The Spatio-temporal Variation Pattern of Thermospheric Mass Density Revealed by Co-clustering

Zhaoyuan Yu[1,2,3], Zhenxia Liu[1], Zengjie Wang[1], Lingling Sun[1], Hong Gao[4], Wen Luo[1,2,3], Linwang Yuan*[1,2,3],

[1]Ministry of Education, Key Laboratory of Virtual Geographic Environment (Nanjing Normal University), Nanjing, China;
[2]Jiangsu Center for Collaborative Innovation in Geographical Information Resource Development and Application, Nanjing, China;
[3]State Key Laboratory Cultivation Base of Geographical Environment Evolution (Jiangsu Province), Nanjing, China;
[4]Faculty of Geomatics, Lanzhou Jiaotong University,Lanzhou, Gansu Province, China;

Key Points:

- Analysis of the spatio-temporal structural characteristics of thermospheric mass density variations using GOCE satellite data
- Exploration on the structures of spatial changes of thermospheric mass density in long time series using a co-clustering method
- Comparative analysis of the influence between different temporal and spatial structural characteristics of thermospheric mass density

Corresponding author: Linwang Yuan, `yuanlinwang@njnu.edu.cn`

Abstract

The spatio-temporal distribution characteristics of thermospheric mass density have been given more attention with an increasing demand for spacecraft launches and low Earth orbital prediction. More and more patterns of spatial structure and temporal variation are being discovered. Notwithstanding these developments, the study of spatio-temporal coupling in characteristics analysis remains quite limited. In this study, we use a co-clustering method to explore and analyze the spatio-temporal coupling structural characteristics of thermospheric mass density. The processed GOCE satellite dataset is divided into 5 temporal clusters and 20 spatial clusters by the co-clustering method. In terms of spatial structure, the density has an obvious zonal distribution structure and hemispheric asymmetry. Moreover, due to the influence of the Earth's magnetic field, there is an average angle about 2.00° between the band structure and the latitudinal circle. In terms of temporal structure, the temporal patterns of density can be grouped into five period types, namely the quiet period, the moderate activity period, the event period, the oscillation period and the recovery period. And significant positive correlation can be found between the F10.7 indices and the temporal density variation. This study explores the spatial structure and temporal pattern of thermospheric mass density and its driving forces from the perspective of spatio-temporal coupling based on a statistical method, which can provide a new idea of spatio-temporal coupling method for spatio-temporal evolution of thermospheric mass density.

1 Introduction

Thermospheric mass density has a significant impact on spacecraft launches and low Earth orbital prediction since it is the primary source of the atmospheric drag(Emmert, 2015; Doornbos et al., 2008), which has complex spatio-temporal variations and response characteristics commonly driven by solar ultraviolet irradiance, solar wind, energetic particles from the magnetosphere, and waves originating from the lower atmosphere(P. Wang et al., 2022). It is essential to better analyze and explore the spatio-temporal characteristics of the thermospheric mass density to guide the space exploration.

With continuous research on the observational data of the thermospheric, the variation pattern of thermospheric mass density is not only related to temporal changes (The seasonal changes, the semiannual changes, daily changes, etc.) but also closely related to spatial changes(Changes in latitude and longitude, changes in polar regions), as well as anomalies such as thermospheric storms, night enhancement, equatorial anomalies, longitude distribution, and irregular polar structures(Emmert, 2015; Liu et al., 2007; Qian & Solomon, 2012). In terms of spatial structure, Walterscheid et al. (Walterscheid & Crowley, 2015) studied the spatial structure of thermospheric mass density during geomagnetic calm and active periods using the TMEGCM model and provided a detailed interpretation of its low and high-density regions. Thermospheric mass density exhibits hemispheric asymmetry (Meier et al., 2015), and overall thermospheric mass density exhibits a slightly higher structure in the southern hemisphere than in the northern hemisphere, and the maximum value occurs in the equatorial region. In terms of temporal pattern, related studies (Lei et al., 2011; Danilov et al., 1992; Haris et al., 1994; Weng et al., 2018) show the significant variation in atmospheric mass density with diurnal, seasonal and annual/semi-annual variation, especially, both amplitude and phase of the thermospheric mass density seasonal variations have strong solar activity dependences(Weng et al., 2018). Many scholars have used empirical models and spatial statistics to reveal the characteristics of the temporal variation of the thermospheric(Yuan et al., 2019; Mehta et al., 2017; Ren et al., 2021), but the temporal and spatial characteristics are often not closely linked in these analyses. Considering the universal connection between temporal and spatial features within geographic phenomena(Deng et al., 2013), the density distribution characteristics of spatial and temporal are not independent but coupled with each other. Therefore, it is especially important to adopt an analytical method that can establish the coupling of temporal and spatial features to analyze the temporal and spatial evolution characteristics of thermospheric atmospheric density.

