On the Variation of Column $O/N\textsubscript{2} $ in the upper atmosphere using Principal Component Analysis in 2-dimensional images

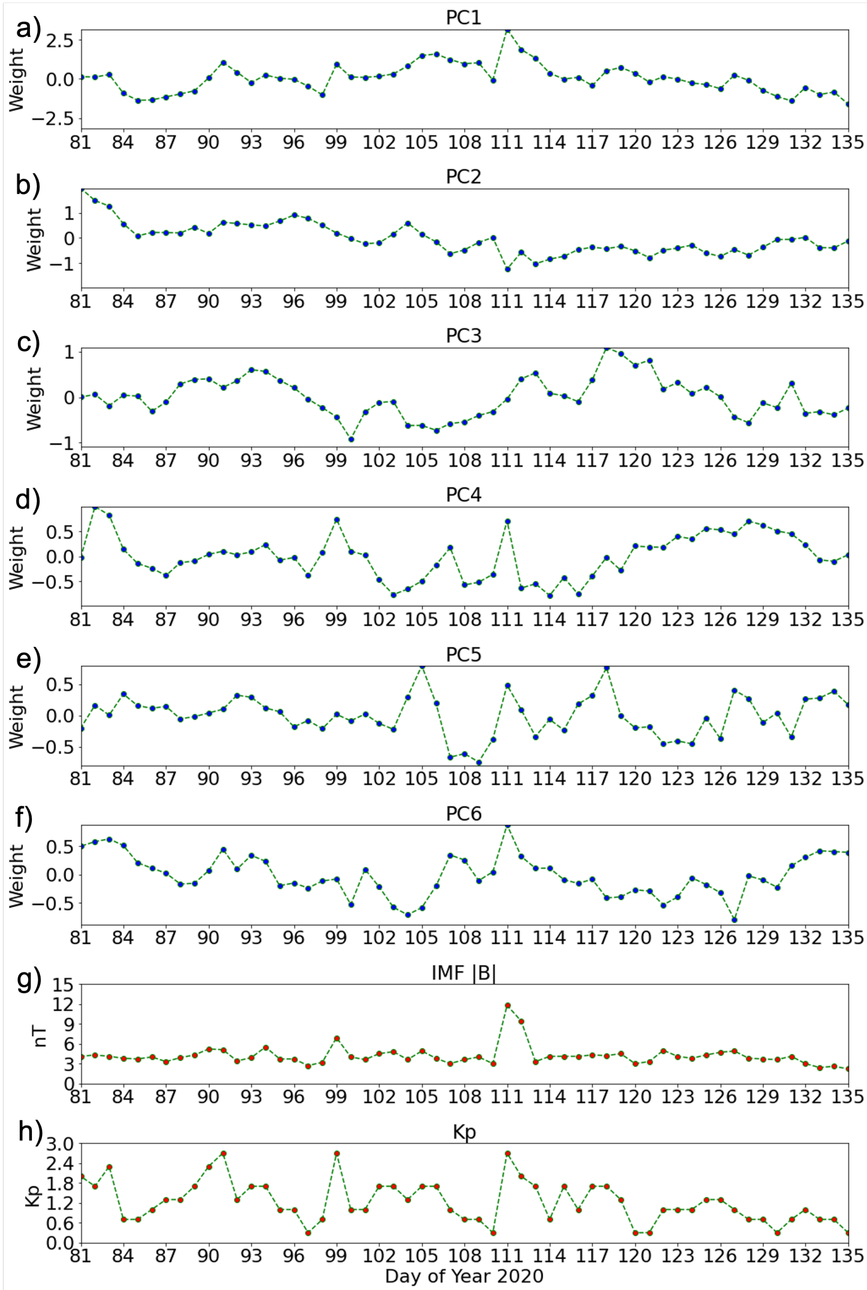
Goel Divyam[1], Wu Yen-Jung Joanne[2], Harding Brian J[2], Triplett Colin[2], Immel Thomas J.[2], Cullens Chihoko[3], and England Scott L[4]

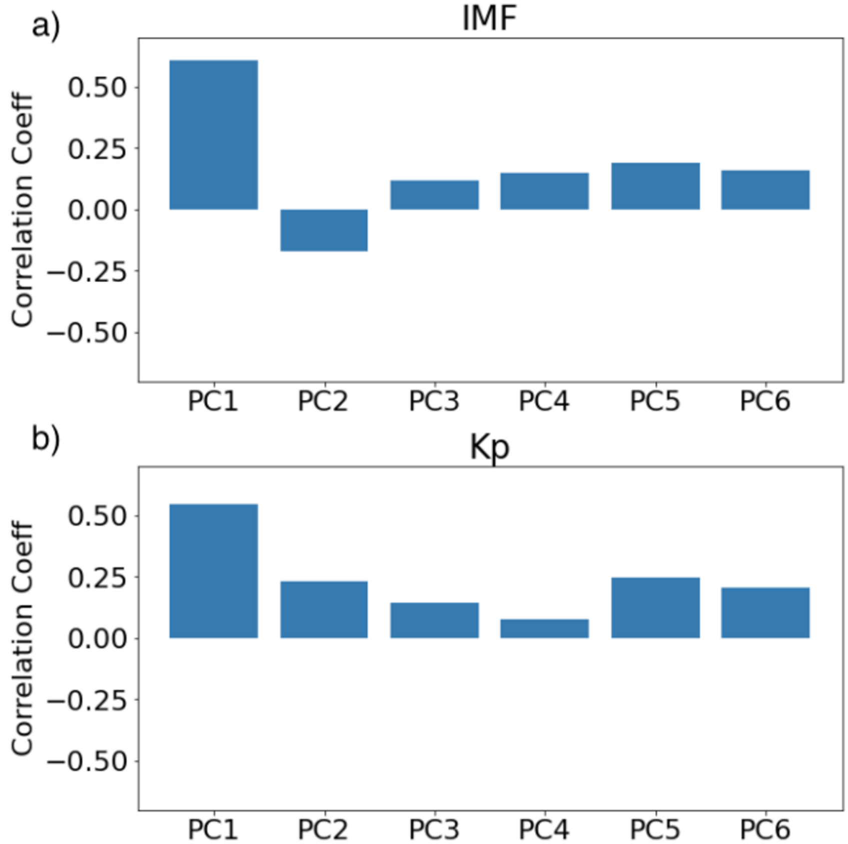
[1]Department of Electrical Engineering and Computer Science, University of California Berkeley
[2]University of California, Berkeley
[3]University of Colorado at Boulder / LASP
[4]Virginia Polytechnic Institute and State University

November 16, 2022

Abstract

Day-to-day variability in thermospheric composition is driven by solar, geomagnetic and meteorological drivers. The ratio of the column density of atomic oxygen and molecular nitrogen (O/N\textsubscript{2}) is a useful parameter for quantifying this variability that has been shown to exhibit close correspondence to F-region electron density, total electron content and upper atmospheric transport. Therefore, understanding the variability in O/N\textsubscript{2} gives an insight into the geophysical variability of other relevant ionospheric and thermospheric parameters. The relative contributions of these drivers for thermospheric variability is not well known. Here we report a new analysis of the variability in O/N\textsubscript{2} to identify the sources of variability in a 55-day time period. Principal Component Analysis (PCA) was performed on thermospheric O/N\textsubscript{2} column density ratio from days 81 to 135 of 2020 from NASA's Global-scale Observations of the Limb and Disk (GOLD) mission. We find that geomagnetic activity is the major source of variability in O/N\textsubscript{2} column density ratio, followed by solar-driven transport and meteorological driving from the lower atmosphere. The first component (PC1) showed a strong correlation to Kp index and IMF, and geomagnetic storm effects are seen in the wavelet analysis of PC1's weights. The fifth component (PC5) showed a strong quasi-6-day oscillation(Q6DO). The higher explained variance ratio of PC1 suggests a stronger effect of geomagnetic activity relative to meteorological forcing from planetary scale waves. The methodology of the present study also demonstrates how PCA can be used to isolate and rank different sources of variability in other IT parameters.


