Earth, Planets and Space

FULL PAPER Open Access

Variations of precipitable water vapor in sandstorm season determined from GNSS data: the case of China's Wuhai

Shihao Han[1], Xin Liu[1*], Xin Jin[1], Fangzhao Zhang[1], Maosheng Zhou[2] and Jinyun Guo[1*]

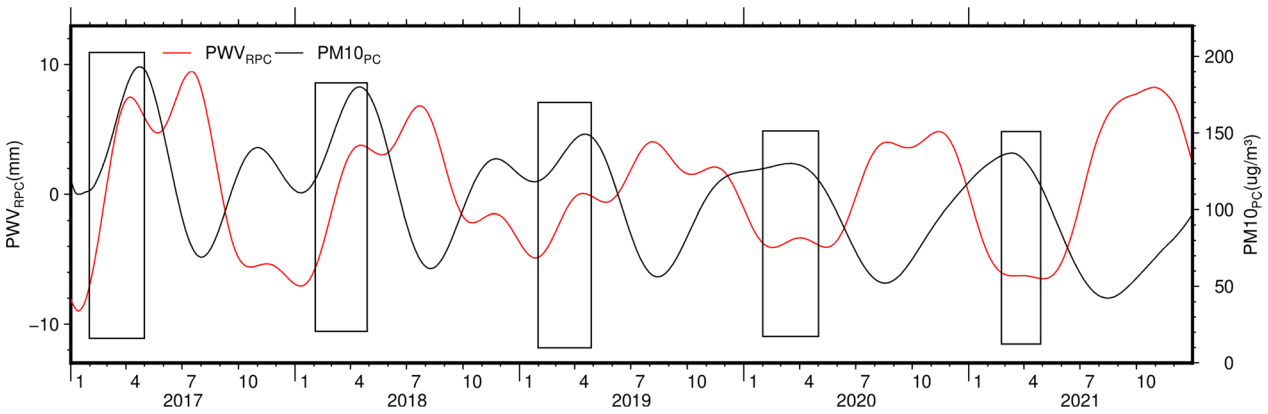
Abstract

In recent years, the Global Navigation Satellite System (GNSS) has witnessed rapid development. However, during the sandstorm season, the precipitable water vapor (PWV_{GNSS}) determined from the GNSS data produces large fluctuations due to the influence of particulate matter, which can indirectly reflect the change in particulate matter concentration. To study the variations of PWV_{GNSS} during the sandstorm season, daily data of PWV_{GNSS}, particulate matter (*PM10*), and precipitation in Wuhai from 2017 to 2021 were used in this study. The principal components of *PWV* residual (PWV_{RPC}) were obtained by using the least-squares linear fitting, singular spectrum analysis, and least-squares spectral analysis on PWV_{GNSS}. The principal components of *PM10* ($PM10_{PC}$) were obtained by using least squares linear fitting and singular spectrum analysis for *PM10*. This study performed a correlation analysis of PWV_{RPC} with $PM10_{PC}$ and precipitation data. The results showed a strong correlation between PWV_{RPC} and $PM10_{PC}$, with a correlation coefficient greater than 0.6. However, it was found that the correlation between PWV_{RPC} and precipitation was not significant. This indicates that during the sandstorm season, *PM10* affects *PWV* determined from GNSS data.

Keywords Global navigation satellite system, Sandstorm season, Precipitable water vapor, Singular spectrum analysis, Least squares spectral analysis, Wuhai

*Correspondence:
Xin Liu
xinliu1969@126.com
Jinyun Guo
jinyunguo1@126.com
Full list of author information is available at the end of the article

Graphical Abstract



Introduction

Precipitable water vapor (*PWV*) is the total amount of water vapor in a column extending from the surface of the Earth to the top of the atmosphere, which is a highly active component of the atmosphere and varies significantly with time and space (Gurbuz 2021; Zhang et al. 2022a). Excessive changes in water vapor can lead to extreme weather, so *PWV* can play an important role in many scientific research fields such as climate forecasting (Chang et al. 2015). As the Global Navigation Satellite System (GNSS) continues to advance, the utilization of GNSS for the estimation of *PWV* has reached a higher level of maturity. In recent years, China has made remarkable progress in the establishment of numerous Continuous Operational Reference System (CORS) stations and the deployment of the Crustal Movement Observation Network of China (CMONOC) through the utilization of GNSS technology. (Yu et al. 2019; Zhao et al. 2015). Compared to conventional methods of water vapor detection, the utilization of GNSS technology for *PWV* estimation offers distinct advantages, including continuous operation, all-weather capability, high accuracy, and enhanced resolution (Wang et al. 2016; Zhai and Eskridge 1996; Tao et al. 2021; Liu et al. 2000; Bevis et al. 1992; Yang 2022; Gurbuz and Bayik 2023). Studying the variations of precipitable water vapor (PWV_{GNSS}) determined from the GNSS data can indirectly reflect climate change and extreme weather events in the region. During the inversion process, PWV_{GNSS} is easily affected by extreme weather events such as haze and sandstorms, resulting in fluctuations in the calculation results (Guo et al. 2020).