As one of the common methods to study the spatial and temporal differentiation of geographical phenomena, clustering analysis groups the data with similarities to identify the internal structure and numerical divergence patterns of data sets, and discovers the divergence pattern within the data(Andrienko et al., 2009), it is particularly effective for exploring the hidden features within geographical phenomena. Clustering methods are particularly useful in atmospheric and space weather studies because the results of clustering over large regions and long time scales are more practical than those summarized from short time observations of several locations(Zirlewagen & Wilpert, 2010). The spatial clustering and temporal clustering methods have been used by many scholars to study the atmospheric spatial and temporal divergence patterns. For example, Wang et al. investigated the spatial and temporal patterns of atmospheric ozone pollution by Spatio-temporal Shared Nearest Neighbor clustering algorithm(S. Wang et al., 2013), Zhao et al.used clustering to explore anomaly correlation and to reveal how global climate model(GCM) forecasts perform as time progresses(Zhao et al., 2017). Ma et al. clustered potential flaring active regions by applying Distance Density clustering on individual parameters and further organized the clustering results into a multivariate time series decision tree and addressed the problem of flare prediction(Ma et al., 2017). However, the spatial divergence patterns obtained using spatial clustering methods alone do not reveal time-varying behavior, and vice versa. Thus, a clustering method that can obtain spatial and temporal divergence patterns is needed to explore the spatial and temporal coupling characteristics. Co-clustering is able to show its variation with time and space by performing both temporal and spatial clustering of the thermospheric atmospheric density, establishing its coupling characteristics in both temporal and spatial dimensions. Therefore, we use a co-clustering algorithm to cluster thermospheric mass density data, and study the spatio-temporal structure and patterns of thermospheric mass density under the interaction and interplay of time and space.

In the first section, we systematically analyze changes in spatio-temporal characteristics of thermospheric mass density. The second section briefly introduces the data sources and overall framework of the methods. The third section analyzes the temporal and spatial characteristics of thermospheric mass density based on the spatio-temporal co-clustering, and discusses the dynamic evolution of spatial structure and temporal patterns and driving force based on the co-cluster result. Finally, we summarized conclusion, and explained the feasibility of our results.

2 Data and Method

2.1 GOCE Data

GOCE (Gravity field and steady-state ocean circulation explore) satellite thermospheric mass density data with better observation continuity and quality, have been used to study the characteristics of gravity wave distribution at various scales and the seasonal variation distribution pattern, increasing our understanding of the variation pattern of the thermospheric mass density (Forbes et al., 2016; Liu et al., 2017). We use the GOCE satellite data from November 1, 2009, to October 20, 2013, including 1450 days. Its latitude coverage is from 83. 5° S to 83° N and the average altitude naturalized to 270km altitude. Due to the inconsistency between the sampling interval of the satellite and the revisit period of the satellite, it is difficult to have multiple consecutive data at the same location, which results in extremely sparse satellite data in space. So we divide the globe into a 2°x2° grid so as to organize discrete GOCE satellite data regularly, and treat one day of sample points within the same grid as one data set. If there are multiple sample points on the same grid, the average value of the sampling points is taken as thermospheric mass density. In addition, since the thermosphere has a thickness, the whole thermosphere is a three-dimensional space. We stretch the three-dimensional thermospheric space into a two-dimensional thermospheric plane to study the horizontal structural characteristics of the thermospheric mass density. Then, we divide the globe into 16,200 grids by unfolding the two-dimensional plane as a 2°x2° grid to achieve one-dimensional spatial properties. Finally, one day as a basic unit in temporal dimension. In this way, the dimensional division of spatio-temporal co-clustering is achieved.

2.2 Overall framework of the methods

After preprocessing the data, the data are analyzed using co-clustering to explore the internal spatial and temporal characteristics, and divided the spatial and temporal clusters. Then, geoderector is used to determine the most appropriate number of clusters. For the co-clustering results, they are spread spatially to explore their spatial characteristics and the angle of the zonal distribution structure, and spread temporally to explore the temporal characteristics and compared with the relevant space weather indices.

2.2.1 Co-clustering

Co-clustering is used to explore the internal spatial and temporal characteristics, and divided the spatial and temporal clusters. Clustering analysis is an important unsupervised learning method. The essential idea of clustering analysis is to divide data objects with similarities into one group. It divides the data set into several groups by a certain criterion (such as distance or similarity of objects), and each group is called a cluster. Traditional clustering methods can be divided into two types: one is to cluster samples based on all variables, and the other is to cluster variables based on all samples(Gore, 2000). These two types of methods only cluster samples or variables, so they are often referred to as unidirectional or one-way clustering. Different from one-way clustering,co-clustering is the clustering of data blocks (subsets of data), which is based on the similarity of element values within the data matrix, forming a sub-matrix partitioning scheme so that the sub-matrices elements are as similar as possible, thus achieving simultaneous clustering in both spatial and temporal dimensions.

Both temporal clustering and spatial clustering are one-way clustering, which is only based on a certain dimension. Figure 1(a) shows thermospheric mass density example data (The values do not represent the real density), Figure 1(b) treats space as an object and time as its property to obtain a spatial group with similarity in time, Figure 1(c) treats time as an object and space as its property to obtain temporal group with similarity in space. Unlike one-way clustering, two-way clustering can cluster both time and space, dividing data that are similar in both time and space into the same spatio-temporal group, in which both spatial and temporal values are similar, as in Figure 1(d). Due to the sparseness of data, the clustering method we used is Biclustermd (Li et al., 2020; Reisner et al., 2019), which is suitable for datasets with missing values and proved to have good performance for very sparse data in an agricultural application and a movie rater application(Li et al., 2020).

