Qualitative and Quantitative Assessment of the SET HASDM Database

Richard J. Licata[1]*, Piyush M. Mehta[1], W. Kent Tobiska[2], Bruce R. Bowman[2], and Marcin D. Pilinski[3]

[1]Department of Mechanical and Aerospace Engineering, West Virginia University, Morgantown, West Virginia, USA.

[2]Space Environment Technologies, Pacific Palisades, California, USA.

[3]Freestream, Boulder, Colorado, USA.

Key Points:

- Principal Component Analysis of HASDM and JB2008 Database covering almost two solar cycles.
- Dominant modes of variation are identified and compared.
- Statistical comparison with CHAMP and GRACE observations.

*1306 Evansdale Drive, Morgantown, West Virginia 26506-6106.

Corresponding author: Richard J. Licata, rjlicata@mix.wvu.edu

Abstract

The High Accuracy Satellite Drag Model (HASDM) is the operational thermospheric density model used by the US Space Force (USSF) Combined Space Operations Center (CSpOC). By using real-time data assimilation, HASDM can provide density estimates with increased accuracy over empirical models. With historical HASDM density data being released publicly for the first time, we can analyze the data to identify dominant modes of variations in the upper atmosphere. As HASDM is a close relative to the Jacchia-Bowman 2008 Empirical Thermospheric Density Model (JB2008), we look at time-matched density data to better understand the models' characteristics. This model comparison is conducted through the use of Principal Component Analysis (PCA). We then compare both datasets to the CHAllenging Minisatellie Payload (CHAMP) and Gravity Recovery and Climate Experiment (GRACE) accelerometer-derived density estimates. By looking at the principal components and PCA scores from the two models, we confirm the increased complexity of the HASDM dataset while the CHAMP and GRACE comparisons show that HASDM more closely matches the accelerometer-derived densities with mean absolute differences of 23.81% and 30.84% compared to CHAMP and GRACE-A, respectively.

1 Introduction

Over the past seven decades, the scientific community has developed and advanced thermospheric density models. A significant subset of these models are empirical. Empirical models use long-term trends from measurements over an array of instruments to fit parametric equations that describe the system. Even within this subset, there are multiple families/series of models that use different types of measurements and have evolved over decades. Three of these series, discussed by Emmert (2015), are the MSIS (*Picone et al.* 2002), DTM (*Bruinsma, Sean* 2015), and Jacchia series (*Bowman et al.* 2008).

Mass Spectrometer Incoherent Scatter Radar (MSIS) models typically used mass spectrometer and incoherent scatter radar measurements but have evolved and now incorporate additional data (e.g. accelerometer-derived density estimates). The Drag Temperature Model (DTM) series used orbit-derived density data but more recently incorporates accelerometer-derived density and mass spectrometer data. The Jacchia series of models

(e.g. Jacchia-70 and JB2008) strictly use both orbit- and accelerometer-derived density estimates.

The most recent in the Jacchia series is the JB2008 density model. JB2008 was an improvement to its predecessors and incorporated new solar and geomagnetic indices to drive the model. It uses the F_{10}, S_{10}, M_{10}, and Y_{10} indices and proxies to model variations caused by solar heating. In addition to a_p, JB2008 utilizes Dst to better model density during geomagnetic storms.

The thermosphere is a dynamic, highly-driven system impacted by external forces (e.g. space weather events) and internal dynamics. Solar irradiance is a major source of variation, providing the baseline average density (*Qian and Solomon* 2011). This process is well-represented by solar indices, particularly at low latitudes (*Vickers et al.* 2014). However, these indices are not adequate in characterizing the thermosphere during solar minimum (*Bowman et al.* 2008), when composition changes and other processes become more relevant (*Mehta et al.* 2019). During events like solar flares or coronal mass ejections (CMEs), mass and energy from the Sun interact with the magnetosphere causing Joule heating and auroral particle precipitation into the thermosphere (*Fedrizzi et al.* 2012; *Deng et al.* 2013). This causes sudden and often large changes in mass density. Due to a lack of pre-storm conditioning and inability to model traveling atmospheric disturbances, empirical models frequently under-perform during these events (*Bruinsma et al.* 2021).

The challenge of accurately modeling thermospheric mass density over a multitude of conditions has severe repercussions in the context of orbit determination and Space Situational Awareness (SSA). Operators rely on these density models for decision making in regards to collision avoidance operations, where inaccuracies and uncertainties can have dire consequences. Insufficient knowledge of future satellite position can result in a collision between two objects. A potential collision would drastically increase the number of objects in a given orbital regime, increasing the probability of future collisions. This could result in a cascade, known as Kessler Syndrome (*Kessler and Cour-Palais* 1978), which could make certain orbital regimes inaccessible. In an effort to avoid this, we look to improve our modeling and forecasting capabilities. A significant improvement in model accuracy came from the implementation of real-time data assimilation, notably utilized by HASDM.

HASDM was developed by Storz et al. (2005) and is an assimilative extension of the Jacchia 1970 upper atmosphere density model (*Jacchia* 1970). HASDM employs Dynamic Calibration of Atmosphere (DCA) which uses calibration satellite observations to make corrections to its background empirical density model. This assimilation technique was introduced as an application for HASDM by Casili and Barker (2002), but was expanded later to estimate 13 global density correction parameters (*Storz et al.* 2005). HASDM is not available for public use, but the global density outputs from the model were released to the public for the first time by Tobiska et al. (2021). It is called the SET HASDM density database. This database contains three-dimensional density grid from the start of 2000 to the end of 2019 at a three-hour cadence.

In this work, we will leverage Principal Component Analysis (PCA) in order to study the most dominant sources of variance within the HASDM dataset and a spatiotemporally-matched JB2008 dataset. The resulting principal components and PCA scores give insight into the processes that drive the variance within the models. This methodology has been used to analyze thermospheric density datasets previously and is often used in the development of reduced-order models (*Mehta and Linares* 2017; *Mehta et al.* 2018; *Gondelach and Linares* 2020). For this paper, the use of PCA is restricted to scientific investigation.

The availability of accelerometer-derived density estimates has been advantageous for model development and assessment. Over the lifetime of satellites with on board accelerometers (e.g. CHAMP and GRACE), we accumulate measurements over a plethora of altitudes and space weather conditions (*Luhr et al.* 2002; *Bettadpur* 2012). Researchers have used these measurements to derive density estimates by removing accelerations from other sources (*Sutton* 2008; *Doornbos* 2012; *Mehta et al.* 2017). We use the estimates from (*Mehta et al.* 2017) for comparisong with the HASDM and JB2008 models.

The paper is organized as follows, we start by detailing the HASDM and JB2008 models. Then, we discuss the use of PCA as an investigatory tool followed by the results of the analysis. After, we compare the HASDM and JB2008 densities to CHAMP and GRACE density estimate over the entire availability of their measurements, with a focus on storm-time and quiet conditions.

2 Model Background

The most recent Jacchia model, JB2008, achieved improved accuracy largely due to its incorporations of new solar and geomagnetic indices. These indices are used in temperature corrections, semiannual functions, and new Dst temperature equations . The model reduced non-storm density errors by $>$ 5% and reduced storm-time density errors from Jacchia-70 by $>$ 60%, from NRLMSIS by $>$ 35% and from JB2008 (with only a_p) by 16% (*Bowman et al.* 2008).

Using a similar background density model, HASDM is able to further reduce these errors. By building on the density correction work of Marcos et al. (1998) and Nazerenko et al. (1998), HASDM can provide dynamic global density corrections via 13 spherical harmonic coefficients through its DCA algorithm. HASDM also exploits a prediction filter for its DCA corrections. Through this filter, the model adjusts an extrapolated time series of 27 days (one solar rotation) for the correction coefficients using wavelet and Fourier analysis (*Storz et al.* 2005). For satellite trajectory estimation, HASDM uses a technique, called segmented solution for ballistic coefficient (SSB), that enables the estimated ballistic coefficient to deviate over the fitting period.

2.1 Model Drivers

The most common solar proxy used in density modeling is $F_{10.7}$, referred to in this paper as F_{10}. Originally identified and measured by Covington (1948), F_{10} serves as a proxy for solar extreme ultraviolet (EUV) emissions which deposit energy into the thermosphere. The 10.7 in the subscript refers to the 10.7 cm wavelength of the solar radio flux being measured. While this does not directly interact with Earth's atmosphere, it has been shown to be a reliable proxy for thermospheric heating (*Tobiska et al.* 2008a). F_{10} is measured in solar flux units ($10^{-22} W\ m^{-2}\ Hz^{-1}$) indicated as *sfu*.

The S_{10} index characterizes the integrated 26-32 nm solar EUV emission. This index is influenced by temperatures in the chromosphere and solar corona (*Tobiska et al.* 2008a). These emissions penetrate into the middle thermosphere and are absorbed by atomic oxygen (*Tobiska et al.* 2008b). While the emissions that S_{10} represents have no relationship to the 10.7 cm wavelength, they are normalized and converted to *sfu* through linear regression. Similar fits are done for M_{10} and Y_{10} to convert to uniform units.

M_{10} is a proxy representative of far ultraviolet (FUV) photospheric 160 nm Schumann-Runge Continuum emissions. The proxy corresponds to processes in the lower thermosphere and is consistent with molecular oxygen dissociation (*Tobiska et al.* 2008b). The final solar driver for JB2008 is Y_{10}, which is a composite index. This hybrid represents X-ray emissions in the 0.1-0.8 nm range and H Lyman-α 121 nm emissions. During solar maximum, the X-ray emissions are more heavily weighted, and the opposite is true for solar minimum. For each of these four solar drivers, 81-day centered averages are generated and used for prediction in JB2008.

The first of the two geomagnetic drivers for JB2008 is the geomagnetic planetary amplitude, a_p. a_p is the linear equivalent of the geomagnetic planetary index, K_p, which has a quasi-logarithmic scale (*McClain and Vallado* 2001). It has a 3-hour cadence and is often used in density models. However, using Dst during geomagnetic storms results in increased accuracy over a_p for density modeling (*Bowman et al.* 2008). The Dst index is largely driven by the strength of the ring current in the inner magnetosphere. This makes it an ideal indicator of ring current strength and therefore geomagnetic storms (*Ganushkina et al.* 2017).

For operational use of HASDM, forecasts of these drivers are required. Space Environment Technologies (SET) provides the driver forecasts using multiple algorithms/sources. The solar drivers are forecasted using the *SOLAR200* algorithm (*Tobiska et al.* 2000). a_p forecasts come from the National Oceanic and Atmospheric Administration (NOAA) Space Weather Prediction Center (SWPC) forecasters, and Dst forecasts are a produced by the *Anemomilos* algorithm (*Tobiska et al.* 2013). Error statistics of historical forecasts for all six drivers were presented as a community benchmark by Licata et al. (2020).

In addition to these space weather drivers, the models use temporal inputs (e.g. universal time (UT) and day of year). To model/correlate seasonal and annual trends, we generated sinusoidal inputs based on the day of year. The first two are sine and cosine functions with periods of six months. This is used to test correlations with semiannual trends. The last two are sine and cosine functions with periods of one year to correlate with annual trends. We did not include any functions of UT, because it is difficult to find linear correlations with these short period fluctuations, particularly when the cadence of the data is only $1/8$ of the function's period.

3 Principal Component Analysis

The spatial resolution of these models are 15° longitude, 10° latitude, and 25 km altitude spanning from 175-825 km. This results in 12,312 grid points for every three hours between the start of 2000 to the end of 2019. Principal Component Analysis is a dimensionality reduction technique that translates the dataset using linear functions to maximize variability and preserve information. The resulting information's dimensions are ordered from most-to-least contribution to the system's variance. Therefore, you can truncate the data at the desired point and only sacrifice a nominal amount of information (*Jolliffe and Cadima* 2016). This compaction forces important information from the dataset to be represented in fewer dimensions, simplifying the analysis.

In previous work, Mehta and Linares (2017) developed a methodology for reduced order modeling where PCA was used to reduce the dimensionality of MSIS densities. This allowed the authors to analyze the dominant sources of variance. We apply that same methodology (PCA implemenation) to identify principal components for both the HASDM and JB2008 datasets. The steps to achieve this will be described, but the original work of (2017) provides a thorough description of the process. This methodology was used on a TIE-GCM dataset to create a linear dynamic reduced order model (*Mehta et al.* 2018).

Initially, the spatial dimensions are flattened to make the spatiotemporal dataset two-dimensional. Then a common logarithm of the density values in taken in order to reduce the variance of the dataset from five orders of magnitude to less than one. Next, we subtract the mean values for each cell to center the data. Finally, we perform a singular value decomposition using the *svds* function in *MATLAB* to obtain the $\mathbf{U}$, Σ, and $\mathbf{V}$ matrices. The normalized and centered density data is denoted by $\mathbf{M}$. Equation 1 shows the relationship between these four matrices and the basis of PCA.

$$\mathbf{M} = \mathbf{U}\Sigma\mathbf{V}^{\mathbf{T}} \tag{1}$$

The $\mathbf{U}$ matrix is the left unitary matrix, and it is made of orthogonal vectors that represent the principal components. Σ is a diagonal matrix consisting of the squares of the eigenvalues that correspond to the vectors in $\mathbf{U}$. We can extract temporal coefficients (shown in Equation 2) by performing matrix multiplication between two of the components in Equation 1.

$$\mathbf{c}(\mathbf{t}) = \Sigma\mathbf{V}^{\mathbf{T}} \tag{2}$$

Using Σ, the energy/variance contribution of each principal component can be acquired as detailed in (*Mehta and Linares* 2017).

4 PCA Results

We begin by perform PCA on the entire dataset (2000-2020) to get insight into the general density formulations. Then, we look into specific conditions, such as solar maximum and solar minimum, where different processes drive the global density variations.

4.1 2000-2020 Analysis

Figure 1 shows both the individual and cumulative variance captured by the first 20 principal components (PCs). In the left subplot, it becomes clear that the contribution of the first PC for both models is significant, capturing over 60% of the system's total variance. More importantly, the first PC for JB2008 captures over 10% more variance than it does for HASDM. There is also more variance captured by the second PC for JB2008, but beyond that, the individual variance captured is marginally greater for HASDM. This is due to the $\sim$ 75% captured in the first PC for JB2008.

