marcos, F., Huang, C.Y., Lin, C.S., Burke, W.J., 2008. A New Empirical Thermospheric Density Model JB2008 Using New Solar and Geomagnetic Indices. Presentation at AIAA/AAS Astrodynamics Specialist Conference, Honolulu, Hawaii. https://doi.org/10.2514/6.2008-6438.

Bruinsma, S.L., 2015. The DTM-2013 thermosphere model. J. Space Weather Space Clim. 2, A04. https://doi.org/10.1051/swsc/2012005.

Bruinsma, S., Boniface, C., 2021. The DTM2020 thermosphere models. J. Space Weather Space Clim. 11, 47. https://doi.org/10.1051/swsc/2021032.

Bruinsma, S.L., Doornbos, E., Bowman, B.R., 2014. Validation of GOCE densities and thermosphere model evaluation. Adv. Space Res. 54, 576–585. https://doi.org/10.1016/j.asr.2014.04.008.

Bruinsma, S., Sutton, E., Solomon, S.C., Fuller-Rowell, T., Fedrizzi, M., 2018. Space Weather Modeling Capabilities Assessment: Neutral Density and Orbit Determination at LEO. Space Weather 16 (11), 1806–1816. https://doi.org/10.1029/2018SW002027.

Bruinsma, S., Siemes, C., Emmert, J.T., Mlynczak, M.G., 2022. Description and comparison of 21st century thermosphere data. Adv. Space Res. https://doi.org/10.1016/j.asr.2022.09.038.

Bussy-Virat, C.D., Ridley, A.J., Getchius, J.W., 2018. Effects of uncertainties in the atmospheric density on the probability of collision between space objects. Space Weather 16 (5), 519–537. https://doi.org/10.1029/2017SW001705.

Calabia, A., Jin, S., 2016. New modes and mechanisms of thermospheric mass density variations from GRACE accelerometers. J. Geophys. Res. 121, 11191–11212. https://doi.org/10.1002/2016JA022594.

Carignan, G.R., Block, B.P., Maurer, J.C., Hedin, A.E., Reber, C.A., Spencer, N.W., 1981. The neutral mass spectrometer on Dynamics Explorer. Space Science Instruments 5, 429–441.

Casali, S., Barker, W., 2002. Dynamic Calibration Atmosphere (DCA) for the High Accuracy Satellite Drag Model (HASDM), AIAA 2002–4888. AIAA/AAS Astrodynamics Specialist Conference and Exhibit, Monterey, CA. https://doi.org/10.2514/6.2002-4888.

Cessateur, G., Dudok, T., de Wit, M., Kretzschmar, J.L., Hochedez, J.F., 2011. Monitoring the solar UV irradiance spectrum from the observation of a few passbands. Astronomy Astrophys. 528, A68. https://doi.org/10.1051/0004-6361/201015903.

Chakraborty, S., Morley, S.K., 2020. Probabilistic prediction of geomagnetic storms and the Kp index. J. Space Weather Space Clim. 10, 36. https://doi.org/10.1051/swsc/2020037.

Chamberlin, P.C., Woods, T.N., Eparvier, F.G., 2008. Flare Irradiance Spectral Model (FISM): Flare component algorithms and results. Space Weather 6, S05001. https://doi.org/10.1029/2007SW000372.

Charbonneau, P., 2020. Dynamo models of the solar cycle. Living Rev. Solar Phys. 17, 1. https://doi.org/10.1007/s41116-020-00025-6.

Chatterjee, T.N., 2001. On the application of information theory to the optimum state-space reconstruction of the short-term solar radio flux (10.7cm), and its prediction via a neural network. Monthly Notices Roy. Astronomical Soc. 323, 101–108. https://doi.org/10.1046/j.1365-8711.2001.04110.x.

Coddington, O., Lean, J., Pilewskie, P., et al., 2019. Solar irradiance variability: comparisons of models and measurements. Earth Space Sci. 6, 2525–2555. https://doi.org/10.1029/2019EA000693.