2.2.2 Geodetector

Geodetector is used to determine the most appropriate number of clusters. Geodetector is a statistical method to detect spatial dissimilarity and to reveal the driving forces behind it(J. Wang & Xu, 2017), which has been widely used in natural sciences, social sciences, environmental sciences and human health(Dong et al., 2017; Ju et al., 2016; Gao et al., 2016). In this study, the reliability of co-clustering results is assessed by the q-statistic

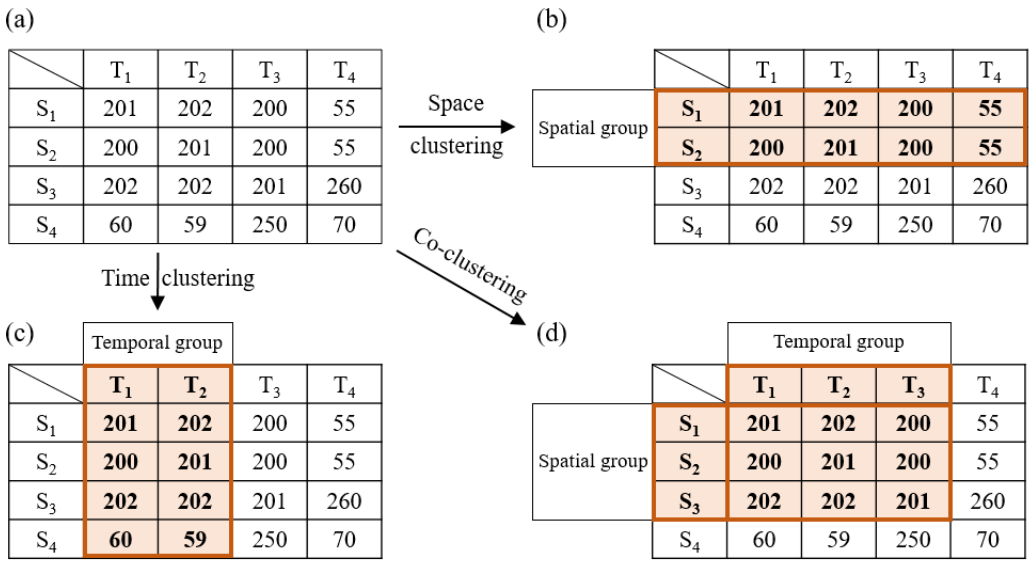

Figure 1. Schematic diagram of co-clustering (Modified from (Wu et al., 2020)). (a) Thermospheric mass density data matrix, T_i denotes the time dimension, S_i denotes the spatial dimension. (b) A group with temporal similarity obtained by spatial clustering. (c) A group with spatial similarity obtained by temporal clustering. (d) Co-clustering to obtain spatio-temporal groups with similarity in both time and space dimension.

in geodetector to measure, and the expression is as follow:

$$\mathrm{q} = 1 - \frac{\sum\limits_{\mathrm{h}=1}^{\mathrm{L}} \mathrm{N_h}\sigma_{\mathrm{h}}^2}{\mathrm{N}\sigma^2}$$

Where: L is the strata of variables,that is cluster in this paper; N_h and N are the number of units in strata h and the whole area, respectively; σ_h^2 and σ^2 are the squares in strata h and the whole area, respectively. The value range of q is [0,1]. Larger value of q indicates more pronounces heterogeneity.

2.2.3 Calculation of angle of zonal distribution structure

The process of calculating the angle of zonal distribution structure is divided into four main steps. Firstly, for structures composed of clusters of the same spatial clusters, the average of the latitudes on the same longitude is obtained to get the center point $A(X_{lon}, Y_{lat})$. Secondly, binomial fitting is carried out with the center points of these different longitudes. The vertex of binomial is the lowest point or the highest point of the whole structure, and the plane passing through this vertex is the axial plane. Thirdly, these centers are projected on the plane where the axis plane is located, and their corresponding positions on the plane are obtained. For a given central point $A(X_{lon}, Y_{lat})$, the equation for its projection onto the axis plane is as follows:

$$\begin{cases} X = R * (1 - \cos X_{lon}) \\ Y = R * \sin Y_{lat} \end{cases}$$

Where X_{lon} is the longitude of the center point A; Y_{lat} is the latitude of the center point A; X is the horizontal coordinate of the projection of the center point A on the axis plane;

Y is the vertical coordinate of the projection of the center point A on the axis plane; X_0 is the longitude of the longitude circle where the axis plane is located; R is the radius of the earth. Finally, the slope of the projection of the zonal distribution structure on the axial plane is fitted according to the projected coordinates, and the angle is calculated by the slope. Considering that the zonal distribution structure is symmetrical about the axis plane, each zonal distribution structure needs to find the inclination angle on both sides of the axis plane by this method, and the average value of the inclination angle on both sides is used as the inclination angle of this zonal distribution structure.

3 Results and Discussion

3.1 Spatio-temporal pattern

Based on the pre-processed two-dimensional spatio-temporal dataset of thermospheric mass density, the data is organized into a data matrix, where the rows represent total 14940 (83×180) grids for the spatial dimension and the columns represent 1450 days for the temporal dimension. The analysis is then performed using a co-clustering method in Section 2.2. To determine the spatial clusters and temporal clusters, clustering analysis is performed using the number of spatial clusters and temporal clusters with step lengths of 2 from 10 to 40 and 1 from 3 to 7, respectively. The clustering results are evaluated using the geodetector q-value in Section 2.3. Finally, the results show that the clustering effect is better when the temporal clusters is 5 and the spatial clusters is 20. At this time, the q-value of the geodetector is 0.7948, indicating that the result has good differentiation(Dong et al., 2017).