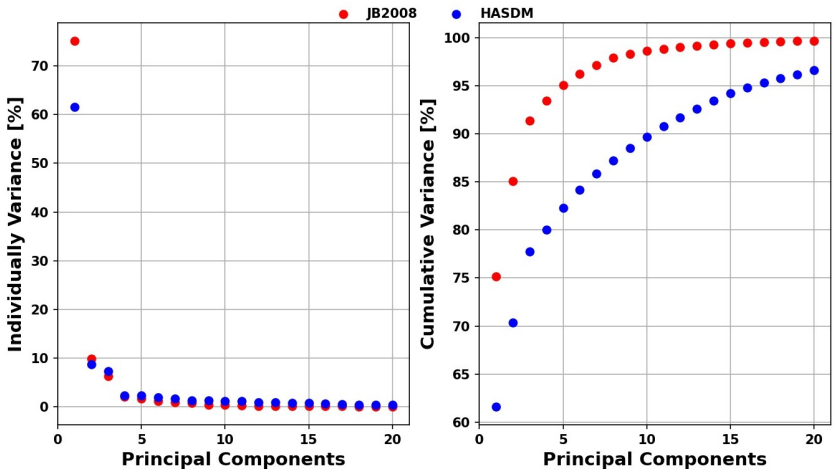


Figure 1. Individual (left) and cumulative (right) variance captured by the first 20 principal components for the two density datasets.

The cumulative variance is shown in the right subplot. It is abundantly clear that the variance captured by JB2008 is more substantial than by HASDM. Since JB2008 is an empirical model, there are well-defined relationships between the various drivers and the

overall system. Therefore, PCA is able to capture those processes in fewer PCs. HASDM, being an assimilative model, produces densities that account for an increased number of processes, making its structure more complex and more difficult to be captured by PCA.

Figures 2, 3, and 4 display the first five principal components (from **U**) for both models at 400 km, the first ten PCA scores (commonly referred to as temporal coefficients), and the results from a Pearson's correlation coefficient analysis (*Schober et al.* 2018) between the scores and drivers, respectively.

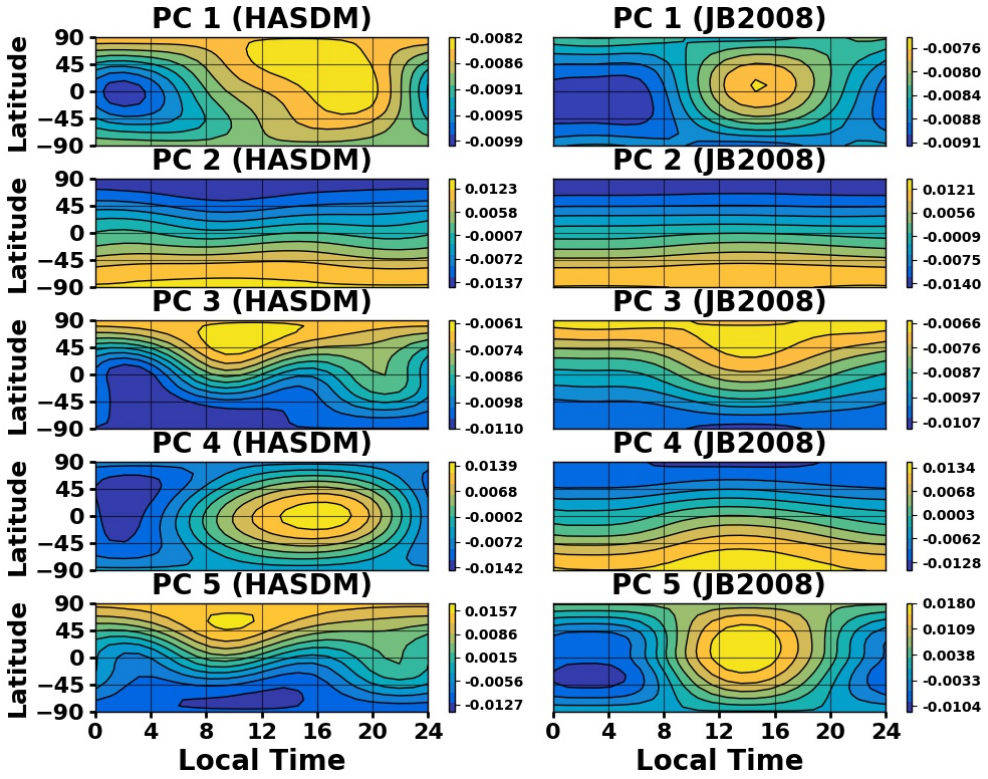


Figure 2. First five principal components for HASDM (left) and JB2008 (right) at 400 km.

The first PC for both models represents solar heating. There is peak around 14 hours local time and a minimum at 2 hours. Looking at this in relation to Figure 1, we can deduce that the effect of solar heating contributes significantly more to JB2008 than it does for HASDM. From Figure 4, the first coefficient is highly correlated to F_{10} and the other solar indices; there is a 90% or greater correlation to all four solar indices/proxies and their centered averages. This explains the larger magnitude at solar maximum around 2001 compared to around 2013. What also stands out is the not insignificant correlation to the

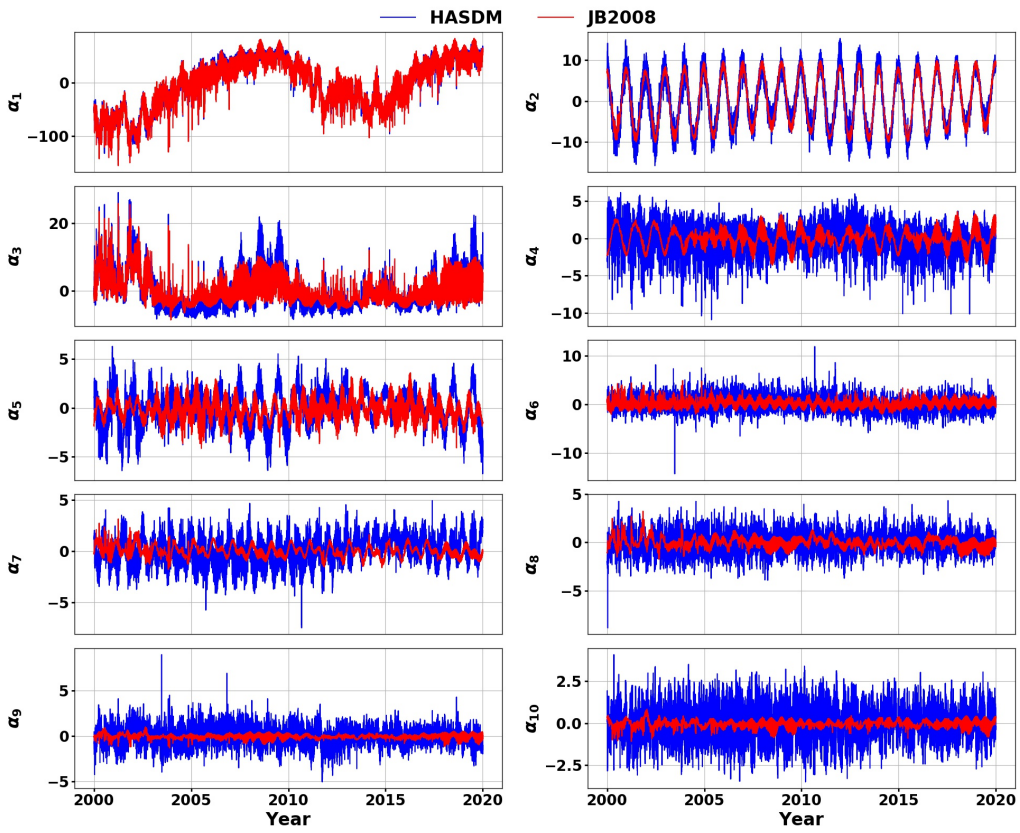


Figure 3. PCA scores/coefficients for both HASDM and JB2008 across two solar cycles.

geomagnetic indices. There are large spikes that coincide with events such as the Halloween storm of 2003 and the St. Patrick's day storm of 2017.

The second PC likely represents annual variations. It shows how the density varies fairly linearly with respect to latitude. Based on the day of year, this PC can change in intensity and orientation. This is caused by the sinusoidal trend of α_2 with a period of $\sim$ 365 days. The signs of the PCs and scores are important as they are multiplied for their contribution.

The third PC is representing the same process(es) between the models. It is important to note that PCA does not guarantee that a each PC corresponds to a single process. Based on the first two, there is an evident dominant process representing it, but this is not always the case. The third coefficient is the most complex of the first three. During solar maximum, the trends seem to mimic coefficient 1, but its magnitude increases again dur-

HASDM(2000 – 2020)

	α_1	α_2	α_3	α_4	α_5	α_6	α_7	α_8	α_9	α_{10}
F_{10}	-0.91	-0.04	0.20	0.15	-0.02	0.01	0.03	-0.00	0.04	0.00
F_{81c}	-0.94	-0.04	0.16	0.15	-0.00	0.02	0.02	-0.01	0.02	-0.01
S_{10}	-0.96	-0.09	0.07	0.17	0.01	0.02	0.04	0.00	0.02	0.00
S_{81c}	-0.95	-0.09	0.04	0.16	0.02	0.03	0.03	-0.00	0.01	-0.01
M_{10}	-0.94	-0.09	0.16	0.16	-0.00	0.00	0.06	-0.01	0.03	0.01
M_{81c}	-0.94	-0.09	0.12	0.14	0.01	0.01	0.06	-0.01	0.01	-0.01
Y_{10}	-0.94	-0.10	0.02	0.15	0.03	0.01	0.08	-0.02	-0.02	-0.00
Y_{81c}	-0.94	-0.10	0.00	0.14	0.04	0.01	0.07	-0.02	-0.02	-0.02
ap	-0.26	-0.04	0.00	-0.47	-0.04	0.08	0.05	-0.01	0.01	-0.03
Dst	0.34	-0.01	-0.01	0.51	0.06	0.15	0.10	0.01	-0.09	0.03
$semi_{sin}$	0.06	-0.01	0.01	0.04	0.01	0.09	-0.04	0.01	-0.01	-0.03
$semi_{cos}$	0.11	0.05	0.13	0.09	0.18	0.01	0.58	-0.06	0.05	-0.11
ann_{sin}	-0.04	-0.17	-0.03	-0.03	-0.06	0.00	-0.05	-0.09	0.00	-0.05
ann_{cos}	-0.14	0.94	-0.10	0.00	-0.14	0.00	-0.01	-0.02	0.01	0.00

JB2008(2000 – 2020)

	α_1	α_2	α_3	α_4	α_5	α_6	α_7	α_8	α_9	α_{10}
F_{10}	-0.90	-0.03	0.18	-0.01	-0.01	-0.19	-0.07	0.02	0.11	0.07
F_{81c}	-0.92	-0.04	0.13	-0.02	-0.02	-0.16	-0.03	-0.02	0.10	0.09
S_{10}	-0.95	-0.08	0.06	-0.01	0.01	-0.19	-0.02	0.01	0.06	0.05
S_{81c}	-0.94	-0.08	0.02	-0.01	0.02	-0.17	-0.00	-0.01	0.05	0.05
M_{10}	-0.93	-0.08	0.13	-0.01	-0.02	-0.18	-0.04	-0.01	0.08	0.07
M_{81c}	-0.93	-0.09	0.09	-0.02	-0.01	-0.15	-0.02	-0.04	0.07	0.08
Y_{10}	-0.93	-0.10	-0.00	-0.01	-0.03	-0.18	-0.06	0.00	0.04	0.03
Y_{81c}	-0.93	-0.10	-0.02	-0.01	-0.02	-0.15	-0.03	-0.02	0.04	0.04
ap	-0.25	-0.05	0.02	0.01	-0.18	0.31	0.06	0.01	-0.12	-0.14
Dst	0.35	-0.01	-0.08	0.01	0.21	-0.42	-0.07	-0.06	0.14	0.14
$semi_{sin}$	0.07	-0.00	-0.09	-0.10	-0.23	-0.49	0.57	-0.14	-0.01	-0.08
$semi_{cos}$	0.12	0.02	-0.03	-0.17	-0.69	-0.06	0.05	0.11	0.10	0.14
ann_{sin}	-0.06	-0.18	0.00	0.04	0.16	0.09	0.65	-0.33	0.00	0.07
ann_{cos}	-0.14	0.97	-0.07	0.02	0.03	0.01	0.12	-0.05	-0.03	0.05

Figure 4. Pearson correlation coefficients between all inputs and PCA scores for HASDM and JB2008 between 2000 and 2020. The sum of the absolute values for HASDM and JB2008 are 16.95 and 20.70, respectively. The colors represent the correlation coefficients with blue being -1.0, white being 0.0, and red being 1.0.

ing solar minimum. This increase no longer seems to relate to the first coefficient. Based on the correlation values, it can be deduced that the third PC is mostly representative of solar activity in regards to the entire time period. The combination of processes captured by this principal component makes linear correlation analysis difficult, especially over this time span. Looking at shorter windows (see the next two subsections) uncovers different processes captured by α_3, depending on the conditions.

In Supplementary Material SM1, the movement of the peak in α_3 provides insight to HASDM. At lower altitudes (250-450 km), the HASDM peak has a $\sim$ 4 hour shift relative to JB2008 which is seen in Figure 2. The JB2008 peak is located between 12 and 14 hours local time while HASDM's are around 9 hours. Beyond 450 km, the peak in both models shifts to 2 hours local time and towards the equator. They exhibit similar trends up to 825 km which hints at the reliance on HASDM's background model when the signal decreases at higher altitudes. We suspect that this represents the Winter helium bulge based on the local time of the peak (*Keating and Prior* 1968; *Reber and Hays* 1973).

The last two PCs shown in Figure 2 for the two models are flipped, meaning the fourth for HASDM has the same source as the fifth for JB2008 and vice versa. There is only a 2.25% difference between PCs 4 and 5 for HASDM which signifies that their respective contribution to the system's overall variance is similar. PC 4 for HASDM and 5 for JB2008 seem to be further effects of solar heating, while the other two are similar to PC 3. α_4 for HASDM and α_5 for JB2008 both have some correlation to geomagnetic activity, while α_5 for JB2008 has moderately strong correlation to the semiannual cosine wave. In Figure 2, the difference between PCs 3 and 5 for HASDM and PCs 3 and 4 for JB2008 is the location of the peak present in either the northern or southern poles.

Note that in the caption, the absolute sum for both models is shown, and the value for JB2008 is over 20% larger than for HASDM. This is as expected due to the simplicity in the density formulation of an empirical model relative to an assimilative one. Beyond these discussed principal components and scores, there is not an abundance of information. These correlation coefficients were generated using the entire twenty year period. However, looking at the coefficient plots and correlations for a single year shows that there is more information that cannot be seen by looking at such a broad period.

4.2 Solar Maximum

Figure 5 shows the first ten PCA scores for both models during 2001 (solar maximum), and Figure 6 contains the corresponding correlation coefficients.

The first coefficients for both models in Figure 5 are nearly identical, alluding to the similarities in how solar drivers impact the resulting density grids in both models. There is lower correlation between α_1 and the solar drivers relative to the twenty-year analysis. There is now increased correlation with the other drivers (e.g. geomagnetic and temporal).