a)
PC1
b)
PC2
c)
PC3
d)
PC4
e)
PC5
f)
PC6
g)
IMF |B|
h)
Kp
Weight
nT
Kp
Day of Year 2020


a)
IMF
Correlation Coeff
0.50
0.25
0.00
−0.25
−0.50
PC1
PC2
PC3
PC4
PC5
PC6
b)
Kp
Correlation Coeff
0.50
0.25
0.00
−0.25
−0.50
PC1
PC2
PC3
PC4
PC5
PC6

On the Variation of Column O/N_2 in the upper atmosphere using Principal Component Analysis in 2-dimensional images

Divyam Goel[1], Yen-Jung J. Wu[2], Brian J. Harding[2], Colin C. Triplett[2], Thomas J. Immel[2], Chihoko Cullens[3], Scott England[4]

[1]Department of Electrical Engineering and Computer Science, University of California, Berkeley, CA
[2]Space Sciences Laboratory, University of California, Berkeley, CA
[3]University of Colorado, Boulder, CO
[4]Virginia Polytechnic Institute and State University, VA

Key Points:

- A principal component analysis for two dimensional images is applied to thermospheric column O/N_2 to characterize its variation.
- 64% of the variability in a 55-day period is captured in 6 components, whose structure and temporal behavior indicate driving processes.
- Clear indications of auroral forcing, seasonal trends, atmospheric tides and planetary waves are identified in components 1, 2, 3 and 5.

Corresponding author: Yen-Jung J. Wu, `yjwu@ssl.berkeley.edu`

Abstract

Day-to-day variability in thermospheric composition is driven by solar, geomagnetic and meteorological drivers. The ratio of the column density of atomic oxygen and molecular nitrogen (O/N_2) is a useful parameter for quantifying this variability that has been shown to exhibit close correspondence to F-region electron density, total electron content and upper atmospheric transport. Therefore, understanding the variability in O/N_2 gives an insight into the geophysical variability of other relevant ionospheric and thermospheric parameters. The relative contributions of these drivers for thermospheric variability is not well known. Here we report a new analysis of the variability in O/N_2 to identify the sources of variability in a 55-day time period. Principal Component Analysis (PCA) was performed on thermospheric O/N_2 column density ratio from days 81 to 135 of 2020 from NASA's Global-scale Observations of the Limb and Disk (GOLD) mission. We find that geomagnetic activity is the major source of variability in O/N_2 column density ratio, followed by solar-driven transport and meteorological driving from the lower atmosphere. The first component (PC1) showed a strong correlation to Kp index and IMF, and geomagnetic storm effects are seen in the wavelet analysis of PC1's weights. The fifth component (PC5) showed a strong quasi-6-day oscillation(Q6DO). The higher explained variance ratio of PC1 suggests a stronger effect of geomagnetic activity relative to meteorological forcing from planetary scale waves. The methodology of the present study also demonstrates how PCA can be used to isolate and rank different sources of variability in other IT parameters.

Plain Language Summary

Day-to-day variability in the ionosphere and thermosphere is driven by changes in solar radiation, the solar wind, and in meteorological forcing from the lower atmosphere. The thermospheric column O/N_2 responds to changes in thermospheric circulation and vertical transport, parameters which themselves are modified by the aforementioned drivers. Principal Component Analysis (PCA) is the algorithm that identifies the characteristic of human faces in the facial recognition technology, this study investigates the day-to-day variability of thermospheric column O/N_2 from NASA's GOLD mission. With the powerful open-source tool in hand, the characteristic spatial variations of column O/N_2 are identified. The present study reveals a strong response to a solar-wind driven geomagnetic storm and a smaller response to a quasi 6-day atmospheric wave. This provides a demonstration of the importance of geomagnetic effects relative to planetary scale waves, with presence in components 1 and 5 respectively.

1 Introduction

The day-to-day variability in the ionosphere and thermosphere (IT) system is primarily connected to variations in three different drivers: 1) solar radiation, 2) solar wind and magnetospheric inputs and 3) upward propagating atmospheric waves (Liu et al., 2021). Planetary waves (PWs), ultra fast Kelvin waves (UFKW) and a range of atmospheric tides are sources of variability that reflect changes in the lower atmosphere. The waves may interact in non-linear fashion and so influence the atmosphere-ionosphere coupling unpredictably (Forbes, 2021). Given the interest of the day-to-day variability, the attention has been on the oscillation of IT parameters by normal mode westward-propagating PWs with approximate periods near 2, 6, 10 and 16 days (Forbes et al., 2018). These waves are now commonly referred to as Quasi-2-day waves (Q2DW), Q6DW, Q10DW and Q16DW,respectively.

Numerous investigations have been conducted to understand the nature of PWs in the upper atmosphere. Yue et al. (2016) summarize possible mechanisms of PW-tide interactions and their impacts on IT system in the context of Q2DW interactions. The dissipation of the westward propagating PWs drives the change of the thermospheric wind circulation in the lower thermosphere, that results in decreases in the mixing ratio of atomic O and and increase in those of N_2 and O_2 in both the lower and upper thermosphere (Yue & Wang, 2014; Yue et al., 2016, and references therein). Chang et al. (2014) first reported observations of a decrease in GUVI O/N_2 ratio in response to six distinct Q2DW events. Gan et al. (2015) showed the effects of the dissipative 6.5 wave on the IT system via the mixing mechanism predicted by Yue and Wang (2014).

The measurement of the column O/N_2 density ratio has provided a key capability for detecting varying conditions in the upper atmosphere (Cai et al., 2020; Liou, 2005; Lu et al., 2012; Zhang, 2003; Rishbeth, 1998; Oberheide et al., 2020). NASA's Global-scale Observations of the Limb and Disk (GOLD) (Eastes et al., 2020) mission has provided a ground-breaking new source of column O/N_2 density ratio, after it became operational in October 2018. Several studies using GOLD data have shown the response of the neutral atmosphere and ionosphere to changes in solar wind and solar extreme ultraviolet radiance. Regarding solar wind disturbances, Gan et al. (2020) reported a geomagnetic storm effect on the OI 135.6 nm dayglow, which showed a deep depletion in brightness as well as a striking westward displacement of the intensified dayglow. Cai et al. (2020) used GOLD's O/N_2 observations and simulations from the Thermosphere-Ionosphere Electrodynamics General Circulation Model (TIE-GCM) to show that weak geomagnetic activity during solar minimum conditions created weak geomagnetic storm-like variations in O/N_2. The variations in O/N_2 persisted for more than 10 hours even after the end of weak geomagnetic disturbances, suggesting a pervasive and longer-term influence of geomagnetic activity on the day-to-day variation in the O/N_2. Regarding solar radiance variations, Schmölter et al. (2021) showed a weak correlation of GOLD's O/N_2 data product with the F10.7 index and GOLD's proxy EUV flux over two well-defined 27-day solar rotation periods. A 3% decrease in O/N_2 mean was attributed to the overall decrease in solar activity during one of the periods (Schmölter et al., 2021). However, it was also suggested that the decrease could be attributed to temperature or wind changes in the upper atmosphere. Regardless, the influence of the solar-rotation cycle on the monthly variation of the O/N_2 density ratio can be considered weak in comparison to solar wind forcing.

Principal Component Analysis (PCA), also known as the empirical orthogonal function analysis, has been used extensively to study variability in IT parameters (see detailed introduction on PCA in section 2). In the context of machine learning regime, PCA has a wide application in pattern recognition, feature selection and dimensional reduction. While applying PCA on one-dimensional time series reveals the dominant variations as a function of time (e.g. Chen et al. (2007)), performing PCA on two-dimensional (2D) spatial data is able to extract the characteristic spatial patterns. For example, Flynn et al. (2018) identifies the strongest variation mode of the global thermospheric nitric oxide infrared radiative flux; Alken et al. (2017) demonstrates that using the eigenmodes of the global ionospheric current extracted from TIEGCM. The features of the equatorial electrojet and Sq current systems can be accurately reproduced; The dominant modes of variability in large-scale Field-Aligned Current is also identified by PCA, and can be used in specifying error covariance for a data assimilation procedure (Cousins, Matsuo, Richmond, & Anderson, 2015; Cousins, Matsuo, & Richmond, 2015).