The co-clustering results are shown in Figure 2(a). The horizontal axis represents 14940 grids in the dataset, arranged in the order of belonging to the same spatial clusters, the vertical axis represents the 1450 days in the dataset, arranged in the order of belonging to the same temporal clusters. The black solid lines represent the boundary of different spatial clusters and temporal clusters. And the color indicates the mass density value, especially, white means the data is empty because of data sparsity. It can be seen that the density values of the same cluster are relatively close, the boundaries between different clusters are relatively obvious, which means that co-clustering analysis can be a good way to construct spatial and temporal characteristics of thermospheric mass density considering the sparsity of the data. Since both temporal and spatial parameters are considered in the clustering calculation process, the results of co-clustering well establish the coupling relationship between temporal and spatial dimensions, which is obviously different from the traditional study of thermospheric mass density in which only its spatial structure or temporal characteristics are considered alone. It solves the shortcoming of uncoupled spatio-temporal characteristics in the traditional analysis, which is significant for the construction of a spatio-temporal coupled thermospheric analysis model and discovering more structural characteristics about the thermospheric mass density from the perspective of spatio-temporal coupling.

The results of spatio-temporal co-clustering in Figure 2(a) are expanded in space to obtain the spatial pattern and evolutionary dynamics of thermospheric mass density as shown in Figure 2(b). The spatial structure of thermospheric mass density at different time periods in Figure 2(b) all show obvious latitudinal divergence, that is, the spatial structure is dominated by latitudinal variation and shows the overall characteristics of higher density at the poles and lower density at the equator. And the density change in the southern hemisphere is more pronounced than that in the northern hemisphere.

The results of spatio-temporal co-clustering in Figure 2(a) are expanded in time, and the temporal clustering change curves are plotted according to the five temporal clusters, as shown in Figure 2(c). Curve A represents the original temporal clustering results in co-clustering, and curve B is obtained by removing outliers which is lager differences in cluster with adjacent time, and the same height indicates the same groups, which the height from low to high is Cluster1, Cluster2, Cluster3, Cluster4, Cluster5. Since the purpose of co-clustering is to measure the similarity of the density in the temporal dimension and the spatial dimension at the same time, so the group of time does not show continuous segmentation under the influence of space. Figure 2(c) shows the short-period variation in thermospheric mass density over time periods.

Throughout the study period, there are some series of time periods with overall high thermospheric mass density, which indicates the existence of external driving forces to make the density change with time. And from the spatial perspective, there are obvious spatially divergent characteristics of the thermospheric atmospheric density. In the following, we will reveal its structural features from the spatial and temporal dimensions based on the results of co-clustering.

3.2 Spatial structure and dynamic evolution

The co-clustering method aggregates the density values with spatial-temporal similarity into the same group. The result of Figure 2(b) well reflects the macroscopic spatial structure characteristics of the thermospheric atmosphere at long time scales, i.e., thermospheric mass density shows hemispheric asymmetry, thermospheric mass density in the southern hemisphere is slightly higher than that in the northern hemisphere, and the maximum value occurs in the equatorial region(Meier et al., 2015). This conclusion is similar to many studies on the spatio-temporal pattern of the thermospheric atmosphere, such as zonal distribution structure (Emmert, 2015; Liang et al., 2015), hemispheric asymmetry (Meier et al., 2015), and so on.

It is obvious through Figure 2(b) that there is an angle between the zonal distribution structure of thermospheric mass density and latitudinal circle. Based on this idea, the angle between the approximate axis of the zonal distribution structure and latitudinal circle is obtained by the method in section 2.4, and the results are shown in Table 1(considering the deformation in polar regions, only the complete zonal distribution structure in the range of 60°N to 60°S is selected here). The results of Table 1 show that the multiple zonal distribution structure all exhibit a certain angle of entrainment(average angle about 2.00°), and the axial plane tends to fluctuate around the line from 12°W to 168°E. This phenomenon is close to the dip of the Earth's magnetic axis(70°W to 110°E.), it is statistically significant to show that the spatial density distribution of the thermospheric is influenced not only by various types of solar activity, but also by the Earth's own magnetic field. This result is coupled with studies on the Earth's magnetic field(Laundal et al., 2017; Lei et al., 2012). Not only can the thermospheric influence the distribution of the Earth's magnetic field through solar wind and ionization, but the Earth's magnetic field can also influence the density characteristics of thermospheric by acting on charged particles.

3.3 Temporal patterns and driving force

Based on Figure 2(c), it can be concluded that thermospheric mass density shows a periodic variation under the action of external driving forces. It has been shown by related studies that thermospheric mass density is closely related to changes in factors such as solar activity and geomagnetic activity, such as solar EUV and UV radiation, high-energy particle deposition in the polar regions, magnetospheric plasma convection, and various atmospheres uploaded from the middle and lower atmospheres to the upper atmosphere(Rostoker et al., 1998; Baker & Kanekal, 2008; Liu & Lühr, 2005). This external energy will heat the thermospheric atmosphere, causing the ionization and decomposition of various components in the thermospheric, and then driving the changes in thermospheric mass density.