Codrescu, M.V., Fuller-Rowell, T.J., Minter, C.F., 2004. An ensemble-type Kalman filter for neutral thermospheric composition during geomagnetic storms. Space Weather 2, 11. https://doi.org/10.1029/2004SW000088.

Crowley, G., Pilinski, M., 2017. Reducing conjunction analysis errors with an assimilative tool for satellite drag specification, In: Proceedings of the 33rd Space Symposium, April 2017, Colorado Springs, CO. https://www.spacefoundation.org/wp-content/uploads/2019/07/Crowley_Geoff_Reducing_Conjunction_Analysis.pdf.

DeLand, M.T., Floyd, L.E., Marchenko, S., Tiruchirapalli, R., 2019. Creation of the GSFCSSI2 composite solar spectral irradiance data set. Earth Space Sci. 6, 1284–1298. https://doi.org/10.1029/2019EA000616.

Ding, C., Peng, H., 2005. Minimum redundancy feature selection from microarray gene expression data. Journal of Bioinformatics and Computational Biology 3 (02), 185–205.

Doornbos, E., 2011. Thermospheric density and wind determination from satellite dynamics, Ph.D. Dissertation. University of Delft, p. 188.

Dudok de Wit, T., Lilensten, J., Aboudarham, J., Amblard, P.-O., Kretzschmar, M., 2005. Retrieving the solar EUV spectrum from a reduced set of spectral lines. Annales Geophysicae 23, 3055–3069. https://doi.org/10.5194/angeo-23-3055-2005.

Dudok de Wit, T., Bruinsma, S., Shibasaki, S., 2014. Synoptic radio observations as proxies for upper atmosphere modelling. J. Space Weather Space Clim. 4, A06. https://doi.org/10.1051/swsc/2014003.

Elvidge, S., Angling, M.J., 2019. Using the local ensemble Transform Kalman Filter for upper atmospheric modelling. J. Space Weather Space Clim. 9, A30. https://doi.org/10.1051/swsc/2019018.

Elvidge, S., Godinez, H.C., Angling, M.J., 2016. Improved forecasting of thermospheric densities using multi-model ensembles. Geosci. Model Dev. 9, 2279–2292. https://doi.org/10.5194/gmd-9-2279-2016.

Emmert, J., Byers, J., Warren, H. and Segerman, A., 2014. Propagation of Forecast Errors from the Sun to LEO Trajectories: How Does Drag Affect Conjunction Frequency? In: 15th Advanced Maui Optical and Space Surveillance Technologies Conference. Maui, HI, https://apps.dtic.mil/sti/citations/ADA616724.

Emmert, J.T., Drob, D.P., Picone, J.M., et al., 2020. NRLMSIS 2.0: A whole-atmosphere empirical model of temperature and neutral species densities. Earth Space Sci. 8 (3). https://doi.org/10.1029/2020EA001321, e2020EA001321.

Emmert, J.T., Warren, H.P., Segerman, A.M., Byers, J.M., Picone, J.M., 2017. Propagation of atmospheric density errors to satellite orbits. Adv. Space Res. 59 (1), 147–165. https://doi.org/10.1016/j.asr.2016.07.036.

Fang, T.W., Fuller-Rowell, T.J., Yudin, V., Matsuo, T., Viereck, R., 2018. Quantifying the sources of ionosphere day-to-day variability. J. Geophys. Res. 123, 9682–9696. https://doi.org/10.1029/2018JA025525.

Fang, T.W., Kubaryk, A., Goldstein, D., Li, Z., Fuller-Rowell, T., Millward, G., et al., 2022. Space Weather Environment during the SpaceX Starlink Satellite Loss in February 2022. Space Weather 20, 11. https://doi.org/10.1029/2022SW003193.

Friedman, A.M., Frueh, C., 2022. Observability of light curve inversion for shape and feature determination exemplified by a case of analysis. J. Astronautics Sciences 69, 537–569. https://doi.org/10.1007/s40295-021-00293-w.