In this study, we apply PCA to isolate and rank different sources of variability in other IT parameters. The global-scale day-to-day variability in the thermospheric O/N_2 is reported based on the column O/N_2 from GOLD in day 81 to 135 of 2020. The detailed introduction of the dataset and the methodology are given in Section 2. The primary principal components of spatial variation patterns in column O/N_2, as well as the signature of the geomagnetic storm and oscillations in response to Q6DW are presented and discussed in Section 3. Further discussion and conclusions are given in Section 4 and Section 5, respectively.

2 Data and Methodology

The GOLD mission was launched onboard the SES-14 satellite to a geostationary orbit positioned above the Amazon Delta Basin at 47.5° longitude. It takes global-scale observations of Earth's disk in the far ultraviolet spectrum, providing radiances from 135-180 nm in every imaging pixel. GOLD's far ultraviolet imager has two identical channels, each capable of scanning Earth's entire disk every 30 minutes. Two scans collect spectro-photometric data in alternating northern and southern hemispheres, each taking 12 minutes. In total, GOLD obtains approximately 68 scans of the Earth each day, spanning 06:00 to 12:00 UT. This study uses data from days 81 to day 135 of 2020. This period sits inside a grating yaw mechanism actuation cycle. During the same cycle, all measurements have been obtained from the same part of the detector so that the sensitivity of the instrument is stable. The detail of the instrumental calibration and the update of the data products can be found in `https://gold.cs.ucf.edu/wp-content/documentation/GOLD_Public_Science_Data_Products_Guide_Rev4.4.pdf`. During the chosen period, three minor-to-moderate geomagnetic events (daily mean Kp > 2.5) occurred on day 94, day 99 and day 111. The event on day 111 is the strongest among the three, for which the maximum Kp and minimum Dst reached 47 and -59 nT, respectively. The column O/N_2 density ratio is derived from OI 135.6 nm and N_2 LBH band emission measurements, for more than 6 hours of every day when portions of the the visible disk are sunlit. A detailed description of the instrument and the Level 2 data product algorithm can be found in Eastes et al. (2020). In this study, we collect data between 14:30 UT and 15:30 UT (scans 34 to 39) for each day. The scans from 34 to 39 cover the entire Disk from GOLD's field of view (FOV). GOLD's FOV covers the entire disk visible from its location at 47.5° longitude, between latitudes ±70° and longitudes -115° to 20°. This study uses version 3 of the level 2 product.

The PCA in this work is performed by the PCA function in the Python package scikit-learn 1.0.2 (Pedregosa et al., 2011). PCA is a dimensionality reduction technique used to discover the directions of greatest variance in some given data. It is an optimization problem trying to find an axis through a cloud of data in n-dimensional space that maximizes the variance of the projected data points onto that axis. Location coordinates are arranged as the basis vectors and each day is considered a data point for the data matrix. Based on this arrangement, the weights of each data point for a PC are computed. The weights of each data point for a PC represent how strongly aligned a data point is to the axis of variance of that PC. The weight of a PC for a data point is calculated by taking the inner product between the mean-centered data point eigenvectors of the covariance matrix associated with that particular PC. A large absolute weight for a particular PC suggests that most of the data in that data point are aligned along the axis of that PC. More information on PCA can be found in Preisendorfer and Mobley (1988).

3 Result

The first six principal components (PCs) are shown in Figure 1a-f. For simplicity, we denote PCn as the n-th principal component in the text hereafter. Underneath each plot is the explained variance ratio of the PCs. The black lines indicate ± 30 degrees magnetic latitude, giving a rough poleward boundary of the Equatorial Ionization Anomaly (EIA) in the quiet time (Stolle et al., 2008; Balan et al., 2018). The PC1 and PC2 show primarily latitudinal variation, while PC3 shows a strong longitudinal variations. The PC4 to PC6 displays a mix of longitudinal and latitudinal variation. The first six PCs together represent around 64% of the variance in the data during this 55-day window. The explained variance ratio of PC1 is 0.30, which means 30% of the total variation is explained by the latitudinal pattern shown in PC1. The plots of the principal components represent the axis or modes of variance, therefore, the color of one region alone holds no physical significance. Two oppositely colored regions in the graph indicate that those

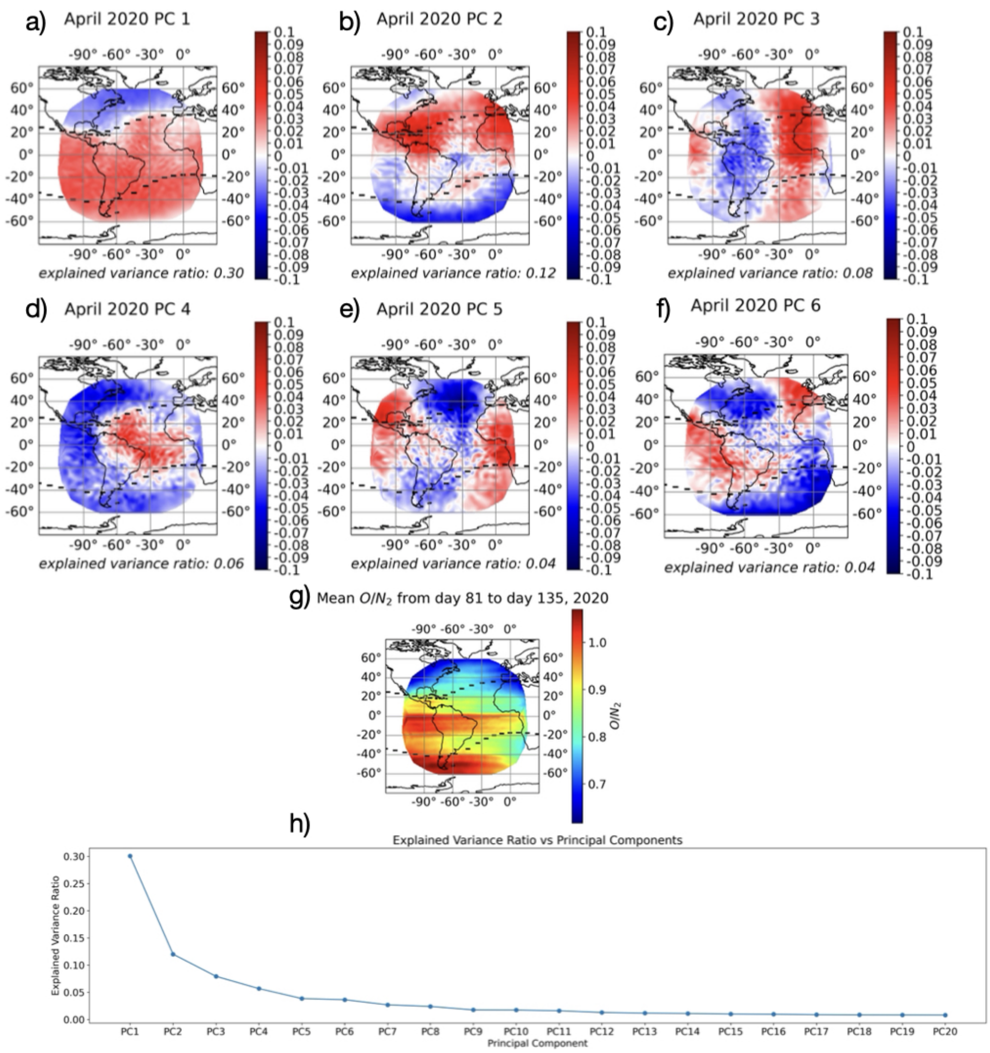


Figure 1. Figure 1(a)-(f) show the principal components 1 to 6, which contribute around 64% of the variation in the target period. The black dashed lines are ± 30 degrees magnetic longitude. Figure 1(g) shows Mean O/N_2 between days 81 and 135. Figure 1(h) shows the explained variance ratio of the first 20 PCs.

regions are oppositely aligned along the axis of variation of the PC and the intensity represents the strength of that polarity.