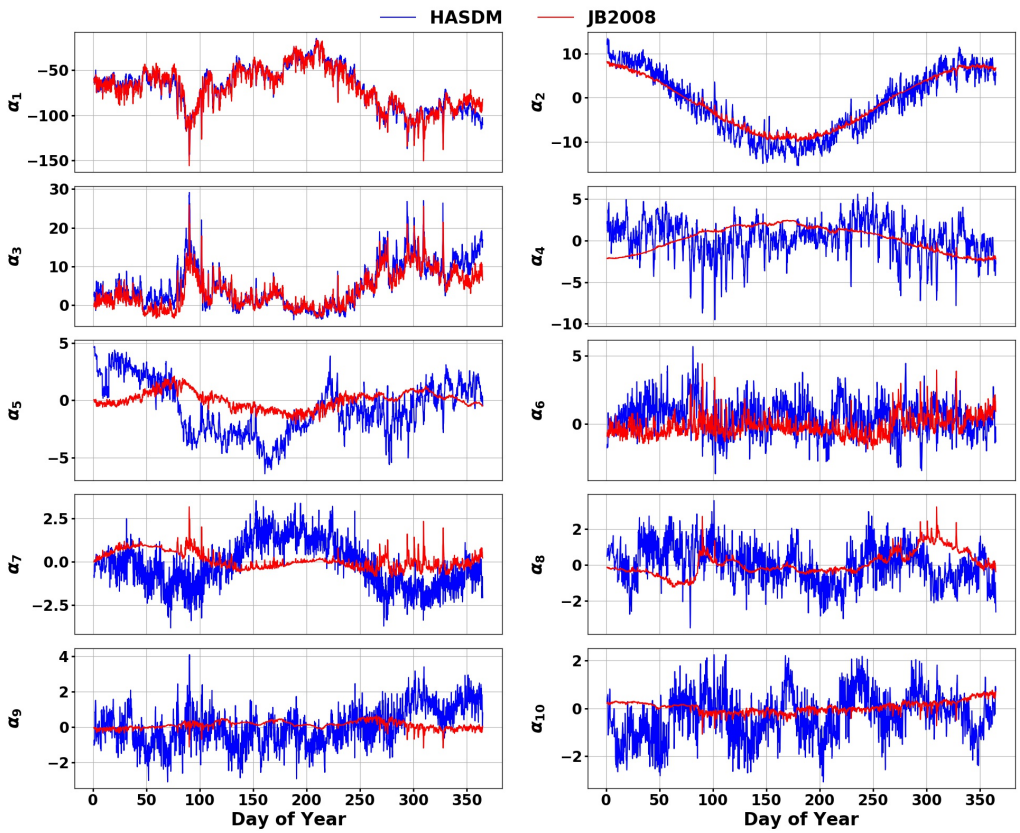


Figure 5. PCA coefficients for both HASDM and JB2008 for 2001.

The geomagnetic indices have more correlation during this year, because there is a larger portion of the time period with moderate or high geomagnetic activity compared to the complete period. There are also increases in correlation for the temporal drivers, but it is not entirely clear if that is coincidental with how solar heating varied over this particular year.

α_2 again shows a distinct annual trend, but there are some differences highlighted by the shortened time period. The variation for HASDM seems to contain additional processes, identified by the more complex structure within the mean annual trend. There is a strong correlation with the annual cosine wave, as was the case for the entire period. Similar to α_1, there are moderate correlations with all of the solar drivers. This is again likely a byproduct of how the solar drivers varied over the year. This is reinforced by the correlation values for JB2008 combined with the near-perfect cosine wave seen in Figure 5.

HASDM(2001)

	α_1	α_2	α_3	α_4	α_5	α_6	α_7	α_8	α_9	α_{10}
F_{10}	-0.77	0.37	0.77	-0.17	0.01	-0.06	-0.48	-0.01	0.54	0.27
F_{81c}	-0.75	0.52	0.72	-0.18	0.18	-0.14	-0.59	-0.06	0.58	0.20
S_{10}	-0.86	0.48	0.86	-0.24	0.07	-0.08	-0.54	-0.01	0.61	0.24
S_{81c}	-0.77	0.63	0.73	-0.18	0.29	-0.12	-0.60	-0.07	0.60	0.14
M_{10}	-0.76	0.41	0.78	-0.20	0.05	-0.12	-0.41	-0.07	0.57	0.17
M_{81c}	-0.69	0.57	0.68	-0.16	0.26	-0.14	-0.49	-0.15	0.60	0.12
Y_{10}	-0.73	0.35	0.73	-0.14	-0.00	-0.02	-0.44	-0.01	0.51	0.30
Y_{81c}	-0.75	0.54	0.72	-0.17	0.21	-0.14	-0.60	-0.05	0.58	0.20
ap	-0.35	0.04	0.42	-0.54	-0.10	-0.04	-0.15	-0.02	0.14	0.10
Dst	0.56	-0.17	-0.57	0.64	0.10	0.30	0.44	-0.00	-0.33	-0.19
$semi_{sin}$	0.41	0.07	-0.41	0.34	0.41	0.15	0.19	0.09	-0.31	-0.06
$semi_{cos}$	0.40	-0.02	-0.31	0.09	0.16	-0.05	0.54	-0.35	-0.00	-0.29
ann_{sin}	0.21	-0.12	-0.26	-0.09	-0.05	0.13	0.06	0.26	-0.35	-0.21
ann_{cos}	-0.59	0.95	0.49	-0.08	0.74	0.03	-0.62	0.12	0.38	-0.11

JB2008(2001)

	α_1	α_2	α_3	α_4	α_5	α_6	α_7	α_8	α_9	α_{10}
F_{10}	-0.74	0.43	0.74	-0.35	0.40	0.28	-0.15	0.66	0.06	0.32
F_{81c}	-0.72	0.61	0.71	-0.59	0.36	0.39	-0.30	0.70	-0.14	0.57
S_{10}	-0.83	0.55	0.83	-0.48	0.38	0.40	-0.08	0.72	-0.05	0.41
S_{81c}	-0.73	0.72	0.70	-0.70	0.35	0.41	-0.23	0.67	-0.24	0.66
M_{10}	-0.73	0.48	0.76	-0.44	0.23	0.34	-0.14	0.68	-0.02	0.40
M_{81c}	-0.65	0.65	0.67	-0.67	0.21	0.40	-0.31	0.67	-0.23	0.66
Y_{10}	-0.71	0.41	0.69	-0.33	0.40	0.26	-0.13	0.60	0.06	0.29
Y_{81c}	-0.72	0.62	0.70	-0.60	0.38	0.38	-0.27	0.68	-0.15	0.58
ap	-0.30	0.01	0.34	0.06	0.08	0.41	0.28	0.24	-0.23	-0.24
Dst	0.60	-0.19	-0.64	0.10	-0.21	-0.67	-0.35	-0.51	0.41	0.21
$semi_{sin}$	0.42	0.00	-0.51	-0.02	0.05	-0.41	0.47	-0.57	-0.03	0.03
$semi_{cos}$	0.41	-0.01	-0.32	-0.16	-0.70	-0.03	-0.09	-0.26	-0.29	0.24
ann_{sin}	0.23	-0.16	-0.37	0.22	0.09	-0.07	0.49	-0.54	-0.07	-0.29
ann_{cos}	-0.57	0.98	0.35	-0.96	0.51	0.22	0.38	0.21	-0.58	0.74

Figure 6. Pearson correlation coefficients between all inputs and PCA scores for HASDM and JB2008 only during 2001 (solar maximum). The sum of the absolute values for HASDM and JB2008 are 45.82 and 56.76, respectively. The colors represent the correlation coefficients with blue being -1.0, white being 0.0, and red being 1.0.

The third coefficient is nearly a mirror image of the first coefficient for both HASDM and JB2008. This was suspected when looking at the coefficient plot for the entire period, but it did not show up in the correlation values from Figure 4. Clearly, α_3 contains multiple processes that are difficult to observe from afar. The correlation values for α_3 with both models is almost exactly the opposite of the values for α_1 which reinforces the observation.

α_4 is quite interesting as it is representing different processes between the models and more than one for HASDM. For JB2008 it is inversely correlated to an annual cosine wave with little influence from another source. However, α_4 for HASDM shows some relation to a semiannual trend with influence from geomagnetic storms, having large spikes that coincide with those in α_1 which is reaffirmed by the correlation coefficients.

The fifth coefficient for HASDM has quite a peculiar set of correlation values. Its most significant correlations are to a semiannual sine wave and an annual cosine wave. Looking at Figure 5, it visually appears noisy, because the spikes do not correspond to spikes in the other coefficients. However, there are likely other contributions not captured by these drivers. For JB2008, there are non-negligible correlations with the solar drivers and the temporal cosine inputs.

Other interesting findings for 2001 include HASDM's α_7 and α_9 along with JB2008's α_8 and α_{10}. HASDM's seventh coefficient visually shows a strong semiannual trend. When looking at Figure 6, there is more correlation to the annual trend and most of the solar indices. This is peculiar, because the qualitative study of the coefficient shows much more signal in relation to a semiannual variation. After consulting the correlation coefficients for HASDM's α_9, the similarity to α_1 and α_3 becomes more clear. It seems to be largely influenced by solar heating but is less evident.

The eighth coefficient for JB2008 at first glance looks to have a semiannual quality, but after looking at the correlation values, its relationship to solar heating becomes more apparent. In fact, it looks like the third coefficient, only with a visually stronger signal. JB2008's α_{10} likely represents an annual cosine wave. However, it is difficult to see due to the y-axis bounds determined by the range of α_{10} for HASDM.

The most distinct difference between Figure 4 and 6 is the increase in the prevalence of color and therefore significant correlations. Nearly every coefficient for both models has at least one driver with 40% or greater correlation. Again, JB2008 has more cumulative correlation than HASDM, due to the increased processes in HASDM not modeled by the current set of drivers.

4.3 Solar Minimum

With the plethora of new information obtained by looking at the coefficients during a shorter period of solar maximum, it is important to see differences during solar minimum. Figure 7 shows the coefficients for both models during 2019, and Figure 8 contains the corresponding correlation coefficients.

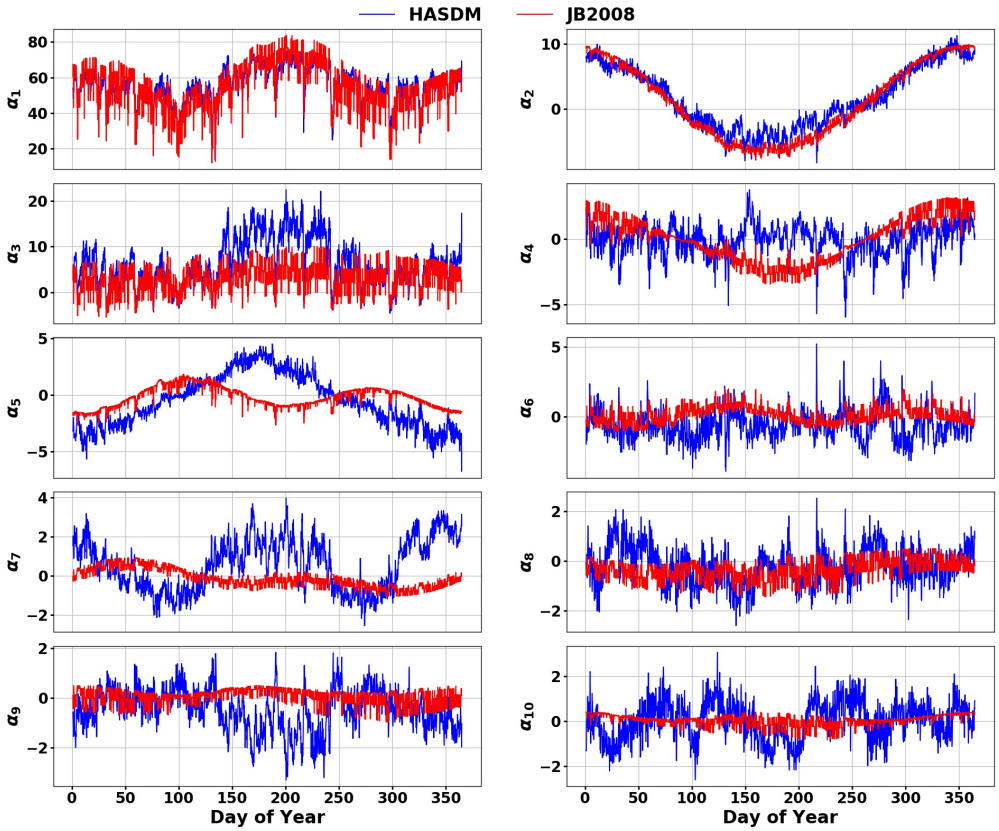


Figure 7. PCA coefficients for both HASDM and JB2008 for 2019.

Visually, some coefficients show stark contrasts to 2001. Primarily, α_1 for both models represents a noisy semiannual cosine wave. The apparent noise is even peculiar; it seems to jump between a fixed range about the mean for most of the year. Another notable difference between the years for α_1 is the change in sign. For 2001, all values were negative and mostly larger in magnitude, but the 2019 values are all positive. The correlations for α_1 deviate from the previous two cases. There is no longer a strong correlation between α_1 and solar heating for either model. In fact, the correlation with geomagnetic

HASDM(2019)

	α_1	α_2	α_3	α_4	α_5	α_6	α_7	α_8	α_9	α_{10}
F_{10}	-0.43	0.21	-0.45	0.09	-0.31	-0.15	-0.11	0.05	0.33	-0.05
F_{81c}	-0.43	0.36	-0.51	0.06	-0.46	-0.11	-0.12	0.03	0.31	0.06
S_{10}	-0.33	-0.25	-0.27	0.08	0.17	-0.14	-0.13	-0.14	0.28	-0.05
S_{81c}	-0.22	-0.34	-0.17	0.08	0.28	-0.08	-0.02	-0.28	0.13	0.06
M_{10}	-0.17	-0.32	-0.08	0.12	0.27	-0.10	-0.02	-0.18	0.12	-0.02
M_{81c}	0.00	-0.59	0.11	0.05	0.58	-0.04	0.08	-0.39	-0.06	0.11
Y_{10}	0.15	-0.28	0.24	0.16	0.34	-0.04	0.36	-0.20	-0.26	0.08
Y_{81c}	0.22	-0.36	0.32	0.13	0.44	0.01	0.39	-0.30	-0.31	0.06
ap	-0.45	-0.10	-0.36	-0.51	0.07	0.25	-0.15	0.12	0.26	-0.02
Dst	0.58	0.03	0.50	0.49	0.02	0.06	0.28	-0.13	-0.41	-0.07
$semi_{sin}$	0.22	0.03	0.19	-0.07	-0.04	0.05	-0.17	0.26	-0.04	-0.05
$semi_{cos}$	0.45	0.08	0.49	0.31	0.08	0.07	0.80	-0.01	-0.48	-0.31
ann_{sin}	-0.17	-0.19	-0.16	0.09	0.10	-0.10	-0.14	-0.09	0.19	-0.07
ann_{cos}	-0.37	0.95	-0.51	0.04	-0.95	-0.05	0.06	0.31	0.21	0.02