F10.7 and Ap are two typical space weather indices that are often considered as important driving force of thermospheric changes, so it is important to explore the relationship between these two indices and the co-clustering results. The joint analysis of the density clustering characteristic curves and the associated space weather indices resulted in Figure 3 and Table 2, Table 3. From the Figure 3, it can be seen that the tem-

Table 1. The angle between the approximate axis of the zonal distribution structure and the latitudinal circle

Cluster	Angle (°)	Axial Plane
Cluster2N	2.74	29° W to 151° E
Cluster2S	2.01	9° W to 171° E
Cluster3S	1.95	7° W to 173° E
Cluster4N	1.81	3° W to 177° E
Cluster4S	1.43	5° W to 175° E
Cluster5S	1.62	7° W to 173° E
Cluster7S	1.31	15° W to 165° E
Cluster8N	3.94	27° W to 153° E
Cluster9S	1.15	3° W to 177° E
Average	2.00	12° W to 168° E

poral clustering curve and the F10.7 curve fit well, when the density is at Cluster1, F10.7 is low overall, at Cluster2, F10.7 is moderate overall, when Cluster3, F10.7 is high overall, when Cluster4, F10.7 is in a fluctuating period overall and fluctuate in a wide range, while Cluster5 is in a time series that belongs to a period immediately following the large fluctuation of Cluster4 after the small fluctuation period. Taking the clustering results of the three classes Cluster1, Cluster2, and Cluster3 (Considering that the values in cluster4 and cluster5 fluctuate widely and represent a fluctuating period, only the cluster1, cluster2, cluster3 that represent different densities are considered when establishing the correlation link) characterizing the density size and the F10.7 curve for correlation analysis, the Pearson correlation coefficient is calculated to obtain $p = 0.755$, it suggests a significant positive correlation between the temporal patterns of thermospheric mass density and F10.7. The curve characteristics shown in Figure 3 are also more consistent with the statistical results of F10.7 in Table 2. This series of features indicates that the temporal patterns of thermospheric mass density are strongly influenced by the solar activity on the time scale, and the overall temporal density characteristics are closely related to the variation of the F10.7, and to some extent, the variation pattern of the F10.7 index can be used to characterize the temporal density variation pattern of the thermospheric mass density.

According to the comparing results and the intensity of the solar activity represented by F10.7, the five temporal clusters cluster1, cluster2, cluster3, cluster4, and cluster5 of the thermospheric mass density can be classified into the quiet period, the moderately active period, the event period, the oscillation period and the recovery period, respectively, which can characterize well the effect of solar activity on thermospheric mass density. This shows that co-clustering can well analyze the temporal characteristics of thermospheric mass density.

Combining Figure 3 and Table 3, it can be found that there is no obvious correlation between the clustering result and Ap, where $p = 0.065$ (Taking the clustering results of the three classes Cluster1, Cluster2, and Cluster3 characterizing the density size and the AP curve for correlation analysis). This shows that geomagnetic activity has no direct influence on the temporal variation of the thermospheric mass density. Combined with spatial analysis, it can be seen that geomagnetic field has obvious influence on the spatial pattern of characteristics in the thermospheric mass density, but it has no obvious driving effect on the density change in temporal scale.

Table 2. F10.7 intensive statistics under different clusters

Cluster	Total (d)	low	Moderate	High	Very High
$C1$	577	252	280	45	0
$C2$	366	0	55	290	21
$C3$	273	0	27	212	34
$C4$	92	0	7	82	3
$C5$	142	0	17	125	0

Table 3. AP statistics under different clusters

Cluster	Total (d)	Quiet	Unsettled	Active	Minor Storm	Major Storm
$C1$	577	442	101	28	5	1
$C2$	366	205	89	55	15	2
$C3$	273	211	40	14	8	0
$C4$	92	61	18	10	2	1
$C5$	142	84	32	23	0	3

4 Conclusion

In this paper, the spatio-temporal coupling characteristics of thermospheric mass density are analyzed from the spatio-temporal correlation of natural phenomena features, using co-clustering method considering both from the temporal and spatial dimensions, exploring dynamic evolution of spatial structure and driving force of temporal patterns. Temporal structural characteristics are closely related to solar activity. And temporal patterns of thermospheric mass density can be concluded into five periods, namely the quiet period, the moderate activity period, the event period, the oscillation period and the recovery period, which can characterize well the effect of solar activity on thermospheric mass density. Especially, there is a significant positive correlation between the F10.7 indices and the density temporal variation. And the F10.7 indices can be used to some extent to characterize the global thermospheric mass density, considering the difficulty of thermospheric mass density measurement. The spatial structural characteristics show obvious zonal distribution structure, in which the equatorial region and the middle and low latitudes of northern latitude have obvious structures and boundaries, but there is no obvious structure in the southern hemisphere and the poles. There is an average angle about 2.00° between the band structure and the latitudinal circle because of the influence of the Earth's magnetic field. Not only can the thermospheric influence the distribution of the Earth's magnetic field through solar wind and ionization, but the Earth's magnetic field can also influence the density characteristics of thermospheric by acting on charged particles.

At present, this research is only a rough exploration and analysis of the spatio-temporal structure of thermospheric mass density on a large scale. Since the statistical method from data is used in this paper, it is not possible to explain the inner formation mechanism of statistical phenomena, such as the dip angle phenomenon of the band structure. There are still many factors that have not been considered, such as the influence of different parameters (geomagnetic, gravity waves). This paper only analyzes the temporal and spatial characteristics of thermospheric mass density from a holistic perspective, and some local anomalous details need to be further explored. Many spatio-temporal phenomena and spatio-temporal structures are still on the surface, and their internal mechanism needs further explanation.