The mean O/N_2 at each location pixel for the entire time frame is shown in Figure 1g. The features in the Figure 1g represent permanent features within the data, some of which are likely to be systematic artifacts; e.g. the extended stripe at the equator. Since PCA involves mean-centering the data, permanent features do not contribute to the PC calculations and the analysis of the variability. Figure 1h shows how the explained variance ratio decreases with higher PCs. The higher PCs are most likely fitting the variance in noise rather than the variance in actual data.

Of the latitudinally varying PCs, the structure of PC1 reflects well a known feature of geomagnetic storms, the reduction of O/N_2 in the middle-latitude morning sector (Strickland et al., 1999; Immel et al., 2000). That the boundary of the zero-value of

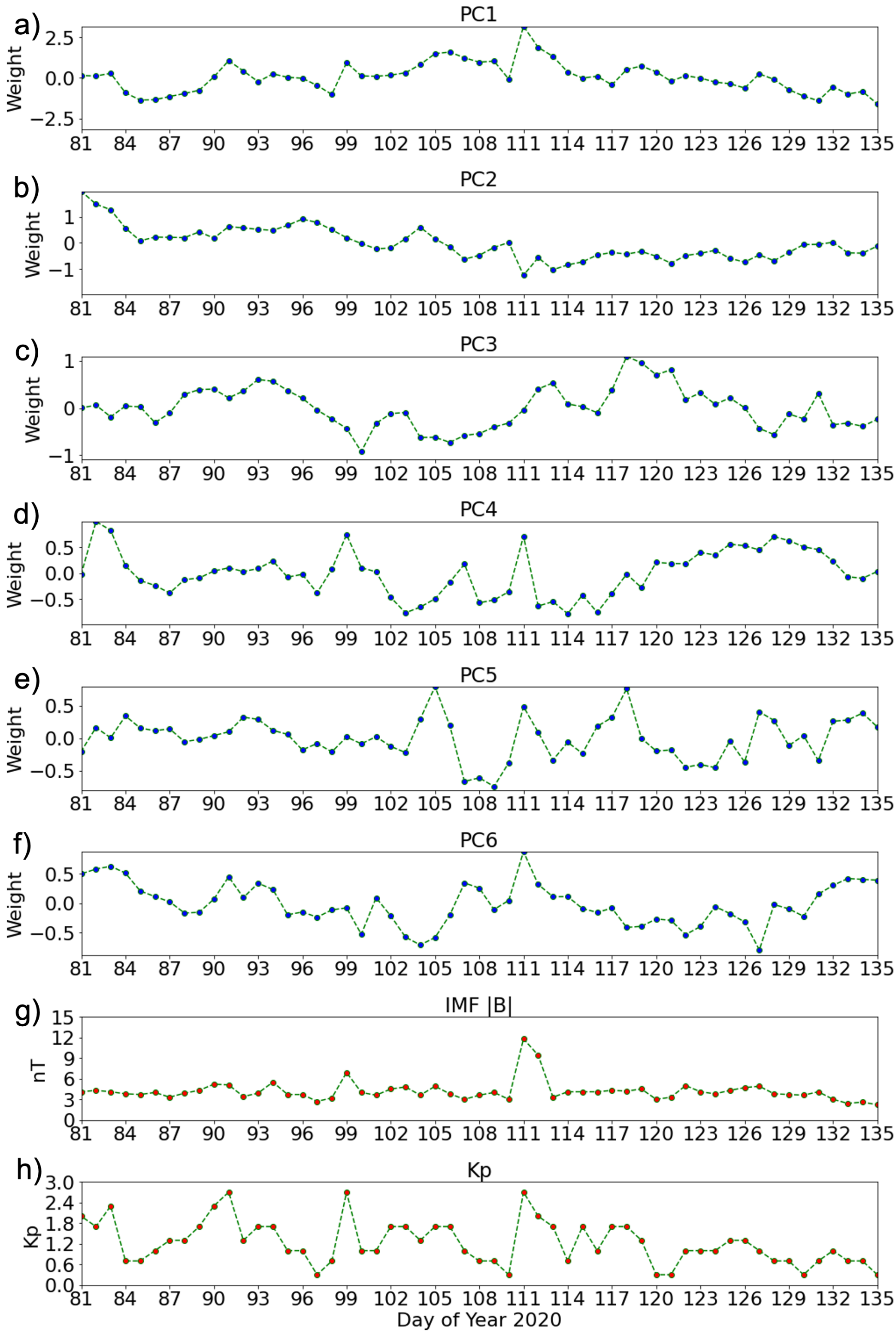


Figure 2. (a)-(f)Time series of daily weights of the first six principal components, (g)IMF amplitude and (h)Kp index from day 81 to day 135 of 2020.

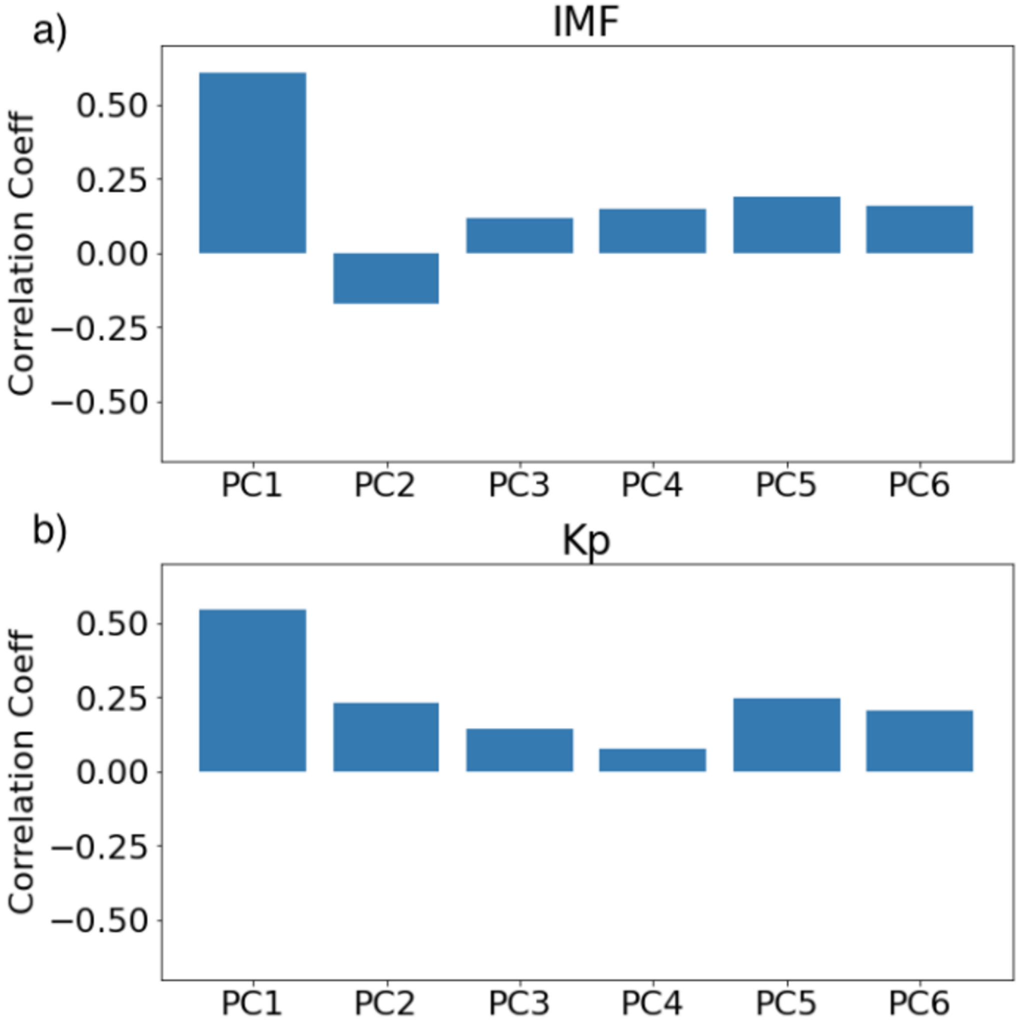


Figure 3. Correlation Coefficient of (a)IMF strength and (b) Kp index over time with the variability of daily weights of each principal component.