JB2008(2019)

	α_1	α_2	α_3	α_4	α_5	α_6	α_7	α_8	α_9	α_{10}
F_{10}	-0.37	0.23	-0.24	0.32	0.17	0.03	0.38	0.06	-0.09	0.31
F_{81c}	-0.43	0.37	-0.22	0.45	0.20	0.08	0.57	-0.09	-0.13	0.40
S_{10}	-0.33	-0.21	-0.16	-0.07	0.37	0.31	0.02	0.08	0.09	0.14
S_{81c}	-0.33	-0.34	-0.02	-0.17	0.55	0.52	0.02	-0.16	0.12	0.08
M_{10}	-0.16	-0.29	-0.02	-0.17	0.31	0.23	-0.17	0.06	0.05	-0.01
M_{81c}	-0.14	-0.60	0.14	-0.45	0.54	0.54	-0.36	-0.10	0.17	-0.16
Y_{10}	0.09	-0.29	0.22	-0.23	0.08	0.27	-0.58	0.07	0.07	-0.06
Y_{81c}	0.11	-0.38	0.27	-0.31	0.12	0.37	-0.72	0.09	0.11	-0.12
ap	-0.30	-0.07	-0.37	-0.06	-0.05	0.28	0.16	0.27	0.17	0.04
Dst	0.42	0.02	0.48	-0.00	0.06	-0.33	-0.22	-0.39	-0.19	-0.13
$semi_{sin}$	0.35	0.01	-0.10	-0.13	-0.39	-0.59	0.55	-0.17	0.03	-0.20
$semi_{cos}$	0.54	0.04	0.11	-0.08	-0.76	-0.13	-0.06	-0.10	0.30	0.26
ann_{sin}	-0.17	-0.20	-0.18	-0.13	0.27	0.10	0.70	-0.39	0.16	0.10
ann_{cos}	-0.29	0.98	-0.15	0.94	-0.31	-0.20	0.20	0.29	-0.36	0.57

Figure 8. Pearson correlation coefficients between all inputs and PCA scores for HASDM and JB2008 only during 2019 (solar minimum). The sum of the absolute values for HASDM and JB2008 are 30.02 and 34.10, respectively. The colors represent the correlation coefficients with blue being -1.0, white being 0.0, and red being 1.0.

activity is more pronounced than solar heating. The correlation with the semiannual cosine wave is the strongest of the inputs for JB2008 and one of the strongest for HASDM.

During this solar minimum, there was very low solar activity and few active regions giving the usual 27-day signal a longer 3-4 month signal, so other processes begin to dominate. This explains why the PCA coefficients are drastically dissimilar to those of solar maximum, and why other drivers become more pronounced. Looking at the cumulative

correlation for the models, there is a substantial decrease from solar maximum; they are approximately 60% and 65% of the 2001 values for HASDM and JB2008, respectively.

α_2 for both models is again a clear annual cosine wave. In comparison to 2001, the values and variance for the two models are more aligned. There is a strong positive correlation to the annual cosine wave, as was the case for the other two periods.

For both HASDM and JB2008, α_3 is difficult to visually discern. There are no significant long-term variations for JB2008. The only drivers with correlations greater than 30% are the two geomagnetic indices. The general appearance of α_3 is similar for HASDM but it has a relatively sharp increase about 1/3 through the year, and a drop around 2/3 through the year. HASDM has similar correlation values for the geomagnetic indices but also has noteworthy correlations to the temporal drivers and to F_{10}.

α_4 for HASDM has no apparent long-term trend (similar to JB2008's α_3), but it is moderately correlated to the geomagnetic drivers. For JB2008, there is a strong positive correlation to the annual cosine wave, which can be seen in the coefficient's trend. α_5 has temporal trends for both models, but it is semiannual for JB2008 and annual for HASDM. Both show a strong signal.

All the higher order coefficients are either temporal variations or indistinct, with the exception of α_7 for JB2008. It appears to have a semiannual component, but the baseline seems to decrease linearly. Oddly, α_7 shows to be more strongly correlated to the annual sine wave than the semiannual sine wave. This coefficient also has a 72% negative correlation to Y_{81c}. However, the variation's source is not definitive.

The largest correlation to any solar driver for HASDM is the correlation to M_{81c} in α_2 and α_5, but that is likely a consequence of those two coefficients being highly correlated to the annual cosine wave. The only other notable connection for HASDM is the strong positive correlation to the semiannual cosine wave in α_7. In fact, the three strongest correlations for HASDM are to temporal drivers.

5 CHAMP & GRACE Density Estimates

The CHAMP and GRACE datasets used in this study originate from Mehta et al. (2017). Both satellites have near polar orbits, covering nearly all latitudes, and over their respective lifetimes, CHAMP and GRACE datasets cover altitudes ranging from 300-535

km. This, in conjunction with the date range covered by the satellites, makes their density estimates invaluable for model comparison. Figure 9 shows altitudes each dataset covers along with orbit-averaged densities over their mission spans. This study only included GRACE-A data due to similarities between the twin satellites' orbits.

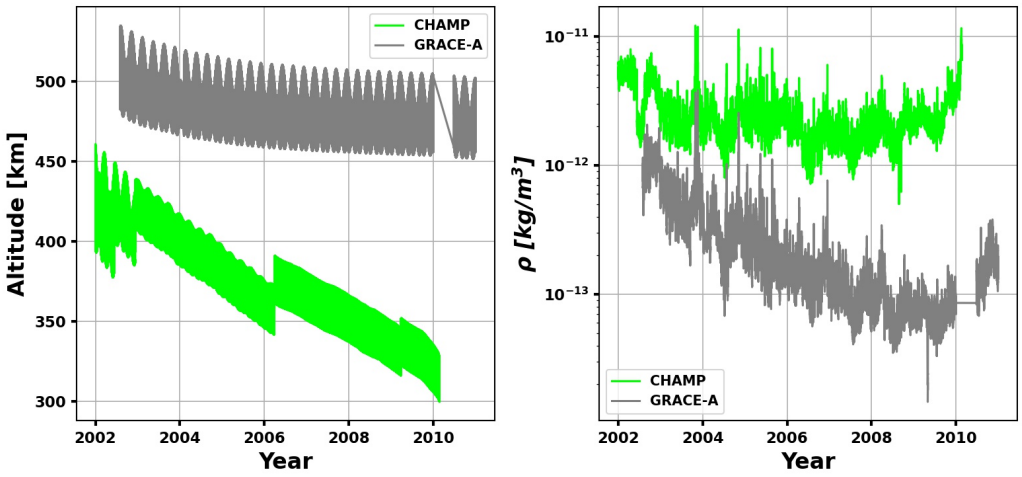


Figure 9. Altitude (left) and orbit-averaged densities (right) for CHAMP and GRACE-A.

There is minimal overlap between the altitudes of the two satellites, resulting in the lower densities encountered by GRACE-A in the right panel. The orbit-averaged densities were computed using a centered window with a span of 90 minutes, approximately one orbit.

Both the CHAMP and GRACE-A datasets contain files for every day containing information such as GPS time, solar local time (SLT), latitude, altitude and density. CHAMP has measurements every 10 seconds, while GRACE-A provides measurements every 5 seconds. In order to compare the satellite density estimates to the two models, we implement a trilinear interpolation algorithm using the global density grids from the models. Since the temporal resolution of the model densities are only every three hours, we maintain the same density grids over each three-hour period. The authors appended the existing CHAMP and GRACE density data of (*Mehta et al.* 2017) with the HASDM and JB2008 densities and have made them publicly available to the community (see Data Statement).

6 Model-Satellite Density Comparison

Figure 10 shows histograms of the log_{10} orbit-averaged densities for both satellites and models. HASDM and JB2008 overpredict relative to CHAMP and GRACE-A. However, the HASDM distributions have a marginally smaller bias for both satellites. The shape of the HASDM distributions more closely matches the high fidelity CHAMP and GRACE-A estimates with the smaller peaks being present on the right side. The JB2008 has similar distributions to both satellites but appears more generalized.

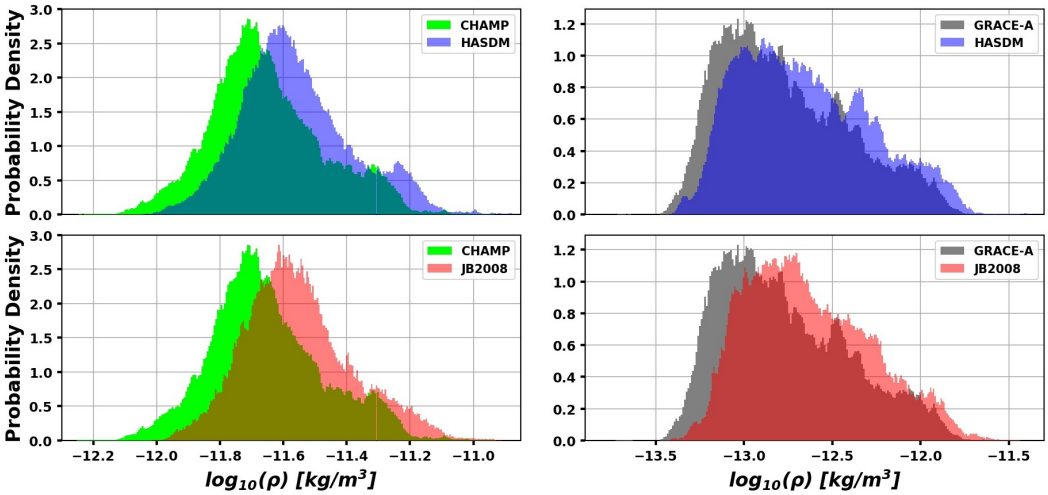


Figure 10. Histograms for CHAMP (left) and GRACE-A (right) orbits and corresponding densities from HASDM (top) and JB2008 (bottom). Values are centered averages with a window of 90 minutes, approximately one orbit.

Having spatiotemporally matched model densities to every measurement, mean absolute differences were computed for both the orbit and orbit-averaged densities. The results are shown in Table 1, broken down by year.

HASDM densities more closely match both CHAMP and GRACE-A estimates overall. However, there are years for both satellites where JB2008 predicts densities closer to the satellite estimates. The general trend is increased similarities to the satellite densities towards solar maximum, which decreases towards solar minimum. The models have lower percent differences when looking at orbit-averaged values, because they are tracking general density trends much better than the short period disturbances. The decrease in the density differences range from $3-13\%$ and $4-15\%$ for HASDM and JB2008, respec-

Table 1. Mean absolute percent difference error statistics for both models and satellites. "OA" denotes orbit-averaged.

	CHAMP		CHAMP (OA)		GRACE-A		GRACE-A (OA)	
Year	**HASDM**	**JB2008**	**HASDM**	**JB2008**	**HASDM**	**JB2008**	**HASDM**	**JB2008**
All	29.35%	33.96%	23.81%	27.63%	41.80%	50.52%	30.84%	39.75%
2002	22.10%	24.31%	19.31%	20.37%	27.70%	29.66%	23.03%	24.50%
2003	23.04%	23.26%	19.48%	18.58%	35.91%	36.97%	29.64%	28.81%
2004	25.23%	22.88%	21.57%	17.88%	41.05%	38.04%	34.22%	30.45%
2005	28.24%	33.29%	23.92%	28.52%	45.06%	57.95%	37.02%	50.45%
2006	29.72%	39.85%	24.47%	34.38%	36.90%	53.41%	30.67%	48.31%
2007	31.33%	37.99%	25.29%	31.77%	48.20%	63.89%	30.47%	45.20%
2008	41.31%	50.44%	28.07%	35.59%	40.65%	52.55%	24.67%	38.43%
2009	33.66%	41.80%	27.93%	35.62%	49.11%	63.35%	29.06%	45.25%
2010	32.33%	24.52%	28.05%	20.49%	44.69%	41.75%	37.44%	34.05%

tively when comparing the orbit to obit-averaged differences for CHAMP. For GRACE-A, the differences are more pronounced, being $5-20\%$ for HASDM and $5-19\%$ for JB2008. Considering the similarity in orbit inclination, this disagreement between the orbit and orbit-averaged differences is likely attributed to the altitude. To more closely examine the densities, we look at both active and quiet six-day periods (Figures 11 and 12 respectively). Figure 11 shows densities along CHAMP and GRACE-A orbits during the 2003 Halloween storm.

Looking at the storm time drivers, F_{10} starts off with substantial magnitude before decaying during the geomagnetic storms. a_p peaks at the maximum possible value of 400 twice during this period; it happening even once is an extremely rare occurrence. Outside of these storms, there is little geomagnetic activity. For the quiet period, F_{10} remains at solar minimum levels, and stays between 69 and 70 *sfu*. Concurrently, a_p varies continuously but never exceeds 12 (K_p = 3-). The densities spike on two occasions, coinciding with the two geomagnetic storms. When a_p initially reaches 400, the density responds to about half the magnitude increase of the first peak. a_p then drops before maintaining 300 *2nT* for a few hours; this is immediately before the density reaches its maximum value for this storm as there is a delayed density response.

