The Correlation Analysis of Atmospheric Model Accuracy Based on the Pearson Correlation Criterion

Wenhui Cui¹, Zhenyu Sun¹, Hao Ma¹, Si Wu¹
¹State Key Laboratory of Astronautic Dynamics, Xi’an 710043, China

Abstract. The atmospheric density above the Beijing area in 2012 is calculated by the DTM2013 atmospheric model and the MSIS00 atmospheric model, and the accuracies of the two models are derived by comparing with the measured data of GOCE in the same region and the same period. The correlations of the accuracies of the two atmospheric models and the variances of their corresponding solar radiation proxies are derived based on the Pearson correlation criterion. It is proved that the accuracy of the two atmospheric models is strongly correlated with the corresponding solar radiation proxy variances in the same period. The solar activity factors of the DTM2013 atmospheric model with higher accuracy are discovered, and the better application value of F30 solar radiation proxy in the atmospheric models is proved.

1. Introduction
Atmospheric drag perturbation has a crucial influence on the accuracy of orbit determination and long-term prediction of low-orbiting satellites. The study of the atmospheric density of the near-earth orbit has been carried out for a long time. Sentman and Cook used the atmospheric damping coefficient to calculated atmospheric density [1][2]. March et al. analysed the differences in various satellite measured atmospheric density datasets [3].

In the study of the effect of solar radiation on atmospheric density, Jacchia pointed out the influence of solar activity on atmospheric density in 1959 [4]. Roemer and Rhoden et al. discovered the significance of solar EUV and MgII proxy in described solar activity [5][6]. Tobiska proposed a new solar radiation proxy E10.7 and claimed the new proxy could significantly improve the prediction accuracy of SME satellite orbital half-length path attenuation instead of the F10.7 proxy in atmospheric models in 2001 [7]. Eastes et al. found that the solar activity cycle (about 27 days) existed in the neutral components of the thermosphere atmosphere based on measured satellite data in 2004 [8]. Bowman et al. proposed two proxies: S10 and Mg10 to represent solar radiation in 2006 [9]. Since F10.7 proxy can achieve more accurate prediction while the other proxies can only be measured by SOHO and other institutions, the NRLMSISE-00 atmospheric model which only uses F10.7 proxy was widely applied. F10.7 proxy has become the standard proxy of atmospheric density model in the early 21st century, and they were continuously used in DTM series models until DTM2013 replaced F10.7 proxy with F30 proxy [10]. NIU J. et al. studied the correlation between F10.7 proxy and the atmospheric density [11], but they did not analyse the relationship between the other solar radiation proxies and atmospheric density, and they also did not analyse the correlation between calculation accuracy of different atmospheric models and their corresponding solar radiation proxy.

2. GOCE atmospheric density data and study area
The GOCE (Gravity field and steady-state Ocean Circulation Explorer) was ESA's earth exploration satellite launched in March 2009 with an orbital inclination of 96.5 degrees. Its scientific exploration
mission lasted from November 2009 to October 2013, and its orbit remained at an altitude of 255km until July 2012. ESA released GOCE satellite measurement data as public products of high-altitude precision atmospheric model [12]. To minimize the change of satellite attitude caused by the space environment disturbance, the GOCE is designed to be a strict symmetrical octagonal prism, and its undamped ion micropropeller provides continuous impetus to compensate the nonconservative force of the satellite. The nonconservative force compensation technology made GOCE maintained its orbit and attitude properly in the flight process, and reduced the influence on satellite gravity gradient observations by the nonconservative force effectively. It further improved the accuracy of the atmospheric density measurement.

The global atmospheric density distribution of GOCE in August 2012 is shown in figure 1. The height range of GOCE in the figure is 250.07~288.82km. The longitude range of GOCE in the figure is –179.998°~180°. The latitude range of GOCE in the figure is –83.444°~83.444°. As can be seen from figure 1, the atmospheric density has a relatively obvious correlation with the latitude of the region in a short period. The atmospheric density of the equatorial region is the highest. With the increase of latitude, the atmospheric density gradually decreases, while the atmospheric density drops sharply in the latitude 70°~80° region.