Fuller-Rowell, T.J., Evans, D.S., 1987. Height integrated Pedersen and Hall conductivity patterns inferred from the TIROS-NOAA satellite data. J. Geophys. Res. 92, 7606–7618. https://doi.org/10.1029/JA092iA07p07606.

Gondelach, D.J., Linares, R., 2020a. Realtime thermospheric density estimation via two-line element data assimilation. Space Weather 18, 2. https://doi.org/10.1029/2019SW002356.

Gondelach, D.J., Linares, R., 2020b. Atmospheric Density Uncertainty Quantification for Satellite Conjunction Assessment, AIAA Scitech Forum, Paper AIAA 2020–0232. Orlando, FL. https://doi.org/10.2514/6.2020-0232.

Haberreiter, M., Schöll, M., Dudok, T., de Wit, M., Kretzschmar, S.M., Tourpali, K., Schmutz, W., 2017. A new observational solar irradiance composite. J. Geophys. Res. 122, 5910–5930. https://doi.org/10.1002/2016JA023492, 2016JA023492.

Hapgood, M., Liu, H., Lugaz, N., 2022. SpaceX - Sailing Close to the Space Weather? Space Weather 20. https://doi.org/10.1029/2022SW003074.

Heelis, R.A., Lowell, J.K., Spiro, R.W., 1982. A model of the high-latitude ionosphere convection pattern. J. Geophys. Res. 87 (A8), 6339–6345. https://doi.org/10.1029/JA087iA08p06339.

Hejduk, M.D., Snow, D.E., 2018. The effect of neutral density estimation errors on satellite conjunction serious event rates. Space Weather 16 (7), 849–869. https://doi.org/10.1029/2017SW001720.

Henney, C.J., Toussaint, W.A., White, S.M., Arge, C.N., 2012. Forecasting F10.7 with solar magnetic flux transport modeling, Space. Weather 10, 2. https://doi.org/10.1029/2011SW000748.

Henney, C.J., Hock, R.A., Schooley, A.K., Toussaint, W.A., White, S.M., Arge, C.N., 2015. Forecasting solar extreme and far ultraviolet irradiance. Space Weather 13 (3), 141–153. https://doi.org/10.1002/2014SW001118.

Ho, T.K., 1995. Random decision forests, In Proceedings of 3rd international conference on document analysis and recognition, 1, pp. 278-282, IEEE, DOI: 10.1109/ICDAR.1995.598994.

Inter-Agency Space Debris Coordination Committee, 2021. IADC Space Debris Mitigation Guidelines, Rev. 3, https://www.iadc-home.org/documents_public/view/id/172#u.

ISO, 2007. 21348:2007 Space environment (natural and artificial) – Process for determining solar irradiances, International Organization for Standardization, Geneva https://www.iso.org/standard/39911.html.

Immel, T.J., England, S.L., Mende, S.B., et al., 2018. The Ionospheric Connection Explorer Mission: Mission Goals and Design. Space Sci Rev 214, 13 https://doi.org/10.1007/s11214-017-0449-2.

Kane, R.P., 2007. A preliminary estimate of the size of the coming solar cycle 24, based on ohl's precursor method. Solar Phys. 243, 205–217. https://doi.org/10.1007/s11207-007-0475-4.

Karak, B.B., Nandy, D., 2012. Turbulent pumping of magnetic flux reduces solar cycle memory and thus impacts predictability of the Sun's activity. Astrophys. J. Lett. 761, 1. https://doi.org/10.1088/2041-8205/761/1/L13.

Lean, J.L., Picone, J.M., Emmert, J.T., 2009. Quantitative forecasting of near-term solar activity and upper atmospheric density. J. Geophys. Res. 114, A7. https://doi.org/10.1029/2009JA014285.

Leka, K.D., Park, S.H., Kusano, K., et al., 2019. A comparison of flare forecasting methods. II. Benchmarks, metrics, and performance results for operational solar flare forecasting systems. Astrophys. J. Supplement Ser. 243, 36. https://doi.org/10.3847/1538-4365/ab2e12.