There is a quicker overall response to the second storm, when a_p maintains its maximum value for a longer duration. Pre-storm conditioning could also cause the more abrupt

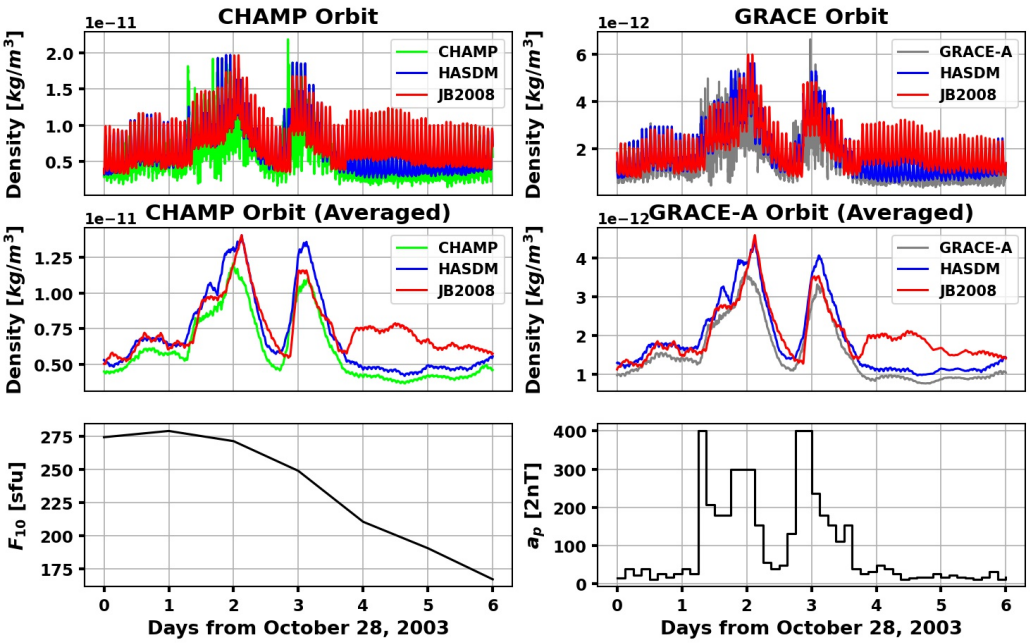
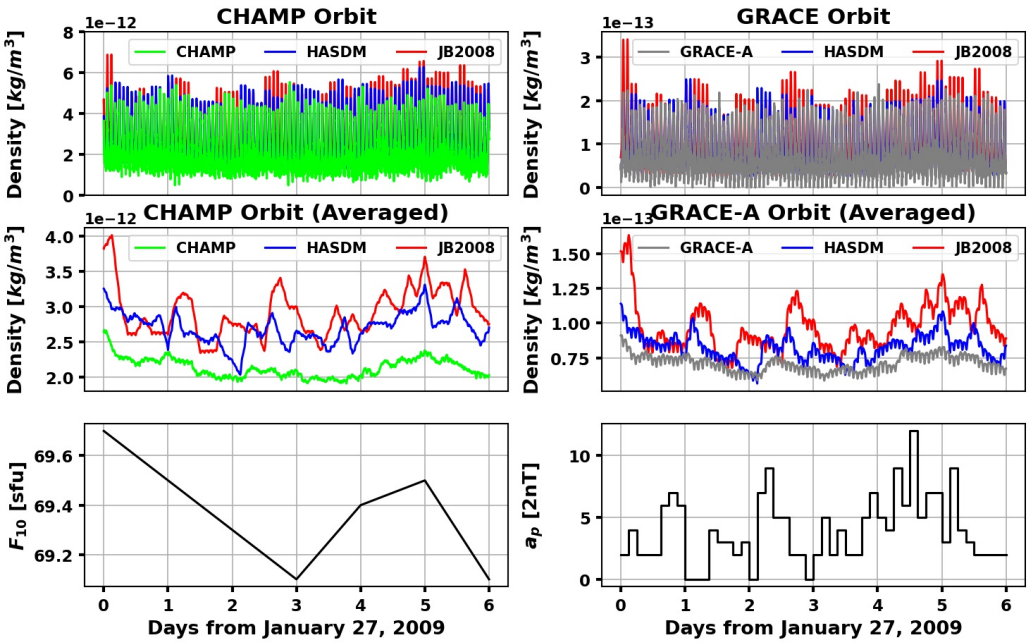


Figure 11. Densities from CHAMP, GRACE-A, HASDM, and JB2008 following a storm period in 2003. Center panels are orbit-averaged densities, and bottom panels show F_{10} and a_p.

density response. These trends are true for both satellite orbits. HASDM and JB2008 both have similar small biases before the storm, and HASDM over-predicts density relative to the satellites and JB2008.

However, the density recovery post-storm is modeled significantly closer to the satellite by HASDM than JB2008. This divergence may be caused by the NO production during the storm, a known cooling mechanism. This flows meridionally towards the equator in the days after the storm and is captured by HASDM but not modeled in JB2008. The mean differences for orbit-averaged densities with respect to CHAMP are 16.27% and 34.33% for HASDM and JB2008, respectively. Relative to GRACE-A, the mean differences are 23.09% and 46.58% for HASDM and JB2008, respectively. In Figure 12, the same information is presented for a quiet period in 2009.

In the orbit density plots, both models follow the oscillations well but have positive biases. Again, HASDM is tracking the satellite densities more closely than JB2008. This is confirmed in the orbit-averaged plots with HASDM being discernibly closer to the satellite densities. JB2008 predicts density closer to HASDM for CHAMP (lower al-



458 **Figure 12.** Densities from CHAMP, GRACE-A, HASDM, and JB2008 during a quiet period
459 in 2009. Center panels are orbit-averaged densities, and bottom panels show F_{10} and a_p.

464 titude) than for GRACE-A. Referring back to Figure 9, the difference in average altitude
465 for the satellites in 2009 is approximately 150 km. This explains the order of magni-
466 tude difference in densities and shows that HASDM more closely matches satellite esti-
467 mates at higher altitudes. The mean differences for orbit-averaged densities with respect
468 to CHAMP are 26.34% and 37.07% for HASDM and JB2008, respectively. Regarding to
469 GRACE-A, the mean differences are 15.81% and 36.28% for HASDM and JB2008, re-
470 spectively.

471 There are considerable peaks in the JB2008 densities, seen in all four panels. It is
472 more pronounced in the orbit-averaged density plots, particularly for the GRACE-A orbit.
473 These seem to be a response to the geomagnetic activity (seen in the bottom-right panel).
474 All JB2008 density peaks lag a_p spikes by about 12 hours and deviate from HASDM and
475 satellite densities. a_p is a good indicator of the source, because JB2008 uses a_p when no
476 storms are detected. JB2008 is overestimating the impact of the a_p fluctuations relative to
477 HASDM and considering the satellite densities.

7 Summary

In this work, we perform scientific investigation into HASDM and JB2008 densities by leveraging PCA, and we conduct an assessment of these models relative to CHAMP and GRACE-A density estimates. To analyze the model data, PCA was applied after normalization and centering. This resulted in a useful covariance matrix and time-dependent coefficients. The covariance matrix can be examined about any axis to identify spatial features that contribute significantly to the variance in the dataset. By looking at SLT-latitude slices at 400 km, we identified key contributions to the system's variance. The most important for both models was solar heating, followed by annual variations. The third mode primarily represented the impact of geomagnetic activity. The fourth and fifth modes were difficult to discern but were flipped between the models; this means that their relative importance is different.

Next, we explored the time-dependent coefficients, or PCA scores. There were striking similarities for the first three coefficients between models, but it becomes challenging to compare for the higher-order coefficients. In the correlation analysis, most driver-coefficient combinations produced weak correlations, with the exception of α_1. To investigate the processes at solar maximum, the window was limited to 2001. This resulted in more distinct trends in the coefficient plots, highlighting the effects of solar and geomagnetic activity along with annual and semiannual trends. JB2008 had more evident variations and subsequently higher correlations to the drivers. The HASDM coefficients had weaker signals and correlations. Relative to the 20-year analysis, the correlation figures displayed stronger correlations across a majority of coefficients. The solar drivers had non-negligible correlations with most of the coefficients due to the variance caused by EUV irradiance.

This study was performed once more on 2019, to investigate the datasets during solar minimum. The coefficients had weaker signal compared to 2001, and the correlations seen in 2001 had changed drastically. The solar drivers had less impact on the variance, and the temporal drivers had more pronounced correlations.

Last, we compared the two models to CHAMP and GRACE-A accelerometer-derived density estimates. The HASDM density distribution more closely match those of CHAMP and GRACE-A, yet both models had similar biases relative to the satellite densities. Mean absolute differences were assessed across the datasets along with yearly values. Overall,

HASDM's predictions were closer to CHAMP's estimates (orbit=29.35% & averaged=23.81%) than JB2008's (orbit=33.96% & averaged=27.63%). This observation was also true with respect to GRACE-A: HASDM (orbit=41.80% & averaged=30.84%) JB2008 (orbit=50.52% & averaged=39.75%).

We looked at both storm and quiet periods to see how well the models tracked satellite densities on a shorter time-scale. In general, HASDM tracked both satellites' estimates better, particularly for the orbit-averaged densities. Over the two geomagnetic storms in the 2003 Halloween storm, HASDM predicted higher peak densities than JB2008 or the satellites, but it modeled the pre- and post-storm densities well. There were spikes for JB2008's density predictions during the solar minimum/quiet period that showed an overestimation of a_p fluctuations on density.

8 Future Work

In the future, we plan to develop machine-learned (ML) models on the PCA coefficients using various drivers. Not only will it generate a computationally efficient predictive model, it will allow us to perform nonlinear analysis into the contribution of these and additional drivers to the PCA scores. These models can leverage ML techniques to also model uncertainty in the system (*Licata et al.* 2021).

9 Data Statement

The JB2008 model is available for download at at `https://spacewx.com/jb2008/`, and requests can be submitted for access to the SET HASDM density database at `https://spacewx.com/hasdm/`. The historical space weather indices used in this study can also be found at the JB2008 link.

Original CHAMP and GRACE density estimates from (*Mehta et al.* 2017) can be found at `http://tinyurl.com/densitysets`. As a product of this work, we appended the HASDM and JB2008 density estimates to those files. These updated files can be found at `https://zenodo.org/record/4602380#.YEwEw-1KhuU`

10 Acknowledgements

This work was made possible by NASA Established Program to Stimulate Competitive Research, Grant #80NSSC19M0054. SET and WVU gratefully acknowledge support

from the NASA SBIR contract #80NSSC20C0292 for Machine learning Enabled Thermosphere Advanced by HASDM (META-HASDM).

References

Bettadpur, S. (2012), Gravity Recovery and Climate Experiment: Product Specification Document, *GRACE 327-720, CSR-GR-03-02*, cent. for Space Res., The Univ. of Texas, Austin, TX.

Bowman, B., W. K. Tobiska, F. Marcos, C. Huang, C. Lin, and W. Burke (2008), A New Empirical Thermospheric Density Model JB2008 Using New Solar and Geomagnetic Indices, in *AIAA/AAS Astrodynamics Specialist Conference*, AIAA 2008-6438.

Bruinsma, S., C. Boniface, E. K. Sutton, and M. Fedrizzi (2021), Thermosphere modeling capabilities assessment: geomagnetic storms, *J. Space Weather Space Clim.*, *11*, 12, doi:10.1051/swsc/2021002.

Bruinsma, Sean (2015), The DTM-2013 thermosphere model, *J. Space Weather Space Clim.*, *5*, A1, doi:10.1051/swsc/2015001.

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

Covington, A. E. (1948), Solar Noise Observations on 10.7 Centimeters, *Proceedings of the IRE*, *36*(4), 454–457, doi:10.1109/JRPROC.1948.234598.

Deng, Y., T. J. Fuller-Rowell, A. J. Ridley, D. Knipp, and R. E. Lopez (2013), Theoretical study: Influence of different energy sources on the cusp neutral density enhancement, *Journal of Geophysical Research: Space Physics*, *118*(5), 2340–2349, doi:https://doi.org/10.1002/jgra.50197.

Doornbos, E. (2012), *Producing Density and Crosswind Data from Satellite Dynamics Observations*, pp. 91–126, Springer Berlin Heidelberg, Berlin, Heidelberg, doi:10.1007/978-3-642-25129-0_4.

Emmert, J. (2015), Thermospheric mass density: A review, *Advances in Space Research*, *56*, doi:10.1016/j.asr.2015.05.038.

Fedrizzi, M., T. J. Fuller-Rowell, and M. V. Codrescu (2012), Global joule heating index derived from thermospheric density physics-based modeling and observations, *Space Weather*, *10*(3), doi:https://doi.org/10.1029/2011SW000724.

Ganushkina, N., A. Jaynes, and M. Liemohn (2017), Space Weather Effects Produced by the Ring Current Particles, *Space Science Reviews*, *212*(3-4), 1315–1344, doi:10.1007/s11214-017-0412-2.

Gondelach, D. J., and R. Linares (2020), Real-Time Thermospheric Density Estimation via Two-Line Element Data Assimilation, *Space Weather*, *18*(2), e2019SW002,356, doi:https://doi.org/10.1029/2019SW002356, e2019SW002356 10.1029/2019SW002356.

Jacchia, L. G. (1970), New Static Models of the Thermosphere and Exosphere with Empirical Temperature Profiles, *SAO Special Report*, *313*.

Jolliffe, I. T., and J. Cadima (2016), Principal component analysis: a review and recent developments, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *374*(2065), 20150,202, doi:10.1098/rsta.2015.0202.

Keating, G. M., and E. J. Prior (1968), The winter helium bulge, in space research VIII, in *Proceedings of the Tenth Cospar Plenary Meeting*, p. 982âĂŞ992.

Kessler, D. J., and B. G. Cour-Palais (1978), Collision frequency of artificial satellites: The creation of a debris belt, *Journal of Geophysical Research: Space Physics*, *83*(A6), 2637–2646, doi:https://doi.org/10.1029/JA083iA06p02637.

Licata, R., P. Mehta, and W. K. Tobiska (2021), Impact of driver and model uncertainty on drag and orbit prediction, in *Proceedings of the 31st AAS/AIAA Space Flight Mechanics Meeting*.

Licata, R. J., W. K. Tobiska, and P. M. Mehta (2020), Benchmarking forecasting models for space weather drivers, *Space Weather*, *18*(10), e2020SW002,496, doi:https://doi.org/10.1029/2020SW002496, e2020SW002496 10.1029/2020SW002496.

Luhr, H., L. Grunwaldt, and C. Forste (2002), CHAMP Reference Systems, Transformations and Standards, *Tech. rep.*, CH-GFZ-RS-002, GFZ-Potsdam, Postdam, Germany.

Marcos, F., M. Kendra, J. Griffin, J. Bass, D. Larson, and J. J. Liu (1998), Precision Low Earth Orbit Determination Using Atmospheric Density Calibration, *Journal of The Astronautical Sciences*, *46*, 395–409.

McClain, W., and D. Vallado (2001), *Fundamentals of Astrodynamics and Applications*, Space Technology Library, 556-557 pp., Springer Netherlands.

Mehta, P. M., and R. Linares (2017), A methodology for reduced order modeling and calibration of the upper atmosphere, *Space Weather*, *15*(10), 1270–1287, doi:10.1002/2017SW001642.

Mehta, P. M., A. C. Walker, E. K. Sutton, and H. C. Godinez (2017), New density estimates derived using accelerometers on board the CHAMP and GRACE satellites, *Space Weather*, *15*(4), 558–576, doi:https://doi.org/10.1002/2016SW001562.

Mehta, P. M., R. Linares, and E. K. Sutton (2018), A Quasi-Physical Dynamic Reduced Order Model for Thermospheric Mass Density via Hermitian Space-Dynamic Mode Decomposition, *Space Weather*, *16*(5), 569–588, doi:10.1029/2018SW001840.

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

Nazarenko, A., P. Cefola, and V. Yurasov (1998), Estimating atmospheric density variations to improve LEO orbit prediction accuracy, in *AIAA/AAS Space Flight Mechanics Meeting*, AAS 98-190.

Picone, J. M., A. E. Hedin, D. P. Drob, and A. C. Aikin (2002), NRLMSISE-00 empirical model of the atmosphere: Statistical comparisons and scientific issues, *Journal of Geophysical Research: Space Physics*, *107*(A12), SIA 15–1–SIA 15–16, doi:10.1029/2002JA009430.

Qian, L., and S. Solomon (2011), Thermospheric density: An overview of temporal and spatial variations, *Space Science Reviews - SPACE SCI REV*, *168*, 1–27, doi:10.1007/s11214-011-9810-z.