The changes of F10.7 and F30 solar radiation proxies from 1980 to 2013 are shown in Figure 2. The high-order fitting curve is also drawn in the figure to facilitate the analysis. As can be seen from the figure 2, both solar radiation proxies conform to the 11-year solar activity cycle. The variance comparison between F10.7 and F30 data from 1980 to 2013 is shown in Figure 3. In high solar activity
years, the variance of the F30 proxy is about 50% of the variance of the F10.7 proxy, which means that F30 has better data fitting features and predictability than F10.7.

GOCE atmospheric density data cover a very wide range. The atmospheric density calculation of the atmospheric models usually needed the longitude, latitude, altitude, and local solar time of the region. In the event of drastic changes in the above parameters, it is difficult to analyse the change rule of atmospheric density within a period and the correlation between solar radiation proxies, various atmospheric models and the measured atmospheric density. According to the research of Doornbos and Emmert [13][14], the equatorial anomaly, longitude distribution, seasonal changes and irregular structure in the near-polar area can lead to a large number of abnormal atmospheric density data, so it is important to choose a limited space to analyse the change rule and correlation of atmospheric density with other data. The study area selected in this manuscript is the airspace over Beijing. The longitude range of the area is 115.40°-117.40°. The latitude range of the area is 39.4°-40.4°. The altitude range of the area is 258.819-272.439km.

3. The atmospheric models and the Pearson correlation criterion

Naval Research Laboratory (NRL) proposed the NRLMSISE-00 atmospheric model [15] (abbreviated as the MSIS00 model) in 2000. ATMOP proposed the DTM2013 model [10] on the basis of DTM94 and DTM2000 in 2013. Compared the document [10] and document [15], the original data of the MSIS00 model and the DTM2013 model are collected from the mass spectrometer, incoherent scattering radar, space craft orbit, and acceleration data. The least-square method is used as the fitting method of the two models. The Bates index is used as the temperature profile form of the two models. The Diurnal term, Semidiurnal term, and the terdiurnal term are both considered in the two models. The 6-order spherical harmonic function are both adopted as the latitude term of the two models. The biggest improvement of the DTM2013 model is that the F30 proxy is proposed to replace the F10.7 proxy in the MSIS00 model, while the fitting coefficients of atmospheric models and geomagnetic parameters have not changed significantly. Therefore, it is of great significance for understanding the role of solar radiation proxy in an atmospheric model to analyse the correlation between solar radiation proxies with the accuracies of the two atmospheric models.

The Pearson correlation analysis is a common criterion to measure the correlation between variables, which is defined as [16]:

\[
R(i) = \frac{\sum_{k=1}^{m} (x_{k,i} - \bar{x}_i)(y_k - \bar{y})}{\sqrt{\sum_{k=1}^{m} (x_{k,i} - \bar{x}_i)^2 \sum_{k=1}^{m} (y_k - \bar{y})^2}}
\]

(1)

Where, \(m\) is the number of samples, \(R(i)\) measures the correlation between the feature \(i\) and the class standard. \(x_{k,i}\) is the feature value of the feature \(i\) of the sample \(k\). \(\bar{x}_i\) is the average value of the feature \(i\). \(y_k\) represents the value of the sample \(k\) and \(\bar{y}\) represents the class standard mean value of the sample \(k\). According to the definition formula, the variation range of \(R(i)\) is between -1 and 1.

When \(R(i)=1\), the feature \(i\) is positively linear correlated with the class standard; when \(R(i)=-1\), the feature \(i\) is negatively linear correlated with the class standard. As the degree of correlation between the feature \(i\) and the class standard varies, the value of \(|R(i)|\) changed between 0 and 1. The larger the value of \(|R(i)|\) is, the greater the contribution of the feature \(i\) to classification is. There was a weak correlation between features and class standard when \(|R(i)|<0.3\). There was a low correlation between features and class standard when \(0.3<|R(i)|<0.5\). There was a significant correlation between features and
class standard when \(0.5 < |R(i)| < 0.8\). There was a high correlation between features and class standard when \(0.8 < |R(i)| < 1\) [17].