5 Data Avaliability Statement

Thermospheric mass density data of GOCE used in the present paper are from TUDelft (ftp://thermosphere.tudelft.nl/)

Acknowledgments

ZYY and ZXL contributed equally to this work and should be considered co-first authors. This work was supported in part by the National Natural Science Foundation of China under Grant 42130103, 41976186, 41971404, and the Postgraduate Research & Practice Innovation Program of Jiangsu Province under Grant KYCX22_1578.

References

Andrienko, G. L., Andrienko, N. V., Rinzivillo, S., Nanni, M., & Giannotti, F. (2009). Interactive visual clustering of large collections of trajectories. In *2009 ieee symposium on visual analytics science and technology* (pp. 3–10). doi: https://doi.org/10.1109/VAST.2009.5332584

Baker, D., & Kanekal, S. (2008). Solar cycle changes, geomagnetic variations, and energetic particle properties in the inner magnetosphere. *Journal of Atmospheric and Solar-Terrestrial Physics*, *70*(2-4), 195–206. doi: https://doi.org/10.1016/j.jastp.2007.08.031

Danilov, A., Kalgin, Y. A., & Pokhunkov, A. (1992). Seasonal and latitudinal variations of turbulence in the lower thermosphere. *Advances in Space Research*, *12*(10), 123–129. doi: https://doi.org/10.1016/0273-1177(92)90453-5

Deng, M., Liu, Q., Wang, J., & Shi, Y. (2013). A general method of spatio-temporal clustering analysis. *Science China Information Sciences*, *56*(10), 1–14.

Dong, Y., Qian, X., Qian, X., Chengdong, X., & Yuying, W. (2017). Delineation of the northern border of the tropical zone ofchina's mainland using geodetector. *Acta Geographica Sinica*, *72*(1), 135.

Doornbos, E., Klinkrad, H., & Visser, P. (2008). Use of two-line element data for thermosphere neutral density model calibration. *Advances in Space Research*, *41*(7), 1115–1122.

Emmert, J. T. (2015). Thermospheric mass density: A review. *Advances in Space Research*, *56*(5), 773–824. doi: https://doi.org/10.1016/j.asr.2015.05.038

Forbes, J. M., Bruinsma, S. L., Doornbos, E., & Zhang, X. (2016). Gravity wave-induced variability of the middle thermosphere. *Journal of Geophysical Research: Space Physics*, *121*(7), 6914–6923. doi: https://doi.org/10.1002/2016JA022923

Gao, F., Li, S., Tan, Z., Wu, Z., Zhang, X., Huang, G., & Huang, Z. (2016). Understanding the modifiable areal unit problem in dockless bike sharing usage and exploring the interactive effects of built environment factors. *International Journal of Geographical Information Science*, *35*(9), 1905-1925. doi: https://doi.org/10.1080/13658816.2020.1863410

Gore, P. A. (2000). 11 - cluster analysis. In H. E. Tinsley & S. D. Brown (Eds.), *Handbook of applied multivariate statistics and mathematical modeling* (p. 297-321). San Diego: Academic Press. Retrieved from `https://www.sciencedirect.com/science/article/pii/B9780126913606500124` doi: https://doi.org/10.1016/B978-012691360-6/50012-4

Haris, P. A. T., Stevens, T. D., Maruvada, S., & Philbrick, C. R. (1994). Latitudinal variation of middle atmospheric density and temperature. *Advances in Space Research*, *14*(9), 83–87. doi: https://doi.org/10.1016/0273-1177(94)90120-1

Ju, H., Zhang, Z., Zuo, L., Wang, J., Zhang, S., Wang, X., & Zhao, X. (2016). Driving forces and their interactions of built-up land expansion based on the geographical detector – a case study of beijing, china. *International Journal of Geographical Information Science*, *30*(11), 2188-2207. doi: https://doi.org/10.1080/13658816.2016.1165228

Laundal, K. M., Cnossen, I., Milan, S. E., Haaland, S., Coxon, J., Pedatella, N., ... Reistad, J. P. (2017). North–south asymmetries in earth's magnetic field. *Space Science Reviews*, *206*(1), 225–257. doi: https://doi.org/10.1007/s11214-016-0273-0

Lei, J., Forbes, J. M., Liu, H.-L., Dou, X., Xue, X., Li, T., & Luan, X. (2011). Latitudinal variations of middle thermosphere: Observations and modeling. *Journal of Geophysical Research: Space Physics*, *116*(A12), A12306. doi: https://doi.org/10.1029/2011JA017067

Lei, J., Thayer, J. P., Wang, W., Richmond, A. D., Roble, R., Luan, X., ... Li, T. (2012). Simulations of the equatorial thermosphere anomaly: Field-aligned ion drag effect. *Journal of Geophysical Research: Space Physics*, *117*(A1). doi: https://doi.org/10.1029/2011JA017114

Li, J., Reisner, J., Pham, H., Olafsson, S., & Vardeman, S. (2020). Biclustering with missing data. *Information Sciences*, *510*, 304-316. doi: https://doi.org/10.1016/j.ins.2019.09.047

Liang, J., Donovan, E., Nishimura, Y., Yang, B., Spanswick, E., Asamura, K., ... Redmon, R. (2015). Low-energy ion precipitation structures associated with pulsating auroral patches. *Journal of Geophysical Research: Space Physics*, *120*(7), 5408–5431. doi: https://doi.org/10.1002/2015JA021094