Licata, R.J., Mehta, P.M., Tobiska, W.K., Bowman, B.R., Pilinski, M.D., 2021a. Qualitative and quantitative assessment of the SET HASDM database. Space e2021SW002798. Weather 19. https://doi.org/10.1029/2021SW002798.

Licata, RJ, Mehta, P.M., Tobiska, W.K., 2021. Impact of driver and model uncertainty on drag and orbit prediction, In Proceedings of the 31st AAS/AIAA Space Flight Mechanics Meeting, https://www.researchgate.net/publication/349117733_Impact_of_Driver_and_Model_Uncertainty_on_Drag_and_Orbit_Prediction.

Licata, R.J., Mehta, P.M., Weimer, D.R., Tobiska, W.K., Yoshii, J., 2022b. MSIS-UQ: Calibrated and enhanced NRLMSIS 2.0 model with uncertainty quantification e2022SW003267. Space Weather 20 (11). https://doi.org/10.1029/2022SW003267.

Licata, R., Mehta, P., Tobiska, W.K., 2020a. Impact of space weather driver forecast uncertainty on drag and orbit prediction, Astrodynamics specialist conference, AAS 20–423. Lake Tahoe, CA.

Licata, R.J., Tobiska, W.K., Mehta, P.M., 2020b. Benchmarking forecasting models for space weather drivers. Space Weather 18, 10. https://doi.org/10.1029/2020SW002496.

Licata, R.J., Mehta, P.M., Tobiska, W.K., Huzurbazar, S., 2022a. Machine-Learned HASDM Model with Uncertainty Quantification. Space Weather 20, 4. https://doi.org/10.1029/2021SW002915.

Liemohn, M.W., McCollough, J.P., Jordanova, V.K., et al., 2018. Model evaluation guidelines for geomagnetic index predictions. Space Weather 16, 2079–2102. https://doi.org/10.1029/2018SW002067.

Lilensten, J., Dudok, T., de Wit, M., Kretzschmar, PO.A., Moussaoui, S., Aboudarham, J., Auchère, F., 2008. Review on the solar spectral variability in the EUV for space weather purposes. Annales Geophysicae 26, 269–279. https://doi.org/10.5194/angeo-26-269-2008.

Lumpe, J.D., McClintock, W.E., Evans, J.S., Correira, J., Veibell, V., Beland, S., Eastes, R., 2020. A new data set of thermospheric molecular oxygen from the Global-scale Observations of the Limb and Disk (GOLD) mission. Journal of Geophysical Research: Space Physics 125 https://doi.org/10.1029/2020JA027812 e2020JA027812.

Luo, J., Zhu, H., Jiang, Y., Yang, J., Huang, Y., 2020. The 10.7-cm radio flux multistep forecasting based on empirical mode decomposition and back propagation neural network. IEEJ Trans. Electrical Electron. Eng. 15 (4), 584–592. https://doi.org/10.1002/tee.23092.

March, G., Doornbos, E., Visser, P., 2019a. High-fidelity geometry models for improving the consistency of CHAMP, GRACE, GOCE and Swarm thermospheric density data sets. Adv. Space Res. 63 (1), 213–238. https://doi.org/10.1016/j.asr.2018.07.009.

March, G., van den Ijssel, J., Siemes, C., Visser, P., Doornbos, E., Pilinski, M., 2019b. Gas-surface interactions modelling influence on satellite aerodynamics and thermosphere mass density. J. Space Weather Space Climate 11, 54. https://doi.org/10.1051/swsc/2021035.

Maruyama, N., Sun, Y.Y., Richards, P.G., Middlecoff, J., Fang, T.W., Fuller-Rowell, T.J., Akmaev, R.A., Liu, J.Y., Valladares, C.E., 2016. A new source of the midlatitude ionospheric peak density structure revealed by a new Ionosphere-Plasmasphere model. Geophys. Res. Lett. 43, 2429–2435. https://doi.org/10.1002/2015GL067312.