Reber, C. A., and P. B. Hays (1973), Thermospheric wind effects on the distribution of helium and argon in the Earth's upper atmosphere, *Journal of Geophysical Research (1896-1977)*, *78*(16), 2977–2991, doi:https://doi.org/10.1029/JA078i016p02977.

Schober, P., C. Boer, and L. Schwarte (2018), Correlation Coefficients: Appropriate Use and Interpretation, *Anesthesia & Analgesia*, *126*, 1763–1768, doi:10.1213/ANE.0000000000002864.

Storz, M., B. Bowman, and J. Branson (2005), High Accuracy Satellite Drag Model (HASDM), in *AIAA/AAS Astrodynamics Specialist Conference and Exhibit*, doi:10.2514/6.2002-4886.

Sutton, E. K. (2008), Effects of solar disturbances on the thermosphere densities and winds from CHAMP and GRACE satellite accelerometer data, Ph.D. thesis, University of Colorado at Boulder.

Tobiska, W., T. Woods, F. Eparvier, R. Viereck, L. Floyd, D. Bouwer, G. Rottman, and O. White (2000), The SOLAR2000 empirical solar irradiance model and forecast tool,

635 *Journal of Atmospheric and Solar-Terrestrial Physics*, *62*(14), 1233 – 1250, doi:https:
636 //doi.org/10.1016/S1364-6826(00)00070-5.

637 Tobiska, W. K., S. D. Bouwer, and B. R. Bowman (2008a), The development of new solar
638 indices for use in thermospheric density modeling, *Journal of Atmospheric and Solar-*
639 *Terrestrial Physics*, *70*(5), 803–819, doi:https://doi.org/10.1016/j.jastp.2007.11.001.

640 Tobiska, W. K., B. Bowman, and S. D. Bouwer (2008b), *Solar and Geomagnetic Indices*
641 *for the JB2008 Thermosphere Density Model*, chap. 4, COSPAR CIRA Draft.

642 Tobiska, W. K., D. Knipp, W. J. Burke, D. Bouwer, J. Bailey, D. Odstrcil, M. P. Hagan,
643 J. Gannon, and B. R. Bowman (2013), The Anemomilos prediction methodology for
644 Dst, *Space Weather*, *11*(9), 490–508, doi:10.1002/swe.20094.

645 Tobiska, W. K., B. R. Bowman, D. Bouwer, A. Cruz, K. Wahl, M. Pilinski, P. M. Mehta,
646 and R. J. Licata (2021), The SET HASDM density database, *Space Weather*, p.
647 e2020SW002682, doi:https://doi.org/10.1029/2020SW002682.

648 Vickers, H., M. J. Kosch, E. Sutton, L. Bjoland, Y. Ogawa, and C. La Hoz (2014), A solar
649 cycle of upper thermosphere density observations from the eiscat svalbard radar, *Jour-*
650 *nal of Geophysical Research: Space Physics*, *119*(8), 6833–6845, doi:https://doi.org/10.
651 1002/2014JA019885.

Figure 1.

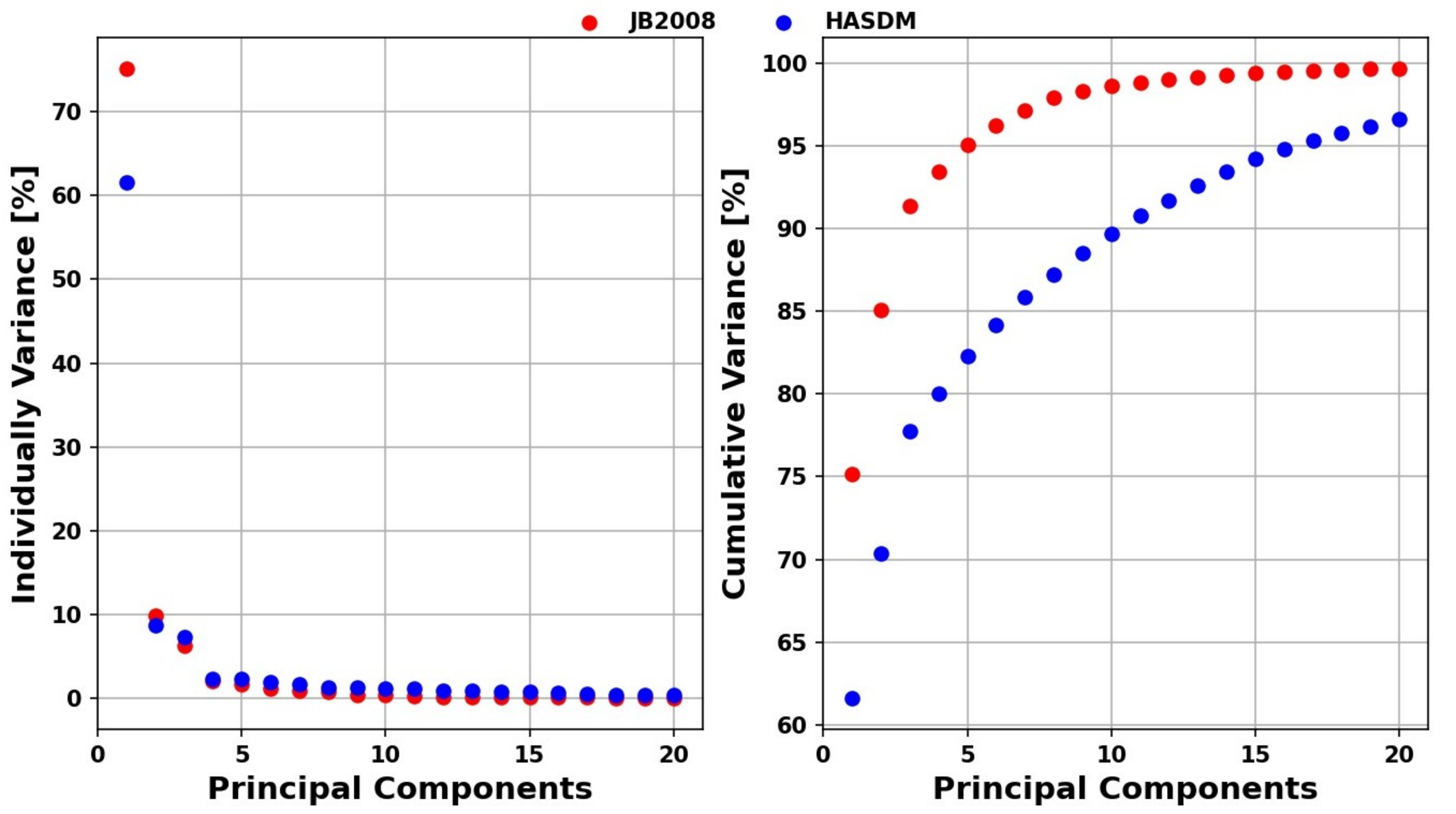

JB2008
HASDM
Individually Variance [%]
Cumulative Variance [%]
Principal Components
Principal Components
0
5
10
15
20
0
10
20
30
40
50
60
70
60
65
70
75
80
85
90
95
100

Figure 2.

PC 1 (HASDM)

PC 1 (JB2008)

PC 2 (HASDM)

PC 2 (JB2008)

PC 3 (HASDM)

PC 3 (JB2008)

PC 4 (HASDM)

PC 4 (JB2008)

PC 5 (HASDM)

PC 5 (JB2008)

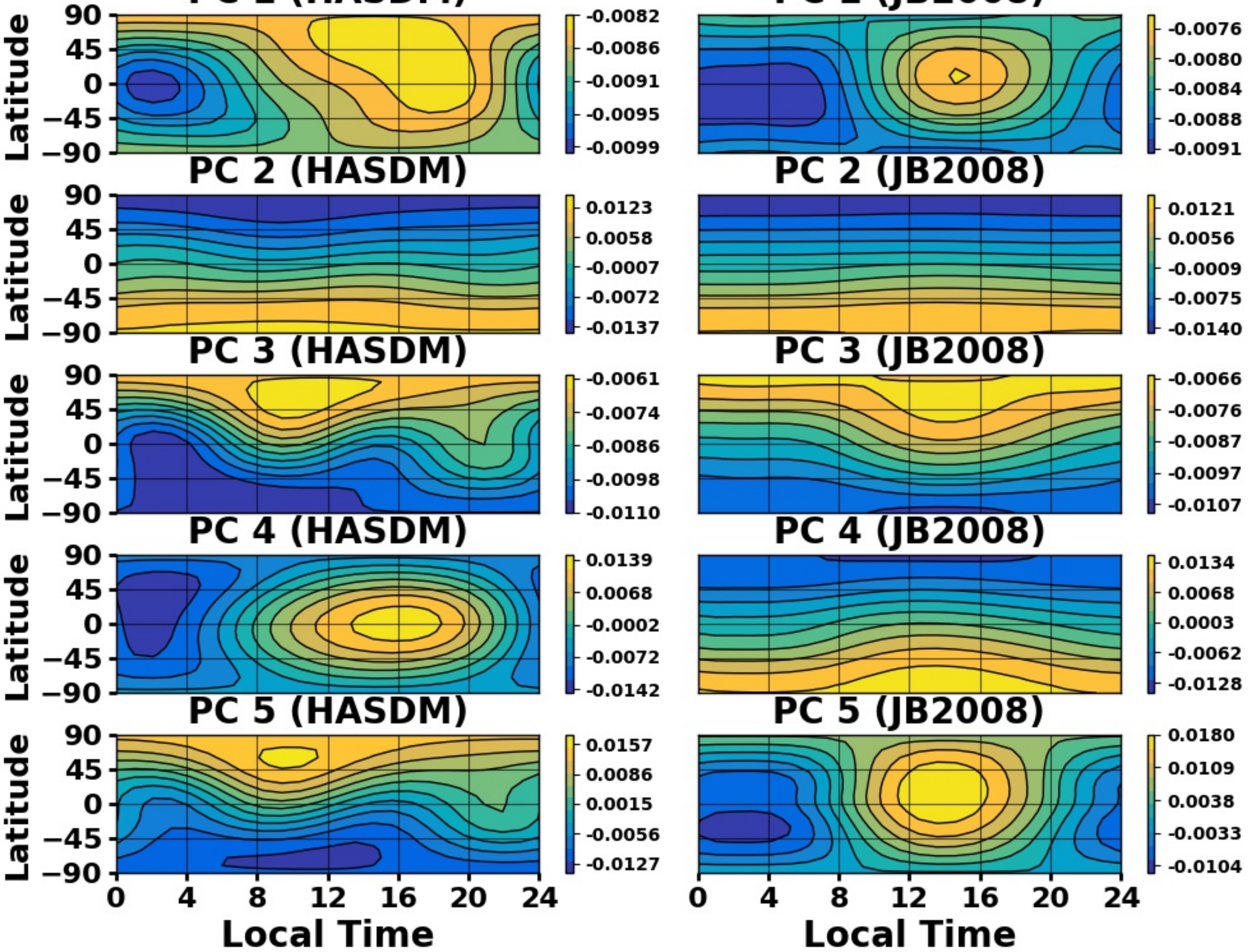


Figure 3.

HASDM JB2008

α_1 α_2 α_3 α_4 α_5 α_6 α_7 α_8 α_9 α_{10}

Year Year

Figure 4.

HASDM(2000 – 2020)

	α_1	α_2	α_3	α_4	α_5	α_6	α_7	α_8	α_9	α_{10}
F_{10}	-0.91	-0.04	0.20	0.15	-0.02	0.01	0.03	-0.00	0.04	0.00
F_{81c}	-0.94	-0.04	0.16	0.15	-0.00	0.02	0.02	-0.01	0.02	-0.01
S_{10}	-0.96	-0.09	0.07	0.17	0.01	0.02	0.04	0.00	0.02	0.00
S_{81c}	-0.95	-0.09	0.04	0.16	0.02	0.03	0.03	-0.00	0.01	-0.01
M_{10}	-0.94	-0.09	0.16	0.16	-0.00	0.00	0.06	-0.01	0.03	0.01
M_{81c}	-0.94	-0.09	0.12	0.14	0.01	0.01	0.06	-0.01	0.01	-0.01
Y_{10}	-0.94	-0.10	0.02	0.15	0.03	0.01	0.08	-0.02	-0.02	-0.00
Y_{81c}	-0.94	-0.10	0.00	0.14	0.04	0.01	0.07	-0.02	-0.02	-0.02
ap	-0.26	-0.04	0.00	-0.47	-0.04	0.08	0.05	-0.01	0.01	-0.03
Dst	0.34	-0.01	-0.01	0.51	0.06	0.15	0.10	0.01	-0.09	0.03
$semi_{sin}$	0.06	-0.01	0.01	0.04	0.01	0.09	-0.04	0.01	-0.01	-0.03
$semi_{cos}$	0.11	0.05	0.13	0.09	0.18	0.01	0.58	-0.06	0.05	-0.11
ann_{sin}	-0.04	-0.17	-0.03	-0.03	-0.06	0.00	-0.05	-0.09	0.00	-0.05
ann_{cos}	-0.14	0.94	-0.10	0.00	-0.14	0.00	-0.01	-0.02	0.01	0.00

JB2008(2000 – 2020)

	α_1	α_2	α_3	α_4	α_5	α_6	α_7	α_8	α_9	α_{10}
F_{10}	-0.90	-0.03	0.18	-0.01	-0.01	-0.19	-0.07	0.02	0.11	0.07
F_{81c}	-0.92	-0.04	0.13	-0.02	-0.02	-0.16	-0.03	-0.02	0.10	0.09
S_{10}	-0.95	-0.08	0.06	-0.01	0.01	-0.19	-0.02	0.01	0.06	0.05
S_{81c}	-0.94	-0.08	0.02	-0.01	0.02	-0.17	-0.00	-0.01	0.05	0.05
M_{10}	-0.93	-0.08	0.13	-0.01	-0.02	-0.18	-0.04	-0.01	0.08	0.07
M_{81c}	-0.93	-0.09	0.09	-0.02	-0.01	-0.15	-0.02	-0.04	0.07	0.08
Y_{10}	-0.93	-0.10	-0.00	-0.01	-0.03	-0.18	-0.06	0.00	0.04	0.03
Y_{81c}	-0.93	-0.10	-0.02	-0.01	-0.02	-0.15	-0.03	-0.02	0.04	0.04
ap	-0.25	-0.05	0.02	0.01	-0.18	0.31	0.06	0.01	-0.12	-0.14
Dst	0.35	-0.01	-0.08	0.01	0.21	-0.42	-0.07	-0.06	0.14	0.14
$semi_{sin}$	0.07	-0.00	-0.09	-0.10	-0.23	-0.49	0.57	-0.14	-0.01	-0.08
$semi_{cos}$	0.12	0.02	-0.03	-0.17	-0.69	-0.06	0.05	0.11	0.10	0.14
ann_{sin}	-0.06	-0.18	0.00	0.04	0.16	0.09	0.65	-0.33	0.00	0.07
ann_{cos}	-0.14	0.97	-0.07	0.02	0.03	0.01	0.12	-0.05	-0.03	0.05

Figure 5.

