

主成分分析を用いた HF ドップラー観測における日変動成分の抽出とその特徴解析

#小林 弥聖¹⁾, 中田 裕之¹⁾, 大矢 浩代¹⁾, 細川 敬祐²⁾, 中野 慎也³⁾

(¹ 千葉大学, (² 電気通信大学, (³ 情報・システム研究機構

Extraction and Analysis of Diurnal Variations in HF Doppler Observations Using PCA

#Misato Kobayashi¹⁾, Hiroyuki NAKATA¹⁾, Hiroyo OHYA¹⁾, Keisuke HOSOKAWA²⁾, Shinya NAKANO³⁾

(¹Chiba University, (²The University of Electro-Communications, (³The Institute of Statistical Mathematics

The ionosphere exhibits complex variations influenced by multiple factors, including diurnal and seasonal variations as well as solar activity, geomagnetic disturbances, gravity waves, and even ground-based phenomena such as earthquakes and volcanic eruptions. To capture the overall picture of such variability, simple time-series averaging or empirical criteria are insufficient, and statistical techniques are required to extract and classify variation patterns from large observational datasets.

In this study, we applied Principal Component Analysis (PCA), a statistical analysis method, to high time-resolution Doppler frequency data obtained from HF Doppler observations to extract representative diurnal variation components and analyze their characteristics. PCA analyzes the correlation structure of multidimensional data and extracts principal components that explain the maximum variance in the data, thereby enabling dimensionality reduction and characterization of variability. The extracted components are ordered as the first, second, and subsequent principal components in descending order of explained variance (contribution rate). Hence, the first principal component best represents the variability in the original data and captures the most dominant variation pattern.

The data used in this study are Doppler frequency shift observations at 5006 kHz, obtained every 10 seconds using an HF Doppler observation network operated by four institutions, primarily the University of Electro-Communications. To investigate the influence of solar activity, PCA was applied to 11 years of data (2005 – 2015).

PCA of the Doppler frequency data obtained at the Sugadaira station revealed that the absolute amplitude of the first principal component was clearly larger than that of the subsequent components. A comparison with the averaged diurnal variation on magnetically quiet days confirmed that the two variations were nearly identical, indicating that the first principal component represents the diurnal variation of the Doppler frequency. In contrast, the amplitudes of the second and higher-order components were considerably smaller, and after applying filtering based on signal intensity, these variations became nearly undetectable. These components are presumed to represent disturbances other than diurnal variation; however, their interpretation as physically meaningful variability remains difficult. Nonetheless, in certain months, the diurnal variation appeared in the second principal component. An examination of correlation coefficients revealed that such cases occurred when the correlation fell below 0.5, suggesting that during periods of frequent disturbances, the diurnal variation may also appear in the second principal component and thus requires consideration.

When comparing the first principal component of Doppler shift at the Sugadaira station over 2005 – 2015, we found that its absolute value varied in accordance with the sunspot numbers. Solar activity follows an 11-year cycle, and within the period from 2005 to 2015, 2009 corresponds to the solar minimum, while 2014 marks the solar maximum. Consistently, the extracted first principal component shows lower absolute values in 2009 and higher values in 2014. Since higher solar activity corresponds to higher maximum electron densities, this variation is likely reflected in the magnitude of the diurnal variation. We anticipate that this amplitude variation can be quantitatively derived from the electron density distribution using the IRI model, and a comparison with these model results is currently in progress.

In this presentation, we report on the validity of diurnal variation extraction by PCA and its relationship to solar activity, based on analyses comparing the extracted principal components to quiet day data and IRI model results.

電離圏の変動は多層的であり、日変動や季節変動に加えて、太陽活動、地磁気擾乱、重力波、さらには地震・火山噴火などの地上起源の影響を受けて複雑に変化する。そのような変動の全体像を把握するためには、単純な時系列平均や経験的基準では不十分であり、膨大な観測データの中から変動パターンを抽出・分類するための統計的手法が必要となる。

そこで本研究では、HF ドップラー観測により得られる高時間分解能のドップラー周波数データに対して、統計的解析手法である主成分分析 (Principal Component Analysis: PCA) を適用することで、代表的な日変動成分を抽出し、その変動の特徴についての解析を行った。

主成分分析は、多次元データから相関構造を解析し、データのばらつきを最大限に説明する主成分を抽出することで、元データの次元削減や変動の特徴の把握を可能にする手法である。抽出された主成分の分散の大きさ (寄与率) によって降順に第一主成分、第二主成分と抽出できる。つまり、第一主成分ほど元データの変動をよく表しており、代表的な変動パターンを捉えていると言える。

本研究では、電気通信大学を中心とした 4 機関で運用する HF ドップラー観測網によって 5006kHz の送信波を用いて取得された 10 秒値のドップラー周波数シフトデータである。

太陽活動の影響を調べるため、11 年分 (2005~2015 年) のデータに対して主成分分析を適用した。

菅平観測点におけるドップラー周波数データに対して主成分分析を行い、抽出された各主成分を確認したところ、第1主成分の変動の振幅の絶対値は、他の主成分と比べ明らかに大きいものであった。そこで静穏日のデータを用いて、その平均的な日変動を導出し比較したところ、両者の変化はほぼ一致することが確認された。このことから、第1主成分は、日変化を表していると考えられる。また、第2主成分以降の変動の絶対値は、かなり小さいだけでなく、受信強度でフィルタリングを行うと、第2主成分以降の変動はほとんど確認できないほどの変動となった。第2主成分以降の変動は、日変化以外の擾乱を表していると考えられるが、物理的な意味をもつ変動としての解釈は困難であると考えられる。ただし、ある月の主成分分析を行った際、第2主成分に日変化が現れる場合があることが確認された。その月の第一主成分と静穏日の変動との相関係数を調べたところ、0.5を下回る時に、そのような事例が確認されたことから、擾乱が頻繁に発生した場合には、日変化を抽出する場合には、第2主成分まで確認する必要がある。

次に、2005年から2015年の菅平観測点におけるドップラーシフトの第一主成分の変動を比較したところ、太陽黒点数の増減と共に第一主成分の絶対値も変動していることがわかった。太陽活動は11年周期で変動しており、2005年から2015年の11年においては、2009年が最も静穏であり、2014年が最も活動的な年となる。それに対して抽出した第一主成分も同様に2009年は絶対値が低く、2014年は絶対値が高くなっていた。太陽活動が高い時期においては、電子密度の最大値も大きくなることから、その変化の大きさが日変化の大きさとなって現れたと考えられる。この変化量については、現在IRIモデルを用いた電子密度分布を用いて、導出できると考えられることから、その結果との比較を進めていく予定である。

本発表では、抽出した主成分に対して静穏日やIRIモデルと比較した解析結果をもとに、主成分分析による日変動抽出の妥当性と太陽活動との関連性について報告する予定である。