Matsuo, T., Fedrizzi, M., Fuller-Rowell, T.J., Codrescu, M.V., 2012. Data assimilation of thermospheric mass density. Space Weather 10, S05002. https://doi.org/10.1029/2012SW000773.

Mehta, P.M., Walker, A.C., Sutton, E.K., Godinez, H.C., 2017. New density estimates derived using accelerometers on-board the CHAMP and GRACE satellites. Space Weather 15, 558–576. https://doi.org/10.1002/2016SW001562.

Mehta, P.M., McLaughlin, C.A., Sutton, E.K., 2013. A Drag Coefficient Model for GRACE developed using Direct Simulation Monte Carlo. Adv. Space Res. 52 (12), 2035–2051. https://doi.org/10.1016/j.asr.2013.08.033.

Mehta, P.M., Walker, A., Lawrence, E., Linares, R., Higdon, D., Koller, J., 2014b. Modeling Satellite Drag Coefficients with Response Surfaces. Adv. Space Res. 54 (8), 1590–1607. https://doi.org/10.1016/j.asr.2014.06.033.

Mehta, P.M., Linares, R., Walker, A.C., 2018. Photometric Data from Non-Resolved Objects for Improved Drag and Re-entry Prediction. J. Spacecraft Rockets 55 (4), 959–970. https://doi.org/10.2514/1.A33825.

Mehta, P.M., Linares, R., Sutton, E.K., 2019. Data-Driven Inference of Thermosphere Composition During Solar Minimum Conditions. Space Weather 17 (9), 1364–1379. https://doi.org/10.1029/2019SW002264.

Mehta, P.M., Paul, S.N., Crisp, N.H., Sheridan, P.L., Siemes, C., March, G., Bruinsma, S., 2022. Satellite drag coefficient modeling for thermosphere science and operations. Adv. Space Res. https://doi.org/10.1016/j.asr.2022.05.064.

Meier, R.R. et al., 2015. Remote Sensing of Earth's Limb by TIMED/GUVI: Retrieval of thermospheric composition and temperature. Earth and Space Science 2, 1–37. https://doi.org/10.1002/2014EA000035.

Morley, S.K., Welling, D.T., Woodroffe, J.R., 2018. Perturbed input ensemble modeling with the space weather modeling framework. Space Weather 16, 1330–1347. https://doi.org/10.1029/2018SW002000.

Morozov, A.V., Ridley, A.J., Bernstein, D.S., Collins, N., Hoar, T.J., Anderson, J.L., 2013. Data assimilation and driver estimation for the Global Ionosphere-Thermosphere Model using the Ensemble Adjustment Kalman Filter. J. Atmospheric and Solar-Terrestrial Phys. 104, 126–136. https://doi.org/10.1016/j.jastp.2013.08.016.

Murray, S.A., 2018. The importance of ensemble techniques for operational space weather forecasting. Space Weather 16, 777–783. https://doi.org/10.1029/2018SW001861.

Murray, S.A., Henley, E.M., Jackson, D.R., Bruinsma, S.L., 2015. Assessing the performance of thermospheric modeling with data assimilation throughout solar cycles 23 and 24. Space Weather 13, 4. https://doi.org/10.1002/2015SW001163.

Nier, A.O., Potter, W.E., Hickman, D.R., Mauersberger, K., 1973. The open-source neutral-mass spectrometer on Atmosphere Explorer-C, -D, -E. Radio Science 8, 271–276.

Pelz, D.T., Reber, C.A., Hedin, A.E., Carignan, G.C., 1973. A neutral-atmosphere composition experiment for the Atmosphere Explorer-C, -D, -E. Radio Sci. 8, 277–283.

Peng, L.F., Long, F., Ding, C., 2005. Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. IEEE Trans. Pattern Anal. Machine Intell. 27 (8), 1226–1238.