4. Simulation and Analysis
The correlation between GOCE measured atmospheric density and Solar radiation proxies can be derived by taking the GOCE measured atmospheric density timing sequence as \(x_{k,i}\) and the corresponding sequence of F30 and F10.7 proxies as \(y_k\) in the formula (1). In the same way, supposed the deviation rate of the atmospheric model as \(x_{k,i}\) and the deviation rate of the solar radiation proxy as \(y_k\), then the correlation between the deviation rates of the two atmospheric models and their corresponding solar radiation proxy deviations can be calculated by the formula (1).

The atmospheric density of the space whose altitude at 258-272 km above Beijing was calculated by the DTM2013 model and the MSIS00 model, and the model accuracies of the two models are derived based on the GOCE measured atmospheric density in 2012.

![Figure 4. GOCE measured atmospheric density over Beijing area in 2012](image4)

![Figure 5. F30 proxy and F10.7 proxy in 2012](image5)

The GOCE measured atmospheric density over the Beijing area in 2012 is shown in figure 4. After removing redundant data with the same date, the atmospheric density in this area was measured by GOCE for 73 times in 2012. The F30 proxy and F10.7 proxy in the corresponding date in Figure 4 are shown in figure 5. The trend of the two Figures is consistently based on time series.

The average atmospheric density over the Beijing area in 2012 is \(4.35986 \times 10^{-11} \text{kg/m}^3\) at about 250 km height. The average F30 proxy and F107 proxy is 84.5507sfu and 123.0918sfu respectively in the same year. Based on the Pearson correlation criterion, the correlation coefficient between measured atmospheric density and F30 proxy is 0.8431, and the correlation coefficient between measured atmospheric density and F10.7 proxy is 0.7954. The high correlation between GOCE measured atmospheric density and solar radiation proxies are proved by the correlation coefficients, and the correlation between F30 proxy and GOCE atmospheric density is higher than that of F10.7 proxy and GOCE atmospheric density.

The atmospheric density over the Beijing area in 2012 is calculated by the DTM2013 atmospheric model and the MSIS00 atmospheric model. The error rates of the two models are derived based on the GOCE measured atmospheric density as reference. The average error rate of the DTM2013 atmospheric model is 7.73% while the average error rate of MSIS00 atmospheric model is 12.61% in this period. The DTM2013 model and the MSIS00 model uses F30 proxy and F10.7 proxy as the representation of solar radiation respectively. Based on the variance of F30 proxy and F10.7 proxy at the same period in figure 3, the correlation between the variance of F10.7 proxy and the calculation accuracy of the MSIS00
The correlation between the variance of the F30 proxy and the calculation accuracy of the DTM2013 atmospheric model is shown in figure 7. Based on the Pearson correlation criterion, the correlation coefficient between the DTM2013 atmospheric model algorithm accuracy and F30 proxy variance is 0.8431, and the correlation coefficient between the MSIS00 atmospheric model algorithm accuracy and F10.7 proxy variance is 0.8015. The accuracy of the two atmospheric models is highly correlated with the corresponding solar radiation proxy variances. The correlation of F30 proxy is slightly higher than that of F10.7 proxy. Since the variance of F30 proxy is less than that of F10.7 proxy, the solar activity factor that the accuracy of the DTM2013 atmospheric model is higher than that of the MSIS00 atmospheric model is explained.