Liu, H., & Lühr, H. (2005). Strong disturbance of the upper thermospheric density due to magnetic storms: Champ observations. *Journal of Geophysical Research: Space Physics*, *110*(A9). doi: https://doi.org/10.1029/2004JA010908

Liu, H., Lühr, H., & Watanabe, S. (2007). Climatology of the equatorial thermospheric mass density anomaly. *Journal of Geophysical Research: Space Physics*, *112*(A5), A05305. doi: https://doi.org/10.1029/2006JA012199

Liu, H., Pedatella, N., & Hocke, K. (2017). Medium-scale gravity wave activity in the bottomside f region in tropical regions. *Geophysical Research Letters*, *44*(14), 7099–7105. doi: https://doi.org/10.1002/2017GL073855

Ma, R., Boubrahimi, S. F., Hamdi, S. M., & Angryk, R. A. (2017). Solar flare prediction using multivariate time series decision trees. In *2017 ieee international conference on big data (big data)* (pp. 2569–2578). doi: https://doi.org/10.1109/BigData.2017.8258216.

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

Meier, R. R., Picone, J. M., Drob, D., Bishop, J., Emmert, J. T., Lean, J. L., ... others (2015). Remote sensing of earth's limb by timed/guvi: Retrieval of thermospheric composition and temperature. *Earth and Space Science*, *2*(1), 1–37. doi: https://doi.org/10.1002/2014EA000035

Qian, L., & Solomon, S. C. (2012). Thermospheric density: An overview of temporal and spatial variations. *Space science reviews*, *168*(1), 147–173. doi: https://doi.org/10.1007/s11214-011-9810-z

Reisner, J., Pham, H., Olafsson, S., Vardeman, S. B., & Li, J. (2019). biclustermd: An r package for biclustering with missing values. *R J.*, *11*(2), 69.

Ren, D., Lei, J., Wang, W., Burns, A., & Luan, X. (2021). Observations and simulations of the peak response time of thermospheric mass density to the 27-day solar euv flux variation. *Journal of Geophysical Research: Space Physics*, *126*(2), e2020JA028756.

Rostoker, G., Skone, S., & Baker, D. N. (1998). On the origin of relativistic electrons in the magnetosphere associated with some geomagnetic storms. *Geophysical Research Letters*, *25*(19), 3701–3704. doi: https://doi.org/10.1029/98GL02801

Walterscheid, R. L., & Crowley, G. (2015). Thermal cell structures in the high-latitude thermosphere induced by ion drag. *Journal of Geophysical Research: Space Physics*, *120*(8), 6837–6850. doi: https://doi.org/10.1002/2015JA021122

Wang, J., & Xu, C. (2017). Geodetector: Principle and prospective. *Acta Geographica Sinica*, *72*(1), 19.

Wang, P., Chen, Z., Deng, X., Wang, J., Tang, R., Li, H., ... Wu, Z. (2022). The prediction of storm-time thermospheric mass density by lstm-based ensemble learning. *Space Weather*, *20*(3), e2021SW002950.

Wang, S., Cai, T., & Eick, C. F. (2013). New spatiotemporal clustering algorithms and their applications to ozone pollution. In *2013 ieee 13th international conference on data mining workshops* (p. 1061-1068). doi: https://doi.org/10.1109/ICDMW.2013.14

Weng, L., Lei, J., Doornbos, E., Fang, H., & Dou, X. (2018). Seasonal variations of thermospheric mass density at dawn/dusk from goce observations. *Annales Geophysicae*, *36*(2), 489–496. doi: https://doi.org/10.5194/angeo-36-489-2018

Wu, X., Cheng, C., Qiao, C., & Song, C. (2020). Spatio-temporal differentiation of spring phenology in china driven by temperatures and photoperiod from 1979 to 2018. *Science China Earth Sciences*, *63*(10), 1485–1498. doi: https://doi.org/10.1007/s11430-019-9577-5

Yuan, L., Jin, S., & Calabia, A. (2019). Distinct thermospheric mass density variations following the september 2017 geomagnetic storm from grace and swarm. *Journal of Atmospheric and Solar-Terrestrial Physics*, *184*, 30-36. doi: https://doi.org/10.1016/j.jastp.2019.01.007

Zhao, T., Liu, P., Zhang, Y., & Ruan, C. (2017). Relating anomaly correlation to lead time: Clustering analysis of cfsv2 forecasts of summer precipitation in china. *Journal of Geophysical Research: Atmospheres*, *122*(17), 9094–9106.

Zirlewagen, D., & Wilpert, K. V. (2010). Upscaling of environmental information: Support of land-use management decisions by spatio-temporal regionalization approaches. *Environmental Management*, *46*(6), 878-893. doi: https://doi.org/10.1007/s00267-010-9468-4