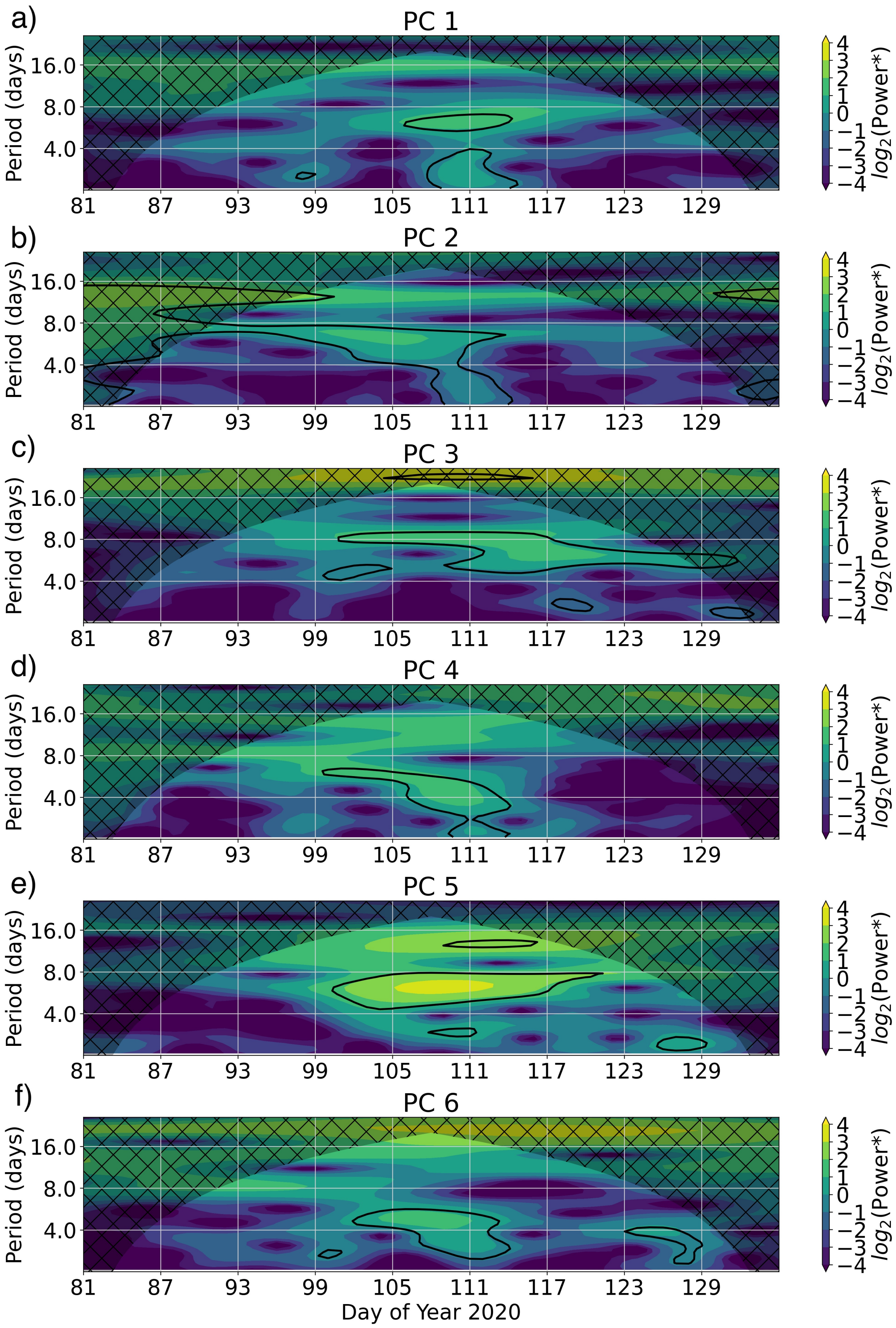


Figure 4. (a)-(f) The wavelet analysis of daily weight time series of the first six principal components. The color bar on the right represents the strength of a particular frequency in the time series. Regions where the confidence level exceed 99% are outlined in black. Cross-hatched regions indicate the cone of influence, where edge effect becomes non-negligible.

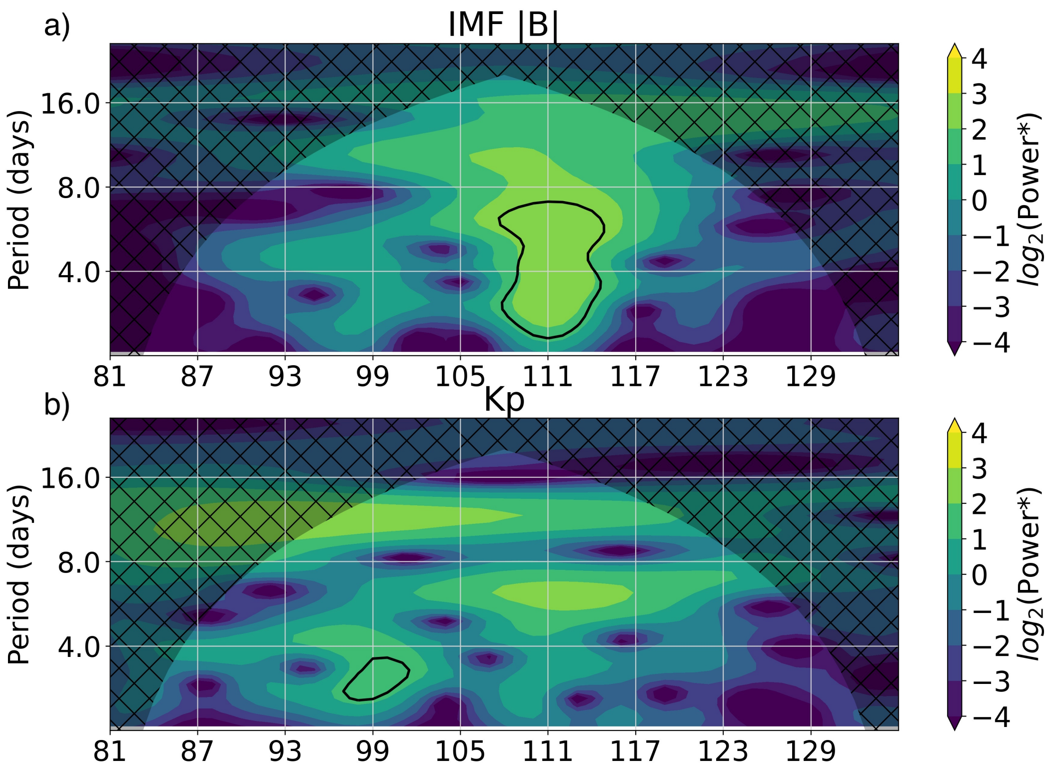


Figure 5. Same as Figure 4 but for (a) IMF strength and (b) Kp Index.

this PC follows a line of magnetic latitude indicates the potentially auroral origin of this feature. The difference between PC1 and PC2 is that PC2 exhibits this zero-value boundary at the equator. Given the natural transition of seasonal highs in O_2 from one hemisphere to the other during this time period that contains equinox (Rishbeth & Müller-Wodarg, 1999), it is possible that this transition is captured by this PC. A time series analysis of these PCs supports these conclusions, as shown later in this report.