5. Conclusion

A correlation model of GOCE measured atmospheric density and the solar radiation proxies based on the Pearson correlation criterion is derived, and the atmospheric density error of the DTM2013 model and the MSIS00 model in the same period is calculated. The Pearson correlation between the calculation accuracy of the DTM2013 model and the MSIS00 model and the variance of corresponding solar radiation proxies was established in a statistical sense, and the contribution of F30 proxy to DTM2013 calculation accuracy in solar activity term is illustrated. Since the correlation coefficient between F30 proxy and atmospheric density is higher than that of F10.7 proxy and atmospheric density, it is also proved that F30 solar radiation proxy has a higher characterization significance for solar activity in the atmospheric calculation. Due to the high correlation between the model accuracy and the variance of the corresponding solar radiation proxies, the model accuracy in high solar activity years is significantly lower than that in low solar activity years. The more intense solar activity is, the lower the accuracy of various atmospheric models will be.

Since the atmospheric density is correlated with the latitude of the studied airspace, the research area of this manuscript is focused on the airspace over the Beijing area to simplify the problem. The following research will extend the research conclusion to a wider space through the spatial autocorrelation study of atmospheric density.

Acknowledgments
Thanks to Zhenyu Sun for his help in obtaining data. Thanks to Hao Ma and Si Wu for their support in algorithm verification.

References
[1] Sentman, L. H. (1961) Free molecule flow theory and its application to the determination of aerodynamic forces. Sunnyvale: Lockheed Missiles and Space Co Inc.
[2] Cook, G. E. (1965) Satellite drag coefficients. Planetary and Space Science, 10: 929-946.
[3] March, G., Doornbos, E. N., and Visser, P.N.A.M. (2019) High-fidelity geometry models for improving the consistency of CHAMP, GRACE, GOCE and Swarm thermospheric density datasets. Advances in Space Research, 63: 213–238.

[4] Jacchia, L. G. (1959) Two atmosphere effects in the orbital acceleration of artificial satellites. Nature, 183: 526-527.

[5] Roemer, M., Framke, W. and Schuchardt, K. G. H. (1983) Solar EUV and decametric indices and thermospheric models. Advances in Space Research, 3: 75-83.

[6] Rhoden, E. A., Forbes, J. M., and Marcos, F. A. (2000) The influence of geomagnetic and solar variabilities on lower thermosphere density. Journal of Atmospheric and Solar-Terrestrial, 62: 999-1013.

[7] Tobiska, W K. (2001) Validating the solar EUV proxy E10.7. Journal of Geophysical Research Space, A12: 29969-29978.

[8] Eastes, R., Bailey, S., and Marcos, F. (2004) The correspondence between thermospheric neutral densities and broadband measurements of the total solar soft X-ray flux. Geophysical Research Letters, L19804.

[9] Bowman, B. R., Tobiska, W. K., Marcos, F. (2006) A new empirical thermospheric density model JB2006 using new solar indices. In: AIAA/AAS Astrodynamics Specialist Conference and Exhibit, Colorado, AIAA 2006-6166.

[10] Bruinsma, S. (2015) The DTM-2013 thermosphere model, Journal of Space Weather and Space Climate, 5: 1-8.

[11] NIU, J., FANG, H., and WENG, L. B. (2014) Correlations between Solar Activity and Thermospheric Density, Journal of Space Science,1: 73-80.

[12] Bruinsma, S.L., Doornbos, E., and Bowman, B.R. (2014) Validation of GOCE densities and thermosphere model evaluation, Advances in Space Research, 54: 576-585.

[13] Doornbos, E. (2012) Thermospheric density and wind determination from satellite dynamics, Springer Science & Business Media, Berlin.

[14] Emmert, J. T. (2015) Thermospheric mass density: A review, Advances in Space Research. 5: 773-824.

[15] Picone, J. M., Hedin, A. E. and Drob, D. P. (2002) NRLMSISE-00 empirical model of the atmosphere: Statistical comparisons and scientific issues, Journal of Geophysical Research,107: 1-16.

[16] Deng, L., Pei, J., Ma, J. and Lee, D. L. (2004) A rank sum test method for informative gene discovery. In: Tenth Acm Sigkdd International Conference on Knowledge Discovery & Data Mining. ACM.

[17] Weston, J., Elisseeff, A., Schölkopf, B., Tipping, M. (2003) Use of the zero norm with linear models and kernel methods. The Journal of Machine Learning Research, 3: 1439–1461.