Pesnell, W.D., 2012. Solar cycle predictions. Solar Phys. 281, 507–532. https://doi.org/10.1007/s11207-012-9997-5.

Pesnell, W.D., 2016. Predictions of solar cycle 24: How are we doing? Space Weather 14 (1), 10–21. https://doi.org/10.1002/2015SW001304.

Petrovay, K., 2020. Solar cycle prediction. Living Rev. Solar Phys. 17, 2. https://doi.org/10.1007/s41116-020-0022-z.

Picone, J.M., Hedin, A.E., Drob, D.P., Aikin, A.C., 2002. NRLMSISE-00 empirical model of the atmosphere: statistical comparisons and scientific issues. J. Geophys. Res. 107 (A12), 1468. https://doi.org/10.1029/2002JA009430.

Pilinski, M.D., Argrow, B.M., Palo, S.E., 2013. Semi-empirical satellite accommodation model for spherical and randomly tumbling objects. J. Spacecraft Rockets 50, 556–571. https://doi.org/10.2514/1.A32348.

Pilinski, M.D., Bowman, B.A., Palo, S.E., Forbes, J.M., Davis, B.L., Moore, R.G., Koehler, C., Sanders, B., 2016. Comparative analysis of satellite aerodynamics and its application to space-object identification. J. Spacecraft Rockets 53, 5. https://doi.org/10.2514/1.A33482.

Pilinski, M., Crowley, G., Seaton, M., Sutton E., 2019. Dragster: an assimilative tool for satellite drag specification. In: Advanced Maui Optical and Space Surveillance Technologies Conference, https://amostech.com/TechnicalPapers/2019/Astrodynamics/Pilinski.pdf.

Poore, A.B., Aristoff, J.M., Horwood, J.T. et al., 2016. Covariance and uncertainty realism in space surveillance and tracking, Tech. report Numerica Corporation, Fort Collins, CO, USA. https://apps.dtic.mil/sti/pdfs/AD1020892.pdf.

Qian, L., Burns, A.G., Chamberlin, P.C., Solomon, S.C., 2011. Variability of thermosphere and ionosphere responses to solar flares. J. Geophys. Res. 116, A10309. https://doi.org/10.1029/2011JA016777.

Qian, L., Solomon, S.C., 2012. Thermospheric density: An overview of temporal and spatial variations. Space Sci. Rev. 168, 147–173. https://doi.org/10.1007/s11214-011-9810-z.

Richmond, A.D., 1992. Assimilative mapping of ionospheric electrodynamics. Adv. Space. Res. 12, 59–68. https://doi.org/10.1016/0273-1177(92)90040-5.

Rostoker, G., 1972. Geomagnetic indices. Rev. Geophys. 10 (4), 935–950. https://doi.org/10.1029/RG010i004p00935.

Schiemenz, F., 2021. Covariance and Uncertainty Realism for Low Earth Orbiting Satellites via Quantification of Dominant Force Model Uncertainties. Doctoral dissertation, University of Würzburg. https://doi.org/10.25972/OPUS-24947.

Schiemenz, F., Utzmann, J., Kayal, H., 2019a. Least squares orbit estimation including atmospheric density uncertainty consideration. Adv. Space Res. 63 (12), 3916–3935. https://doi.org/10.1016/j.asr.2019.02.039.

Schiemenz, F., Utzmann, J., Kayal, H., 2019b. Propagating EUV solar flux uncertainty to atmospheric density uncertainty. Adv. Space Res. 63 (12), 3936–3952. https://doi.org/10.1016/j.asr.2019.02.040.

Schiemenz, F., Utzmann, J., Kayal, H., 2020a. Adaptive Gaussian Mixture based Orbit Determination with combined atmospheric density uncertainty consideration. Adv. Space Res. 66 (7), 1609–1634. https://doi.org/10.1016/j.asr.2020.05.042.

Schiemenz, F., Utzmann, J., Kayal, H., 2020b. Propagation of grid-scale density model uncertainty to orbital uncertainties. Adv. Space Res. 65 (1), 407–418.