The PC with the most prominent longitudinal pattern/polarity is PC3. Several geophysical phenomena may vary with either longitude or local time, and therefore present variability that would be fitted in this PC. Diurnal solar insolation clearly increases with longitude in these 15 UT snapshots, with accompanying temperature changes and possible vertical transport. Atmospheric tides are one significant feature of the thermosphere which exhibit strong zonal organization, accompanied by transport effects that may affect O/N_2. The structure in PC3 has a zonal wavelength about 120° taking at the same UT everyday, which resembles the migrating tide with zonal wavenumber $\bar{3}$, the terdiurnal migrating tide (TW3). However, unlike PC3, a climatology study of TW3 in temperature shows a clear latitudinal structure with minimums of the amplitude at 20N and 20S latitude (Pancheva et al., 2013). This discrepancy between PC3 and TW3 climatology indicates that other non-migrating tides may play the role.

The later PCs (PC4-6) have approximately 5% explained ratio each , but their spatial structures are worth attention as well. PC4 reveals a concentric-like structure at the center of the FOV. A possible explanation to the structure in PC 4 is through the g-factors that is used to parameterize the excitation and ionization rate in the O/N_2 retrieval(Strickland et al., 1997). The g-factors are related to solar zenith angle and column density that both are functions of the radial distance from the center of FOV. PC5 displays a zonal wave structure that is roughly symmetric to the equator, which we identified as a signature of Q6DW. The detailed discussion on PC5 is given in the following paragraphs. A complicated structure is shown in PC6, that is partially aligned with the 30° magnetic latitude and its latitudinal and longitudinal structures are hard to define. The complicity implies multiple factors that can create local variation start taking place to influence the PC spatial pattern.

To reveal the temporal variation of each PC spatial patterns, the weight of each PC is calculated by taking the inner product of the target PC and the daily O/N_2 data. Time series of daily weights with respect to each PC is shown in Figure 2a-f. The daily-averaged Interplanetary Magnetic Field (IMF) and Kp index (Figure 2g-h) from NASA's OMNI dataset (King & Papitashvili, 2005; Matzka et al., 2021) are brought in to compare with the time series of each PC, to see whether the solar wind and geomagnetic condition contribute to the variations. The absolute IMF magnitude (Figure 2g) presents the geomagnetic disturbances measured outside of the magnetosphere, whereas the Kp index reflects the geomagnetic perturbation at the mid-high latitudes on the surface of the earth. In the IMF data, Figure 2g, it shows geomagnetic storms on day 111, while the geomagnetic Kp index, Figure 2h, shows the significant geomagnetic perturbations occurred on day 91, 99 and 111.

In Figure 2, PC1, PC4, PC5 and PC6 show a spike that coincides with the geomagnetic storm occurred on day 111(Figure 2g-h). This correspondence necessitated a study of the relationship of the PC weights with the IMF and Kp index. In Figure 3, we report the correlation between the weights of each PC and the geomagnetic indices. The IMF and Kp index show the greatest correlation with PC1, having a correlation value of 0.6 and 0.55, respectively. The magnitude of the correlation value with the weights of the other PCs is less than 0.25 for both IMF and Kp index. The outstanding relation between PC1 and the geomagnetic indices suggests that geomagnetic disturbance is likely the driver of the spatial pattern in PC1. The contrast of colors in the high latitude and mid-low latitude in PC1 (Figure 1a) indicates the variations are negatively correlated. Such pattern is consistent with the storm time effect to O/N_2, where O/N_2 is decreased due to the thermal expansion at the high latitudes, and is increased due to the circulation and thermal changes on composition in the mid-low latitudes (Burns et al., 1995; Prölss et al., 1991).

The Wavelet Analysis (Torrence & Compo, 1998) is applied on the time series of weight of each PC (Figure 4) and geomagnetic indices (Figure 5) to gain further insights into their respective variation as a function of period and time. The black contour on Figure 4 and Figure 5 indicates the area that exceeds the 99% confidence level with respect to the red noise with lag-1 autocorrelation (Torrence & Compo, 1998). Due to the limit of sampling rate, 1 sample/day, and the size entire examining window, 55 days, the shortest period and the longest period the wavelet analysis can resolve are 2 days and 16 days, respectively. The features mentioned above make analysing the time series of the weight a tool to study the day-to-day variability in the column O/N_2.

The most remarkable feature in Figure 4 is the relative large power that reaches 99% confidence level between the period of 4 days to 8 days in PC5 from day 100 to day 120. The maximum power is identified in the period of 6.2 days on day 110. Since the range of period is near to the Q6DW, the planetary wave with a period about 6.5 days, the oscillation identified in PC5 is likely be driven by the Q6DW. During the same time period, Wu et al. (2022) reported a significant Q6DW event in the neutral wind and temperature data from the Ionospheric Connection Explorer (ICON) data, which reinforces the suggestion of identifying Q6DW by PC5. The spatial pattern in PC5 (Figure 1e) is worth noting, in which the amplitude of the longitudinal structure is larger in the mid-high latitude region, and relatively weaker in the equatorial region. This latitudinal preference is consistent with the symmetric structure of the Hough function solution of the gravest symmetric wavenumber 1 Rossby (1,1) mode for vertical wind and temperature

(Talaat et al., 2001; Forbes et al., 2020). However, from PC5 pattern alone, we can not justify whether the Q6DO is from Q6DW itself or the modulation of Q6DW on the atmospheric tides.

In addition to the effect of the planetary wave, the geomagnetic disturbances is another significant feature in the wavelet spectrum. The spectrum of PC1's weights shows a strong amplitude centered around day 111 (Figure 4a), while IMF amplitude also shows higher power around day 111 in the period range less then 8 days. The agreement of the period of the oscillation in time domain in PC1 and IMF amplitude can explain the significant correlation of the pair. On the other hand, the high correlation between Kp Index and PC1 is predominantly from the enhancement of the period between 4 days and 8 days, from day 100 to 120.

4 Discussion

While Kp index has an enhancement near the period of Q6DO from day 105 to 123 indicating a series of recurrent magnetic disturbances, IMF amplitude shows a boardband amplitude enhancement around the storm day on day 111 due to the exceptional peak. The geomagnetic disturbance of day 94 is not visible in the wavelet analysis. This is probably because the event was most active at 2100 UT, well outside the 14:30-15:30 UT time frame that data were accumulated. Since the IMF amplitude is a space-born measurement outside of the magnetosphere and Kp index is derived from the ground-based magnetometers, the quasi-6 day oscillations in Kp index could be the influence of the strong Q6DW from the lower atmosphere under a relatively quiet solar condition, given the strong Q6DW observed in the neutral atmosphere by ICON (Wu et al., 2021).

Fang et al. (2018) (see also Liu et al. (2021)) demonstrated that geomagnetic activity was the main contributor to ionospheric variability during the period they investigated, followed by the perturbation from the lower atmosphere. The PC ranking present in the current work also reflects the competition between geomagnetic disturbances and influences coming from below. During this 55-day window that accompany with minor geomagnetic storms, the contribution to the thermospheric O/N_2 variability are in the order as following: geomagnetic disturbances (PC1), seasonal interhemispheric transition (PC2), atmospheric tides(PC3) and planetary waves (PC5). One would expect a geomagnetic-disturbances-related pattern may not be PC1 if looking at a geomagnetic quiet period, but whether the contribution ranking of the three meteorological forcing changes is still unknown, and can be an extension work based on the current study.

Liu et al. (2021) raised interest in the spatial and temporal evolution of IT parameters in a global context. The present work showed the eigenmodes of global-scale variability in thermospheric O/N_2. Each PC shown is likely representative of the net effects of the superimposition of more than one geophysical process. Therefore, it is challenging to thoroughly infer the geophysical process behind a PC through visual inspection. In addition to the time series analysis presented in this work, a potential future direction would be to perform PCA on numerical simulations by isolating geophysical processes and looking at the geospatial eigenmodes to infer the spatial influence of different drivers. The PCA patterns that are extracted from the model with isolating effects can be used as a tool to quantify the relative effects of different geophysical drivers.