HASDM JB2008

α_1 α_2 α_3 α_4 α_5 α_6 α_7 α_8 α_9 α_{10}

Day of Year Day of Year

Figure 6.

HASDM(2001)

	α_1	α_2	α_3	α_4	α_5	α_6	α_7	α_8	α_9	α_{10}
F_{10}	-0.77	0.37	0.77	-0.17	0.01	-0.06	-0.48	-0.01	0.54	0.27
F_{81c}	-0.75	0.52	0.72	-0.18	0.18	-0.14	-0.59	-0.06	0.58	0.20
S_{10}	-0.86	0.48	0.86	-0.24	0.07	-0.08	-0.54	-0.01	0.61	0.24
S_{81c}	-0.77	0.63	0.73	-0.18	0.29	-0.12	-0.60	-0.07	0.60	0.14
M_{10}	-0.76	0.41	0.78	-0.20	0.05	-0.12	-0.41	-0.07	0.57	0.17
M_{81c}	-0.69	0.57	0.68	-0.16	0.26	-0.14	-0.49	-0.15	0.60	0.12
Y_{10}	-0.73	0.35	0.73	-0.14	-0.00	-0.02	-0.44	-0.01	0.51	0.30
Y_{81c}	-0.75	0.54	0.72	-0.17	0.21	-0.14	-0.60	-0.05	0.58	0.20
ap	-0.35	0.04	0.42	-0.54	-0.10	-0.04	-0.15	-0.02	0.14	0.10
Dst	0.56	-0.17	-0.57	0.64	0.10	0.30	0.44	-0.00	-0.33	-0.19
$semi_{sin}$	0.41	0.07	-0.41	0.34	0.41	0.15	0.19	0.09	-0.31	-0.06
$semi_{cos}$	0.40	-0.02	-0.31	0.09	0.16	-0.05	0.54	-0.35	-0.00	-0.29
ann_{sin}	0.21	-0.12	-0.26	-0.09	-0.05	0.13	0.06	0.26	-0.35	-0.21
ann_{cos}	-0.59	0.95	0.49	-0.08	0.74	0.03	-0.62	0.12	0.38	-0.11

JB2008(2001)

	α_1	α_2	α_3	α_4	α_5	α_6	α_7	α_8	α_9	α_{10}
F_{10}	-0.74	0.43	0.74	-0.35	0.40	0.28	-0.15	0.66	0.06	0.32
F_{81c}	-0.72	0.61	0.71	-0.59	0.36	0.39	-0.30	0.70	-0.14	0.57
S_{10}	-0.83	0.55	0.83	-0.48	0.38	0.40	-0.08	0.72	-0.05	0.41
S_{81c}	-0.73	0.72	0.70	-0.70	0.35	0.41	-0.23	0.67	-0.24	0.66
M_{10}	-0.73	0.48	0.76	-0.44	0.23	0.34	-0.14	0.68	-0.02	0.40
M_{81c}	-0.65	0.65	0.67	-0.67	0.21	0.40	-0.31	0.67	-0.23	0.66
Y_{10}	-0.71	0.41	0.69	-0.33	0.40	0.26	-0.13	0.60	0.06	0.29
Y_{81c}	-0.72	0.62	0.70	-0.60	0.38	0.38	-0.27	0.68	-0.15	0.58
ap	-0.30	0.01	0.34	0.06	0.08	0.41	0.28	0.24	-0.23	-0.24
Dst	0.60	-0.19	-0.64	0.10	-0.21	-0.67	-0.35	-0.51	0.41	0.21
$semi_{sin}$	0.42	0.00	-0.51	-0.02	0.05	-0.41	0.47	-0.57	-0.03	0.03
$semi_{cos}$	0.41	-0.01	-0.32	-0.16	-0.70	-0.03	-0.09	-0.26	-0.29	0.24
ann_{sin}	0.23	-0.16	-0.37	0.22	0.09	-0.07	0.49	-0.54	-0.07	-0.29
ann_{cos}	-0.57	0.98	0.35	-0.96	0.51	0.22	0.38	0.21	-0.58	0.74

Figure 7.

HASDM
JB2008
α_1
α_2
α_3
α_4
α_5
α_6
α_7
α_8
α_9
α_{10}
Day of Year
Day of Year

Figure 8.

HASDM(2019)

	α_1	α_2	α_3	α_4	α_5	α_6	α_7	α_8	α_9	α_{10}
F_{10}	-0.43	0.21	-0.45	0.09	-0.31	-0.15	-0.11	0.05	0.33	-0.05
F_{81c}	-0.43	0.36	-0.51	0.06	-0.46	-0.11	-0.12	0.03	0.31	0.06
S_{10}	-0.33	-0.25	-0.27	0.08	0.17	-0.14	-0.13	-0.14	0.28	-0.05
S_{81c}	-0.22	-0.34	-0.17	0.08	0.28	-0.08	-0.02	-0.28	0.13	0.06
M_{10}	-0.17	-0.32	-0.08	0.12	0.27	-0.10	-0.02	-0.18	0.12	-0.02
M_{81c}	0.00	-0.59	0.11	0.05	0.58	-0.04	0.08	-0.39	-0.06	0.11
Y_{10}	0.15	-0.28	0.24	0.16	0.34	-0.04	0.36	-0.20	-0.26	0.08
Y_{81c}	0.22	-0.36	0.32	0.13	0.44	0.01	0.39	-0.30	-0.31	0.06
ap	-0.45	-0.10	-0.36	-0.51	0.07	0.25	-0.15	0.12	0.26	-0.02
Dst	0.58	0.03	0.50	0.49	0.02	0.06	0.28	-0.13	-0.41	-0.07
$semi_{sin}$	0.22	0.03	0.19	-0.07	-0.04	0.05	-0.17	0.26	-0.04	-0.05
$semi_{cos}$	0.45	0.08	0.49	0.31	0.08	0.07	0.80	-0.01	-0.48	-0.31
ann_{sin}	-0.17	-0.19	-0.16	0.09	0.10	-0.10	-0.14	-0.09	0.19	-0.07
ann_{cos}	-0.37	0.95	-0.51	0.04	-0.95	-0.05	0.06	0.31	0.21	0.02

JB2008(2019)

	α_1	α_2	α_3	α_4	α_5	α_6	α_7	α_8	α_9	α_{10}
F_{10}	-0.37	0.23	-0.24	0.32	0.17	0.03	0.38	0.06	-0.09	0.31
F_{81c}	-0.43	0.37	-0.22	0.45	0.20	0.08	0.57	-0.09	-0.13	0.40
S_{10}	-0.33	-0.21	-0.16	-0.07	0.37	0.31	0.02	0.08	0.09	0.14
S_{81c}	-0.33	-0.34	-0.02	-0.17	0.55	0.52	0.02	-0.16	0.12	0.08
M_{10}	-0.16	-0.29	-0.02	-0.17	0.31	0.23	-0.17	0.06	0.05	-0.01
M_{81c}	-0.14	-0.60	0.14	-0.45	0.54	0.54	-0.36	-0.10	0.17	-0.16
Y_{10}	0.09	-0.29	0.22	-0.23	0.08	0.27	-0.58	0.07	0.07	-0.06
Y_{81c}	0.11	-0.38	0.27	-0.31	0.12	0.37	-0.72	0.09	0.11	-0.12
ap	-0.30	-0.07	-0.37	-0.06	-0.05	0.28	0.16	0.27	0.17	0.04
Dst	0.42	0.02	0.48	-0.00	0.06	-0.33	-0.22	-0.39	-0.19	-0.13
$semi_{sin}$	0.35	0.01	-0.10	-0.13	-0.39	-0.59	0.55	-0.17	0.03	-0.20
$semi_{cos}$	0.54	0.04	0.11	-0.08	-0.76	-0.13	-0.06	-0.10	0.30	0.26
ann_{sin}	-0.17	-0.20	-0.18	-0.13	0.27	0.10	0.70	-0.39	0.16	0.10
ann_{cos}	-0.29	0.98	-0.15	0.94	-0.31	-0.20	0.20	0.29	-0.36	0.57

Figure 9.

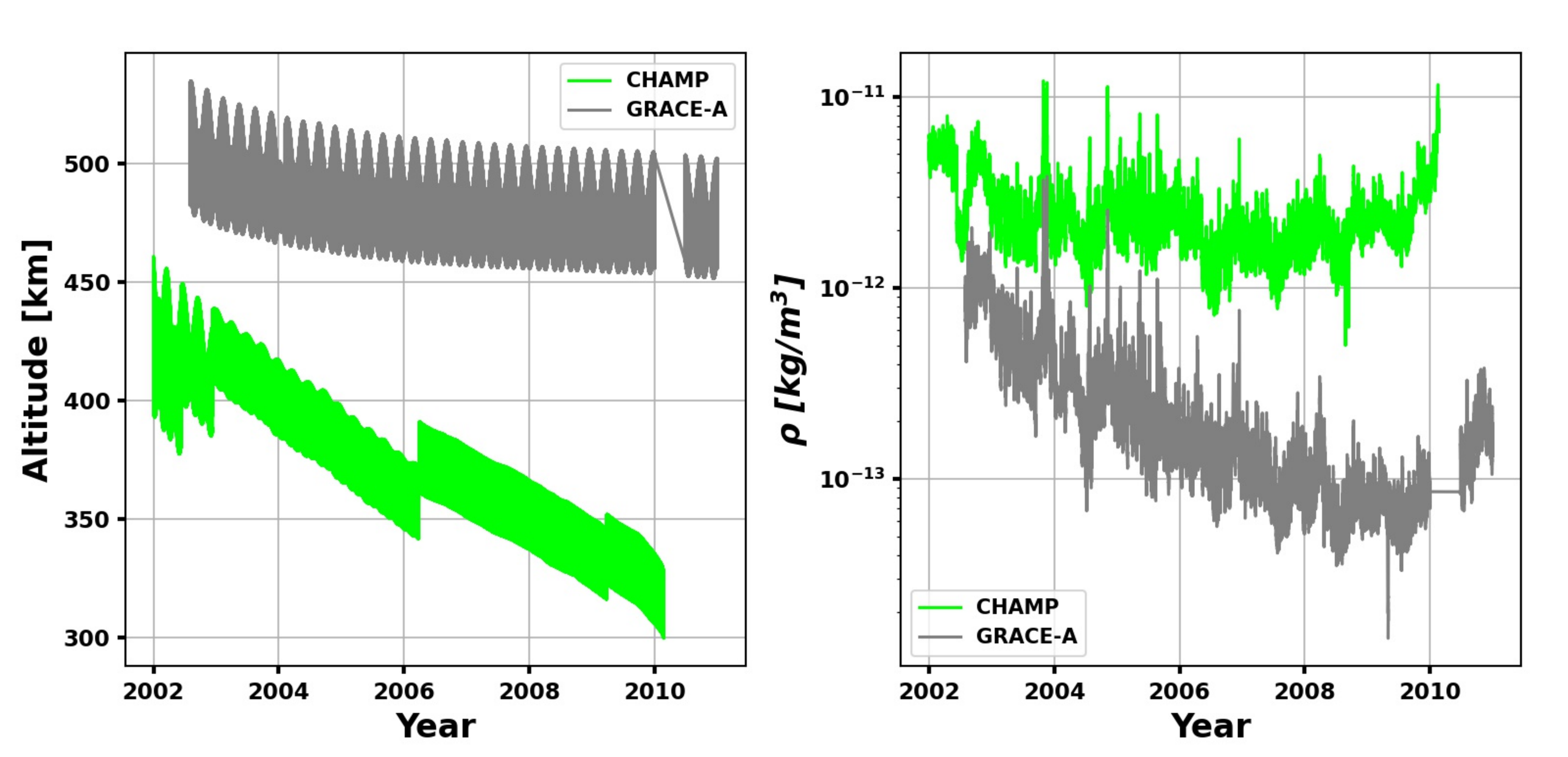

CHAMP
GRACE-A
Altitude [km]
500
450
400
350
300
2002
2004
2006
2008
2010
Year
ρ [kg/m³]
10⁻¹¹
10⁻¹²
10⁻¹³
CHAMP
GRACE-A
2002
2004
2006
2008
2010
Year

Figure 10.

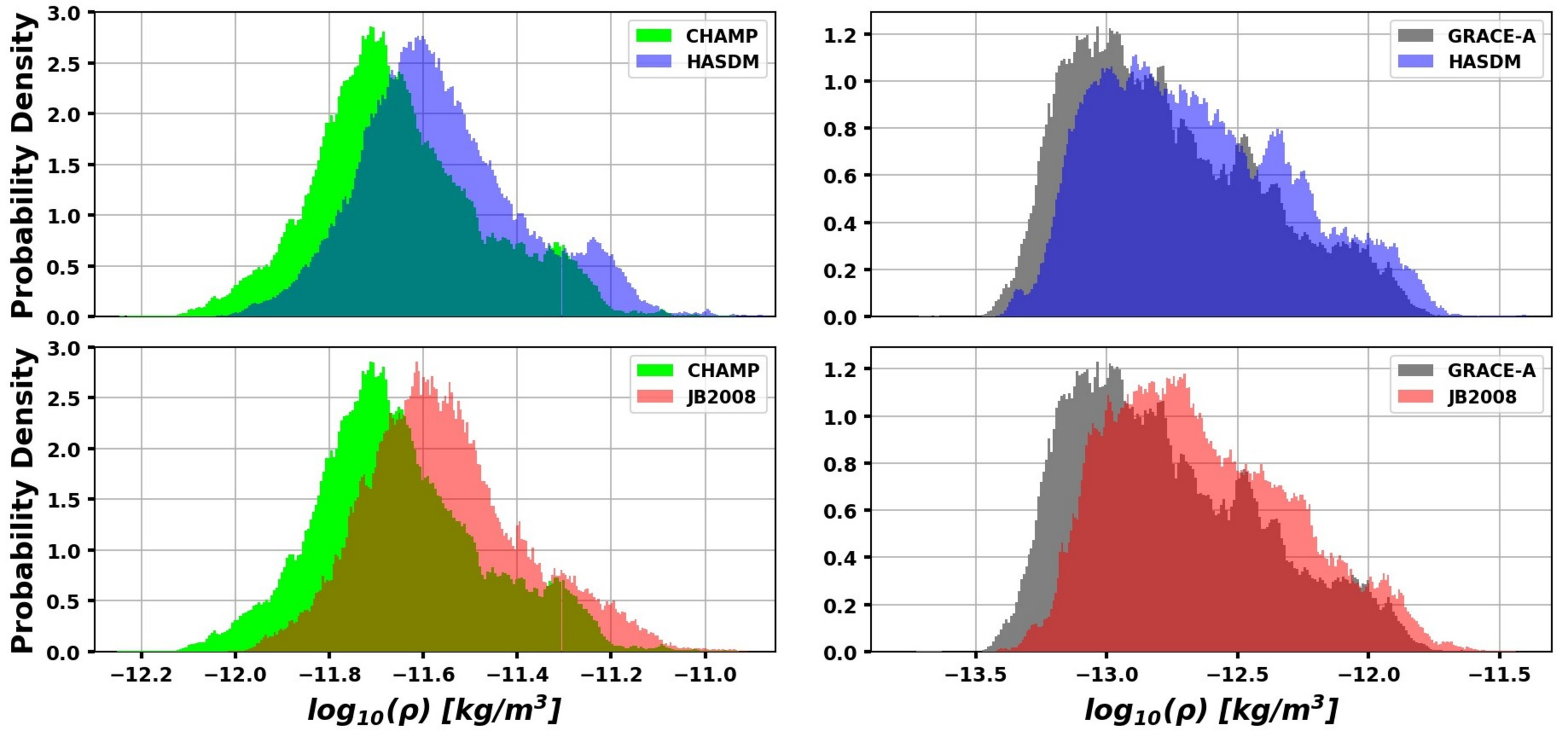

Probability Density
CHAMP
HASDM
GRACE-A
HASDM
Probability Density
CHAMP
JB2008
GRACE-A
JB2008
log10(ρ) [kg/m³]
log10(ρ) [kg/m³]

Figure 11.