As a common form of extreme weather, sandstorms pose a significant threat as a highly hazardous meteorological disaster and act as an important dust transport process that accelerates land desertification. Over the past 500 years, East Asia has experienced recurrent occurrences of sandstorms, with northwest China being particularly vulnerable to their devastating impact. These sandstorm disasters have inflicted significant harm on various aspects, including economic development, human well-being, public health, and the environment (Huo et al. 2019; Zhang et al. 2021). According to their diameters, sand particles are classified as a particulate matter (Watanabe et al. 2016), which refer to the general term for solid and liquid particulate matter that floats in the air. Respirable suspended particles, commonly known as *PM10*, are particles with a diameter of 10 μm or less. Furthermore, there are fine particles called *PM2.5*, which have a diameter of 2.5 μm or less (Lau and He 2017). Both *PM10* and *PM2.5* particles contain numerous harmful substances, posing substantial risks to human health

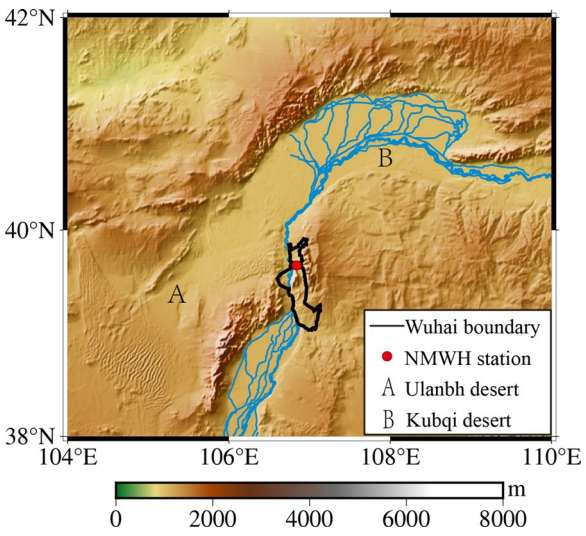


Fig. 1 NMWH station and geographical location of Wuhai

and leading to various societal concerns. (Yao et al. 2015; Kiser et al. 2020; Paun et al. 2015; David and Gao 2016). Based on the long-term occurrence of sandstorms in Wuhai and in combination with the climate bulletins issued by the China Meteorological Administration and the government of Inner Mongolia Autonomous Region, this study concludes that the period from February to April is the sandstorm season in Wuhai. During the sandstorm season, the PWV_{GNSS} is influenced not only by precipitation but also by the presence of particulate matter (Guo et al. 2021a, 2021b; Wen et al. 2020). By studying the correlation between PWV_{GNSS} and particulate matter during sandstorms, this paper aims to identify the relationship between PWV_{GNSS} and particulate matter. By identifying the occurrence time and trajectory of sandstorms within the area, proactive alerts can be issued to mitigate the detrimental impact caused by these sandstorms.

Numerous studies have employed SSA in the analysis of data sets, including *PWV* and *PM10* (Zhou et al. 2022). As a digital signal processing technique, SSA has gained significant prominence in the fields of climatology and measurement due to its wide range of applications. It is capable of effectively performing operations on time series data, including trend identification, period determination, data reconstruction, and data interpolation (Vautard et al. 1992; Wyatt et al. 2012; Chen et al. 2013; Zabalza et al. 2014; Guo et al. 2018; Shen et al. 2018).

To study the variations of *PWV* determined from GNSS data during the sandstorm season, daily average data of PWV_{GNSS}, precipitation, and particulate matter from January 1, 2017 to December 31, 2021 in Wuhai were selected for this study. The focus of this study is on *PM10* due to its strong correlation with *PM2.5*, and also because the proportion of coarse particles in sandstorms during spring is significantly higher than that of fine particles (Alghamdi et al. 2015). The data processing procedure was as follows: firstly, this study aims to obtain principal components of *PWV* residual (PWV_{RPC}) and principal components of *PM10* ($PM10_{PC}$) by using least-squares fitting, singular spectrum analysis, and least-squares spectrum analysis; Secondly, this study aims to obtain the correlation between PWV_{RPC} and $PM10_{PC}$ and between PWV_{RPC} and precipitation.