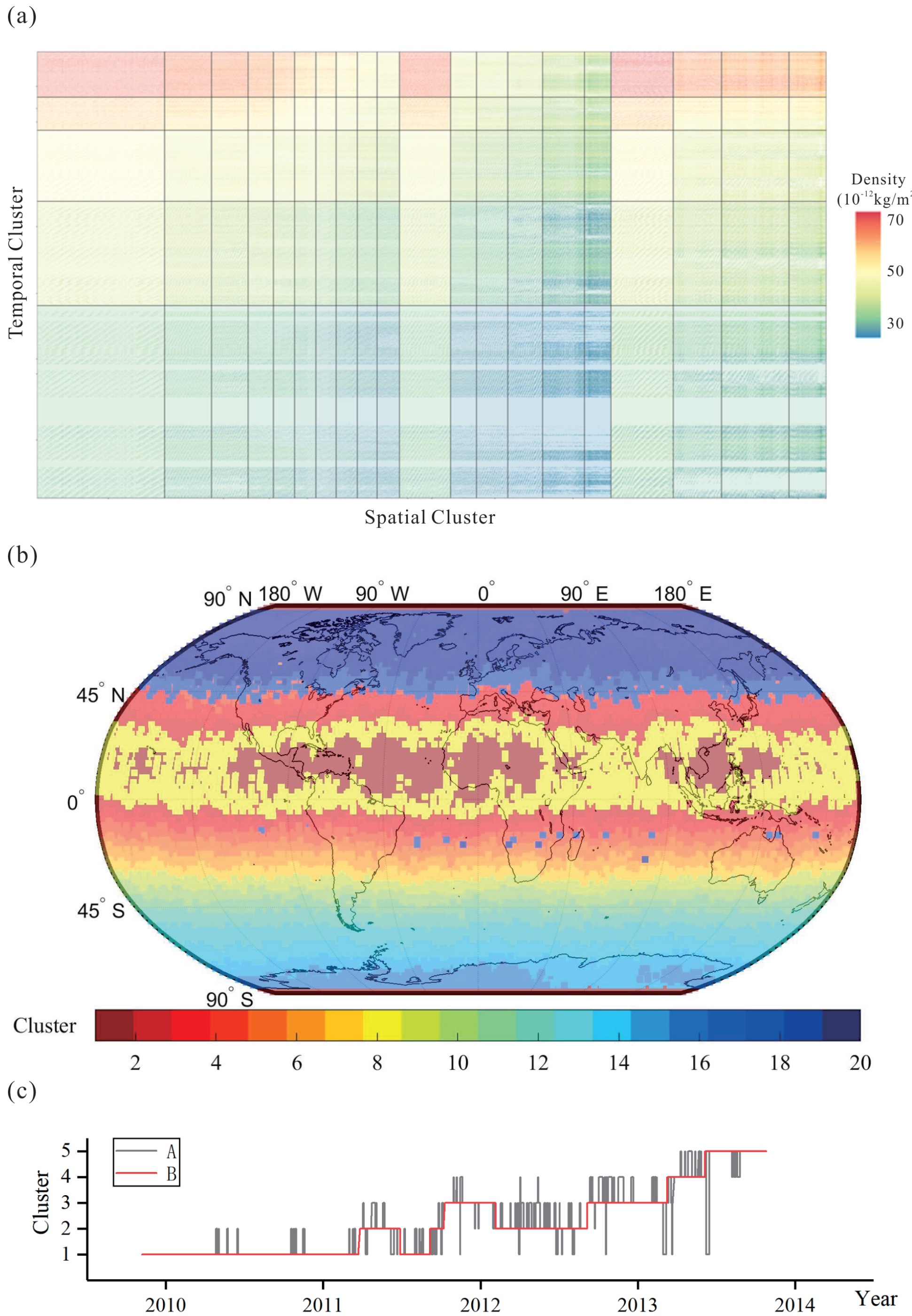

Figure 2. (a) Spatio-temporal co-clustering results of thermospheric mass density, the horizontal axis represents 14940 grids in the dataset, arranged in the order of belonging to the same spatial clusters, the vertical axis represents the 1450 days in the dataset, arranged in the order of belonging to the same temporal clusters, the black solid lines represent the boundary of different groups of time and space, and the color indicates the density value; (b) Spatial clustering result, 1-20 represent 20 groups of spatial clusters; (c) Temporal clustering result, curve A represents original temporal clustering, and curve B represents temporal clustering after removing outliers

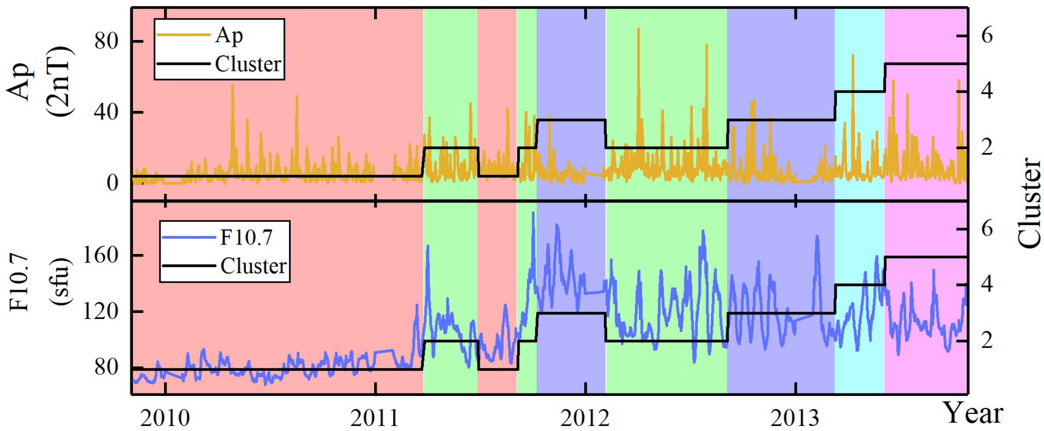

Figure 3. Comparison between temporal patterns and space weather indices