Schiemenz, F., Utzmann, J., Kayal, H., 2020c. Accurate estimation of relative atmospheric density error on the example of uncertain geomagnetic activity information. Adv. Space Res. 65 (1), 251–270. https://doi.org/10.1016/j.asr.2019.10.013.

Schmidtke, G., 1976. EUV indices for solar-terrestrial relations. Geophys. Res. Lett. 3, 573–576. https://doi.org/10.1029/GL003i010p00573.

Sheridan, P.L., Paul, S.N., Avendano-Franco, G., Mehta, P.M., 2022. Updates and improvements to the satellite drag coefficient Response Surface Modeling toolkit. Adv. Space Res. 69 (10), 3828–3846. https://doi.org/10.1016/j.asr.2022.02.044.

Shprits, Y.Y., Vasile, R., Zhelayskaya, I.S., 2019. Nowcasting and predicting the Kp index using historical values and real-time observations. Space Weather 17 (8), 1219–1229. https://doi.org/10.1029/2018SW002141.

Sinpetru, L.A., Crisp, N.H., Mostaza-Prieto, D., Livadiotti, S., Roberts, P.C.E., 2022. ADBSat: Methodology of a novel panel method tool for aerodynamic analysis of satellites. Computer Phys. Commun. 275. https://doi.org/10.1016/j.cpc.2022.108326 108326.

Snow, M., Machol, J., Viereck, R., Woods, T., Weber, M., Woodraska, D., Elliott, J., 2019. Revised Magnesium II Core-to-Wing Ratio From SORCE SOLSTICE. Earth Space Sci. 6 (11), 2106–2114. https://doi.org/10.1029/2019EA000652.

Storz, M.F., Bowman, B.R., Branson, M.J.I., Casali, S.J., Tobiska, W.K., 2005. High accuracy satellite drag model (HASDM). Adv. Space Res. 36 (12), 2497–2505. https://doi.org/10.1016/j.asr.2004.02.020.

Suess, K., Snow, M., Viereck, R., Machol, J., 2016. Solar Spectral Proxy Irradiance from GOES (SSPRING): a model for solar EUV irradiance. J. Space Weather Space Climate 6, A10. https://doi.org/10.1051/swsc/2016003.

Sutton, E.K., 2018. A new method of physics-based data assimilation for the quiet and disturbed thermosphere. Space Weather 16 (6), 736–753. https://doi.org/10.1002/2017SW001785.

Sutton, E.K., Forbes, J.M., Nerem, R.S., Woods, T.N., 2006. Neutral density response to the solar flares of October and November, 2003. Geophys. Res. Lett. 33. https://doi.org/10.1029/2006GL027737.

Tan, Y., Hu, Q., Wang, Z., Zhong, Q., 2018. Geomagnetic index Kp forecasting with LSTM. Space Weather 16, 406–416 https://doi.org/10.1002/2017SW001764.

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

Thiemann, E.M.B., Eparvier, F.G., Woodraska, D., Chamberlin, P.C., Machol, J., Eden, T., Jones, A.R., Meisner, R., Mueller, S., Snow, M., Viereck, R., Woods, T.N., 2019. The GOES-R EUVS model for EUV irradiance variability. J. Space Weather Space Climate 9, A43. https://doi.org/10.1051/swsc/2019041.

Tobiska, W.K., Bouwer, S., Bowman, B., 2008. The development of new solar indices for use in thermospheric density modeling. J. Atmospheric and Solar-Terrestrial Phys. 70, 803–819. https://doi.org/10.1016/j.jastp.2007.11.001.

Tobiska, W.K., Bowman, B.R., Bouwer, S.D., Cruz, A., Wahl, K., Pilinski, M.D., Mehta, P.M., Licata, R.J., 2021. The SET HASDM density database e2020SW002682. Space Weather 19. https://doi.org/10.1029/2020SW002682.