5 Conclusion

Principal Component Analysis was performed on GOLD's Level 2 column O/N_2 data for the period between day 81 and 135 of 2020. The eigenvectors, principal components, from PCA reveals the characteristic spatial patterns. We reported the first 6 PCs, which together explained 68% of the column O/N_2 variation during the examined period. The time series analysis of the daily weight is conducted to study the temporal variation of each PC. Correlation of the time series between the daily weights and IMF and Kp index was performed. A summary of the main findings is as follows:

1) The structure of PC1 reflects well a known feature of geomagnetic storms, the reduction of O/N_2 in the middle-latitude morning sector. The time series analysis shows that Kp index and IMF strength correlate most to PC1 weights. This suggests that geomagnetic activity is the primary driver of global-scale thermospheric O/N_2 variability.

2) The wavelet analysis reveals that the spatial pattern of PC5 is a quasi 6 day oscillation in thermospheric O/N_2 driven by planetary waves as one of the major modes of variability.

3) The order of PCs discloses the competition between geomagnetic disturbances and influences coming from below. The minor geomagnetic storms take the lead following by seasonal interhemispheric transition, atmospheric tides, and planetary waves.

Acknowledgments
The GOLD data is available at `http://gold.cs.ucf.edu/`. The OMNI data were obtained from the GSFC/SPDF OMNIWeb interface at `https://omniweb.gsfc.nasa.gov`. The Python code of this work is avaliable at `https://github.com/divyam123-EECS-Physics/GOLD-PCA.git`

References

Alken, P., Maute, A., Richmond, A. D., & Vanhamäki, G. D., H.and Egbert. (2017, May). An application of principal component analysis to the interpretation of ionospheric current systems: TIEGCM MODELING, PCA, AND DATA FITTING. *J. Geophys. Res. Space Physics*, *122*(5), 5687–5708. Retrieved 2022-07-06, from http://doi.wiley.com/10.1002/2017JA024051 doi: 10.1002/2017JA024051

Balan, N., Liu, L., & Le, H. (2018). A brief review of equatorial ionization anomaly and ionospheric irregularities. *Earth and Planetary Physics*, *2*(4), 257-275. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.26464/epp2018025 doi: https://doi.org/10.26464/epp2018025

Burns, A. G., Killeen, T. L., Deng, W., Carignan, G. R., & Roble, R. G. (1995). Geomagnetic storm effects in the low- to middle-latitude upper thermosphere. *Journal of Geophysical Research: Space Physics*, *100*(A8), 14673-14691. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/94JA03232 doi: https://doi.org/10.1029/94JA03232

Cai, X., Burns, A. G., Wang, W., Qian, L., Solomon, S. C., Eastes, R. W., ... McClintock, W. E. (2020, September). The Two-Dimensional Evolution of Thermospheric O/N $_2$ Response to Weak Geomagnetic Activity During Solar-Minimum Observed by GOLD. *Geophys. Res. Lett.*, *47*(18). Retrieved 2022-07-05, from https://onlinelibrary.wiley.com/doi/10.1029/2020GL088838 doi: 10.1029/2020GL088838

Chang, L. C., Yue, J., Wang, W., Wu, Q., & Meier, R. R. (2014, June). Quasi two day wave-related variability in the background dynamics and composition of the mesosphere/thermosphere and the ionosphere. *J. Geophys. Res. Space Physics*, *119*(6), 4786–4804. Retrieved 2022-07-05, from http://doi.wiley.com/10.1002/2014JA019936 doi: 10.1002/2014JA019936

Chen, G.-X., Xu, W.-Y., Du, A.-M., Wu, Y.-Y., Chen, B., & Liu, X.-C. (2007, June). Statistical characteristics of the day-to-day variability in the geomagnetic Sq field: DAY-TO-DAY-VARIABILITY OF Sq. *J. Geophys. Res.*, *112*(A6), n/a–n/a. Retrieved 2022-07-06, from http://doi.wiley.com/10.1029/2006JA012059 doi: 10.1029/2006JA012059

Cousins, E. D. P., Matsuo, T., & Richmond, A. D. (2015, July). Mapping high-latitude ionospheric electrodynamics with SuperDARN and AMPERE. *J. Geophys. Res. Space Physics*, *120*(7), 5854–5870. Retrieved 2022-07-14, from `https://onlinelibrary.wiley.com/doi/10.1002/2014JA020463` doi: 10.1002/2014JA020463

Cousins, E. D. P., Matsuo, T., Richmond, A. D., & Anderson, B. J. (2015, August). Dominant modes of variability in large-scale Birkeland currents. *J. Geophys. Res. Space Physics*, *120*(8), 6722–6735. Retrieved 2022-07-06, from `https://onlinelibrary.wiley.com/doi/10.1002/2014JA020462` doi: 10.1002/2014JA020462

Eastes, R. W., McClintock, W. E., Burns, A. G., Anderson, D. N., Andersson, L., Aryal, S., … Woods, T. N. (2020, July). Initial Observations by the GOLD Mission. *J. Geophys. Res. Space Physics*, *125*(7). Retrieved 2022-07-05, from `https://onlinelibrary.wiley.com/doi/10.1029/2020JA027823` doi: 10.1029/2020JA027823

Fang, T., Fuller-Rowell, T., Yudin, V., Matsuo, T., & Viereck, R. (2018, November). Quantifying the Sources of Ionosphere Day-To-Day Variability. *J. Geophys. Res. Space Physics*, *123*(11), 9682–9696. Retrieved 2022-07-05, from `https://onlinelibrary.wiley.com/doi/abs/10.1029/2018JA025525` doi: 10.1029/2018JA025525

Flynn, S., Knipp, D. J., Matsuo, T., Mlynczak, M., & Hunt, L. (2018, May). Understanding the Global Variability in Thermospheric Nitric Oxide Flux Using Empirical Orthogonal Functions (EOFs). *J. Geophys. Res. Space Physics*, *123*(5), 4150–4170. Retrieved 2022-07-06, from `https://onlinelibrary.wiley.com/doi/10.1029/2018JA025353` doi: 10.1029/2018JA025353

Forbes, J. M. (2021, March). Atmosphere-Ionosphere (A-I) Coupling by Solar and Lunar Tides. In W. Wang, Y. Zhang, & L. J. Paxton (Eds.), *Geophysical Monograph Series* (1st ed., pp. 157–181). Wiley. Retrieved 2022-07-05, from `https://onlinelibrary.wiley.com/doi/10.1002/9781119815631.ch9` doi: 10.1002/9781119815631.ch9

Forbes, J. M., Maute, A., Zhang, X., & Hagan, M. E. (2018, September). Oscillation of the Ionosphere at Planetary-Wave Periods. *J. Geophys. Res. Space Physics*, *123*(9), 7634–7649. Retrieved 2022-07-05, from `https://onlinelibrary.wiley.com/doi/10.1029/2018JA025720` doi: 10.1029/2018JA025720

Forbes, J. M., Zhang, X., & Maute, A. (2020, 5). Planetary wave (pw) generation in the thermosphere driven by the pw-modulated tidal spectrum. *Journal of Geophysical Research: Space Physics*, *125*. Retrieved from `https://onlinelibrary.wiley.com/doi/abs/10.1029/2019JA027704` doi: 10.1029/2019JA027704

Gan, Q., Eastes, R. W., Burns, A. G., Wang, W., Qian, L., Solomon, S. C., … McClintock, W. E. (2020). First synoptic observations of geomagnetic storm effects on the global-scale oi 135.6-nm dayglow in the thermosphere by the gold mission. *Geophysical Research Letters*, *47*(3), e2019GL085400. Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019GL085400` (e2019GL085400 2019GL085400) doi: https://doi.org/10.1029/2019GL085400