CHAMP Orbit

1e−11

Density [kg/m^3]

2.0
1.5
1.0
0.5

CHAMP
HASDM
JB2008

GRACE Orbit

1e−12

Density [kg/m^3]

6
4
2

GRACE-A
HASDM
JB2008

CHAMP Orbit (Averaged)

1e−11

Density [kg/m^3]

1.25
1.00
0.75
0.50

CHAMP
HASDM
JB2008

GRACE-A Orbit (Averaged)

1e−12

Density [kg/m^3]

4
3
2
1

GRACE-A
HASDM
JB2008

F_{10} [sfu]

275
250
225
200
175

0 1 2 3 4 5 6

Days from October 28, 2003

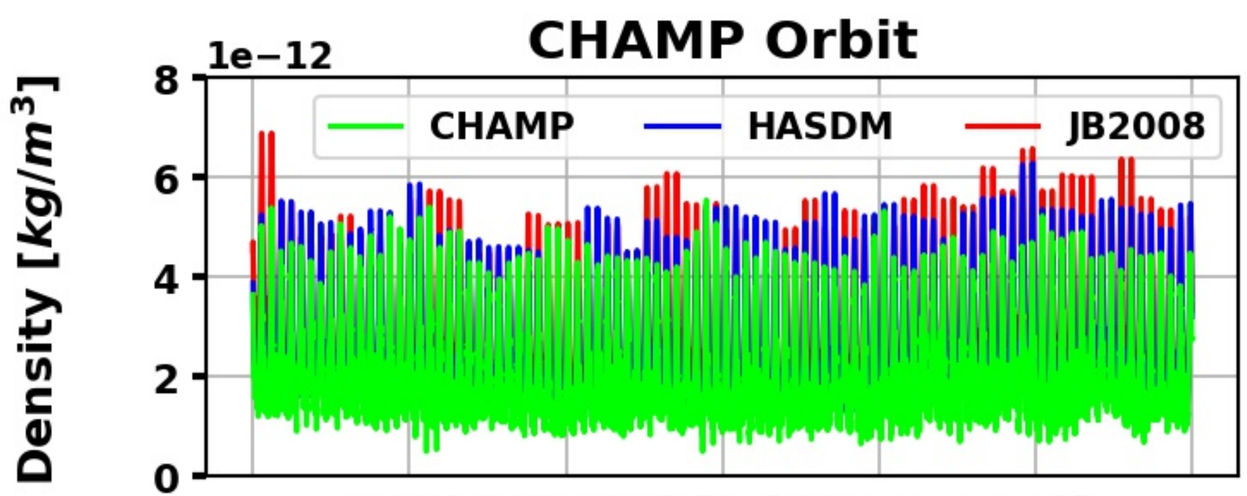
a_p [2nT]

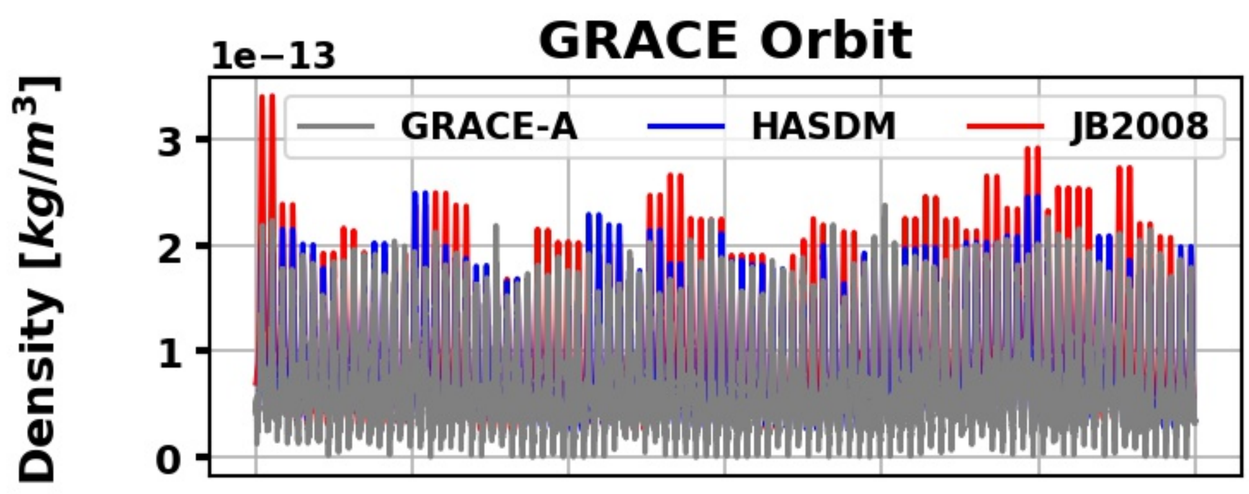
400
300
200
100
0

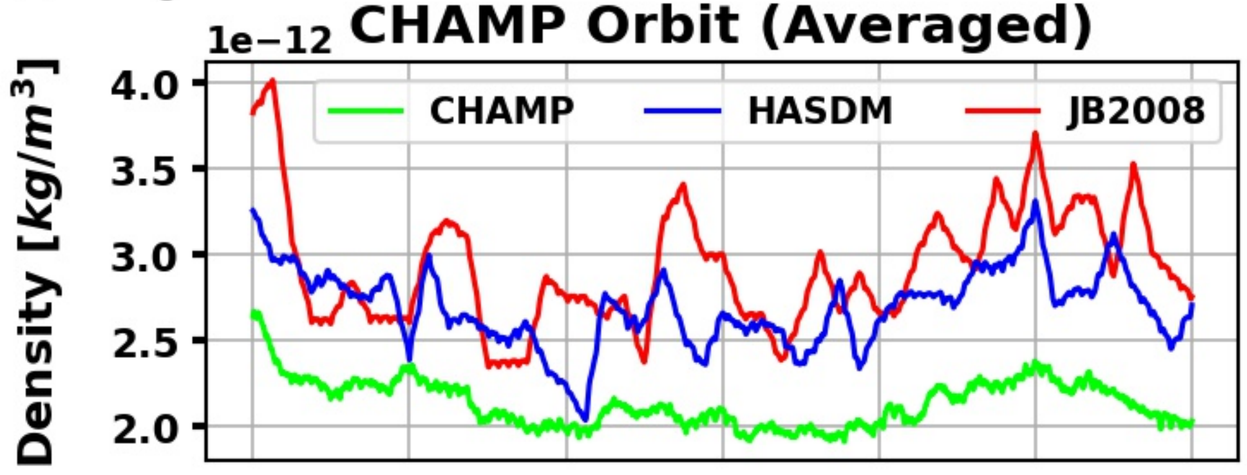
0 1 2 3 4 5 6

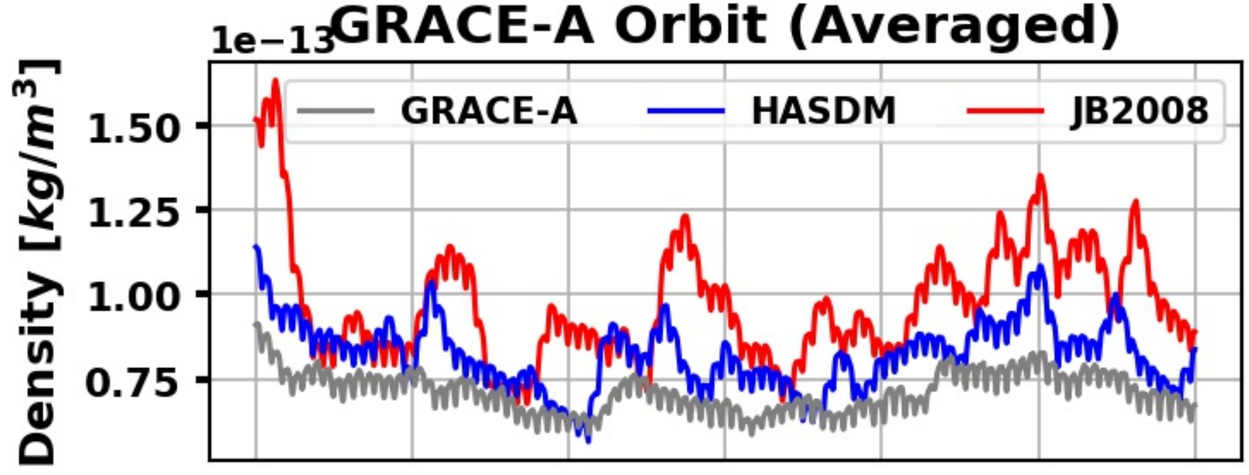
Days from October 28, 2003

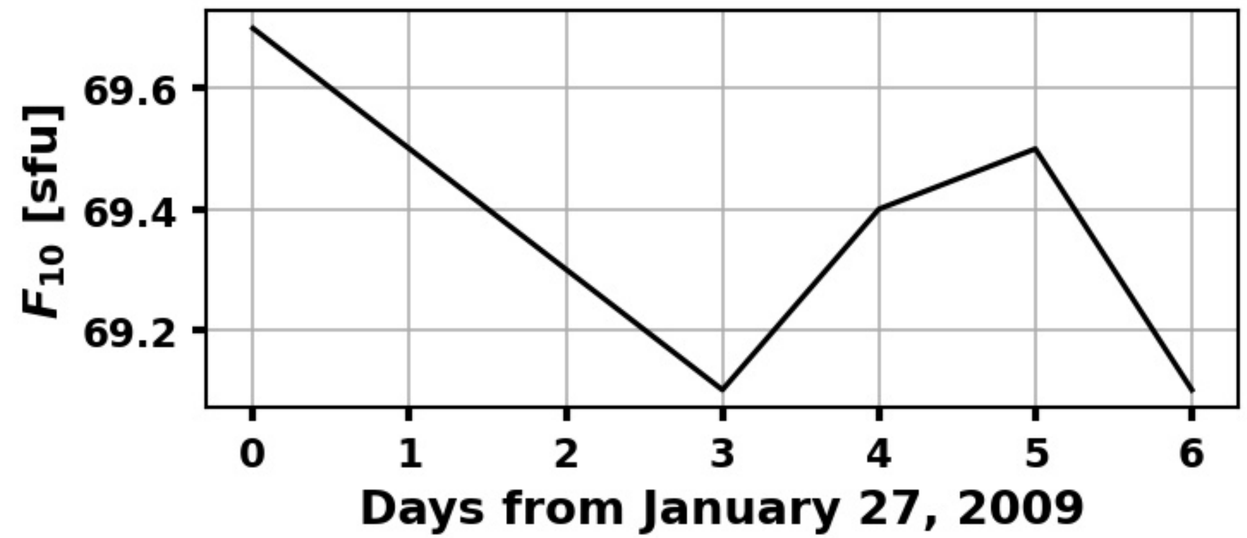
Figure 12.

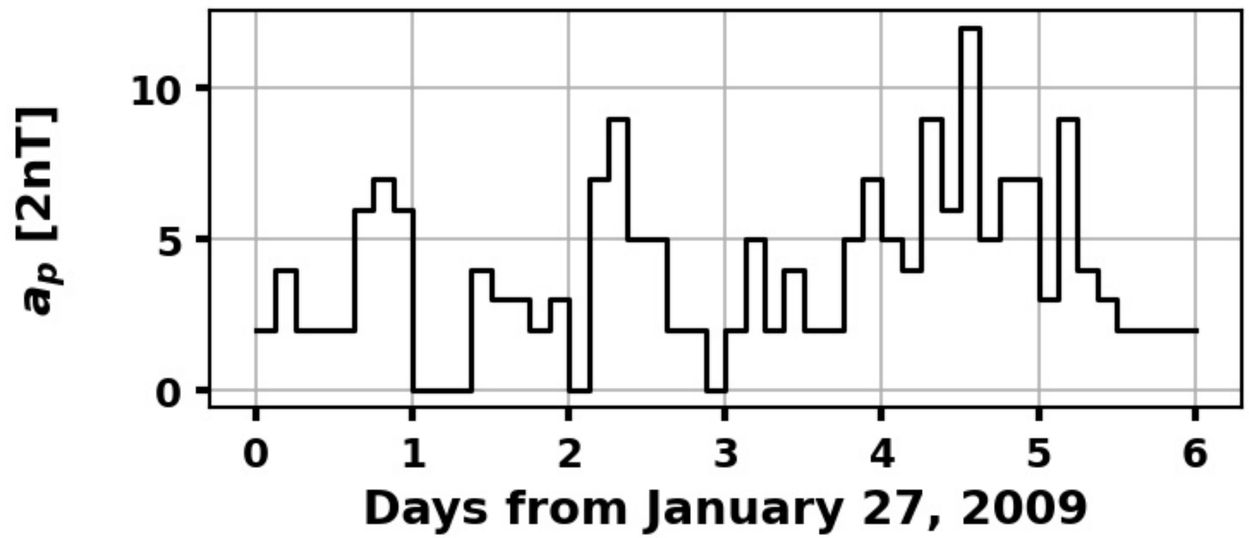

CHAMP Orbit
1e−12
CHAMP
HASDM
JB2008
Density [kg/m³]


GRACE Orbit
1e−13
GRACE-A
HASDM
JB2008
Density [kg/m³]


CHAMP Orbit (Averaged)
1e−12
CHAMP
HASDM
JB2008
Density [kg/m³]


GRACE-A Orbit (Averaged)
1e−13
GRACE-A
HASDM
JB2008
Density [kg/m³]


F10 [sfu]
Days from January 27, 2009


ap [2nT]
Days from January 27, 2009