Tobiska, W.K., Pilinski, M.D., Mutschler, S., Wahl, K., Yoshii, J., Bouwer, D., Mehta, P., Licata, R., 2022. Understanding variability in HASDM to support space traffic management, 23rd AMOS September 27–30, 2022. Maui, Hawaii.

Turner, H., Zhang, M., Gondelach, D., Linares, R., 2020. Machine Learning Algorithms for Improved Thermospheric Density Modeling. In: Darema, F., Blasch, E., Ravela, S., Aved, A. (Eds.) Dynamic Data Driven Applications Systems. DDDAS 2020. Lecture Notes in Computer Science, vol 12312. Springer, Cham. Doi: 10.1007/978-3-030-61725-7_18.

Vallado, D., Finkleman, D., 2014. A critical assessment of satellite drag and atmospheric density modeling. Acta Astronautica 95, 141–165. https://doi.org/10.1016/j.actaastro.2013.10.005.

Viereck, R., Hanser, F., Wise, J., Guha, S., Jones, A., McMullin, D., Plunket, S., Strickland, D., Evans, S., 2007. Solar extreme ultraviolet irradiance observations from GOES: design characteristics and initial performance. In: Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, volume 6689 of Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series. Doi: 10.1117/12.734886.

Vourlidas, A., Bruinsma, S., 2018. EUV Irradiance Inputs to Thermospheric Density Models: Open Issues and Path Forward. Space Weather 16, 5–15. https://doi.org/10.1002/2017SW001725.

Walker, A., Mehta, P.M., Koller, J., 2014. Drag Coefficient Modeling using the Cercignani-Lampis-Lord Gas-Surface Interaction Model. J. Spacecraft Rockets 51 (5), 1544–1563. https://doi.org/10.2514/1.A32677.

Warren, H.P., Emmert, J.T., Crump, N.A., 2017. Linear forecasting of the F10.7 proxy for solar activity. Space Weather 15, 1039–1051. https://doi.org/10.1002/2017SW001637.

Weimer, D.R., 2005. Improved ionospheric electrodynamic models and application to calculating Joule heating rates. J. Geophys. Res. 110, A05306. https://doi.org/10.1029/2004JA010884.

Wintoft, P., Wik, M., Matzka, J., Shprits, Y., 2017. Forecasting Kp from solar wind data: Input parameter study using 3-hour averages and 3-hour range values. Journal of Space Weather and Space Climate 7, A29 https://doi.org/10.1051/swsc/2017027.

Wu, D.L., Yee, J.H., Schlecht, E., Mehdi, I., Siles, J., Drouin, B.J., 2016. THz limbsounder (TLS) for lower thermosphericwind, oxygen density, and temperature. J. Geophys. Res. 121, 7301–7315. https://doi.org/10.1002/2015JA022314.

Yamazaki, Y., Matzka, J., Stolle, C., Kervalishvili, G., Rauberg, J., Bronkalla, O., Morschhauser, A., Bruinsma, S., Shprits, Y.Y., Jackson, D.R., 2022. Geomagnetic activity index Hpo. Geophys. Res. Lett. 49, 10. https://doi.org/10.1029/2022GL098860.

Yaya, P., Hecker, L., Dudok, T., de Wit, C., Le Fèvre, S., Bruinsma, 2017. Solar radio proxies for improved satellite orbit prediction. J. Space Weather Space Climate 7, A35. https://doi.org/10.1051/swsc/2017032.

Yue, J., Yu, W., Pedatella, N., Bruinsma, S., Wang, N., Liu, H., 2022. Contribution of the lower atmosphere to the day-to-day variation of thermospheric density. Adv. Space Res. https://doiorg/10.1016/j.asr.2022.06.011.

Zhelavskaya, I.S., Vasile, R., Shprits, Y.Y., Stolle, C., Matzka, J., 2019. Systematic analysis of machine learning and feature selection techniques for prediction of the Kp Index. Space Weather 17. https://doi.org/10.1029/2019SW002271.