Gan, Q., Yue, J., Chang, L. C., Wang, W. B., Zhang, S. D., & Du, J. (2015, July). Observations of thermosphere and ionosphere changes due to the dissipative 6.5-day wave in the lower thermosphere. *Ann. Geophys.*, *33*(7), 913–922. Retrieved 2022-07-05, from `https://angeo.copernicus.org/articles/33/913/2015/` doi: 10.5194/angeo-33-913-2015

Immel, T. J., Craven, J. D., & Nicholas, A. C. (2000, January). An empirical model of the OI FUV dayglow from DE-1 images. *J. Atmos. and Solar-Terrestrial Phys.*, *62*(1), 47-64. doi: 10.1016/S1364-6826(99)00082-6

King, J. H., & Papitashvili, N. E. (2005). Solar wind spatial scales in and comparisons of hourly wind and ace plasma and magnetic field data. *Journal of Geophysical Research: Space Physics*, *110*(A2). Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2004JA010649 doi: https://doi.org/10.1029/2004JA010649

Liou, K. (2005). Neutral composition effects on ionospheric storms at middle and low latitudes. *J. Geophys. Res.*, *110*(A5), A05309. Retrieved 2022-07-05, from http://doi.wiley.com/10.1029/2004JA010840 doi: 10.1029/2004JA010840

Liu, H., Yamazaki, Y., & Lei, J. (2021, March). Day-to-Day Variability of the Thermosphere and Ionosphere. In W. Wang, Y. Zhang, & L. J. Paxton (Eds.), *Geophysical Monograph Series* (1st ed., pp. 275–300). Wiley. Retrieved 2022-07-05, from https://onlinelibrary.wiley.com/doi/10.1002/9781119815631.ch15 doi: 10.1002/9781119815631.ch15

Lu, G., Goncharenko, L., Nicolls, M. J., Maute, A., Coster, A., & Paxton, L. J. (2012, August). Ionospheric and thermospheric variations associated with prompt penetration electric fields: VARIATIONS ASSOCIATED WITH PPEF. *J. Geophys. Res.*, *117*(A8), n/a–n/a. Retrieved 2022-07-05, from http://doi.wiley.com/10.1029/2012JA017769 doi: 10.1029/2012JA017769

Matzka, J., Stolle, C., Yamazaki, Y., Bronkalla, O., & Morschhauser, A. (2021). The geomagnetic kp index and derived indices of geomagnetic activity. *Space Weather*, *19*(5), e2020SW002641. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2020SW002641 (e2020SW002641 2020SW002641) doi: https://doi.org/10.1029/2020SW002641

Oberheide, J., Pedatella, N. M., Gan, Q., Kumari, K., Burns, A. G., & Eastes, R. W. (2020). Thermospheric composition o/n response to an altered meridional mean circulation during sudden stratospheric warmings observed by gold. *Geophysical Research Letters*, *47*(1), e2019GL086313. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019GL086313 (e2019GL086313 10.1029/2019GL086313) doi: https://doi.org/10.1029/2019GL086313

Pancheva, D., Mukhtarov, P., & Smith, A. K. (2013). Climatology of the migrating terdiurnal tide (tw3) in saber/timed temperatures. *Journal of Geophysical Research: Space Physics*, *118*(4), 1755-1767. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/jgra.50207 doi: https://doi.org/10.1002/jgra.50207

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., … Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Preisendorfer, R. W., & Mobley, C. D. (1988). *Principal component analysis in meteorology and oceanography* (No. 17). Amsterdam ; New York : New York, NY, U.S.A: Elsevier ; Distributors for the U.S. and Canada, Elsevier Science Pub. Co.

Prölss, G. W., Brace, L. H., Mayr, H. G., Carignan, G. R., Killeen, T. L., & Klobuchar, J. A. (1991). Ionospheric storm effects at subauroral latitudes: A case study. *Journal of Geophysical Research: Space Physics*, *96*(A2), 1275-1288. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/90JA02326 doi: https://doi.org/10.1029/90JA02326

Rishbeth, H. (1998, September). How the thermospheric circulation affects the ionospheric F2-layer. *Journal of Atmospheric and Solar-Terrestrial Physics*, *60*(14), 1385–1402. Retrieved 2022-07-05, from https://linkinghub.elsevier.com/retrieve/pii/S1364682698000625 doi: 10.1016/S1364-6826(98)00062-5

Rishbeth, H., & Müller-Wodarg, I. C. F. (1999, June). Vertical circulation and thermospheric composition: a modelling study. *Annales Geophys.*, *17*(6), 794-805. doi: 10.1007/s00585-999-0794-x

Schmölter, E., Berdermann, J., & Codrescu, M. (2021, February). The Delayed Ionospheric Response to the 27-day Solar Rotation Period Analyzed With GOLD and IGS TEC Data. *J. Geophys. Res. Space Physics*, *126*(2). Retrieved 2022-07-05, from `https://onlinelibrary.wiley.com/doi/10.1029/2020JA028861` doi: 10.1029/2020JA028861

Stolle, C., Manoj, C., Lühr, H., Maus, S., & Alken, P. (2008). Estimating the daytime equatorial ionization anomaly strength from electric field proxies. *Journal of Geophysical Research: Space Physics*, *113*(A9). Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2007JA012781` doi: https://doi.org/10.1029/2007JA012781

Strickland, D. J., Cox, R. J., Meier, R. R., & Drob, D. P. (1999). Global O/N_2 derived from DE-1 FUV dayglow data: Technique and examples from two storm periods. , *104*, 4251–4266.

Strickland, D. J., Majeed, T., Evans, J. S., Meier, R. R., & Picone, J. M. (1997). Analytical representation of g factors for rapid, accurate calculation of excitation rates in the dayside thermosphere. *Journal of Geophysical Research: Space Physics*, *102*(A7), 14485-14498. Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/97JA00943` doi: https://doi.org/10.1029/97JA00943

Talaat, E. R., Yee, J.-H., & Zhu, X. (2001). *Observations of the 6.5-day wave in the mesosphere and lower thermosphere* (Vol. 106). doi: 10.1029/2001JD900227

Torrence, C., & Compo, G. P. (1998, January). A Practical Guide to Wavelet Analysis. *Bulletin of the American Meteorological Society*, *79*(1), 61-78. doi: 10.1175/1520-0477(1998)079\textless{}0061:APGTWA\textgreater{}2.0.CO;2

Wu, Y.-J. J., Harding, B. J., Triplett, C. C., Goel, D., Chihoko, C., Lieberman, R., ... Heelis, R. (2021). Multi-parameter observations of the quasi-6-day wave and corresponding oscillations in the upper atmosphere observed by icon and gold. *AGU Fall Meeting Abstracts*, *2021*.

Yue, J., & Wang, W. (2014, March). Changes of thermospheric composition and ionospheric density caused by quasi 2 day wave dissipation: QUASI-TWO-DAY WAVE CAUSED MIXING. *J. Geophys. Res. Space Physics*, *119*(3), 2069–2078. Retrieved 2022-07-05, from `http://doi.wiley.com/10.1002/2013JA019725` doi: 10.1002/2013JA019725

Yue, J., Wang, W., Ruan, H., Chang, L. C., & Lei, J. (2016, April). Impact of the interaction between the quasi-2 day wave and tides on the ionosphere and thermosphere. *J. Geophys. Res. Space Physics*, *121*(4), 3555–3563. Retrieved 2022-07-05, from `https://onlinelibrary.wiley.com/doi/abs/10.1002/2016JA022444` doi: 10.1002/2016JA022444

Zhang, Y. (2003). Negative ionospheric storms seen by the IMAGE FUV instrument. *J. Geophys. Res.*, *108*(A9), 1343. Retrieved 2022-07-05, from `http://doi.wiley.com/10.1029/2002JA009797` doi: 10.1029/2002JA009797