Data and methods

Wuhai is located within the range of 106°40'E-107°09'E and 39°15'N-39°32'N, belonging to the northwest region of China (Fig. 1), close to the Loess Plateau and the border between China and Mongolia, where sandstorms frequently occur in spring. The Wuhai station (NMWH) established by CMONOC is shown in Fig. 1. On the one hand, the vast desert area in southern Mongolia is one of the key sources of sandstorms, from which most of the particulate matter of sandstorms in northern China originates (Han et al. 2021). On the other hand, the unique geographical location and climatic environment make it prone to sandstorms in spring every year.

Data

Precipitable water vapor

The PWV_{GNSS} data is provided by CMONOC, which belongs to GNSS data product service platform of China Earthquake Administration (http://www.cgps.ac.cn). The second sampling observation of GNSS reference stations in the CMONOC system primarily aims to monitor the total atmospheric water vapor content and stratified water vapor content, thereby offering a comprehensive water vapor product (Zhu et al. 2020).

The data processing software is GAMIT. The sampling rate of GNSS observation was 30 s, Rinex format, cut-off angle of 10°, GMF mapping functions. The zenith tropospheric delay (ZTD) correction is estimated hourly and the GAMIT default horizontal gradient are used (Liu et al. 2022).

After the parameters selection is completed, the hourly ZTD can calculated by GAMIT. The ZTD consists of zenith hydrostatic delay (ZHD) and zenith wet delay (ZWD), and ZHD can be accurately calculated with the equation by Saastamoinen (1972):

$$ZHD = \frac{(2.2768 \pm 0.0024)P_S}{1 - 0.00266 \times \cos 2\theta + 0.00028 \times H} \quad (1)$$

where P_s is atmospheric pressure (unit: hPa). θ is the latitude of the GNSS station (unit: °). H is the height of the GNSS station (unit: km).

After obtaining ZHD, the ZWD can be calculated according to Eq. (2):

$$ZWD = ZTD - ZHD \quad (2)$$

The PWV can be calculated by the Eq. (3):

$$PWV = \Pi \times ZWD \quad (3)$$

where $\prod$ is the conversion, which can be calculated by equation (Bevis et al. 1994):

$$\Pi = \frac{10^6}{\left(k_3 \times T_m^{-1} + k_2\prime\right) \times R_v} \quad (4)$$

where k_2' and k_3 are the atmospheric refractivity constants ($k_2\prime = 22.1 \pm 2.2K/hPa, k_3 = 3.375 \times 10^5 \pm 0.012 \times 10^5 K^2/hPa$), and Rv is the gas constant of water vapor ($R_v = 461.495\ J/(kg \cdot K)$). Tm is the weighted mean temperature (unit: K), which can be calculated using radiosonde profiles, or using numerical weather models (Bevis

et al. 1994; Yao et al. 2012), most published studies to date estimate *Tm* by equation (Bevis et al. 1992):

$$T_m = 70.2 + 0.72 \times T_s \tag{5}$$

where *Ts* is the surface temperature, which can be obtained from ground-based weather stations co-located with GNSS stations.

In this study, PWV_{GNSS} data were provided by CMONOC for every two hours from January 1, 2017 to December 31, 2021 and were averaged over 12 PWV_{GNSS} data per day before being used.

Particulate Matter 10

The real-time data platform of the China National Environmental Monitoring Center (CNEMC) provides hourly *PM10* and *PM2.5* data (http://www.cnemc.cn/sssj). The process of calculating the correlation between *PM10* and *PM2.5* is as follows:

$$\begin{aligned}\rho(PM_{10}, PM_{2.5}) &= \frac{\mathrm{cov}(PM_{10}, PM_{2.5})}{\sigma_{PM_{10}}\sigma_{PM_{2.5}}} \\ &= \frac{E[(PM_{10} - E(PM_{10}))(PM_{2.5} - E(PM_{2.5}))]}{\sigma_{PM_{10}}\sigma_{PM_{2.5}}} \\ &= \frac{E(PM_{10}PM_{2.5}) - E(PM_{10})E(PM_{2.5})}{\sqrt{E(PM_{10}^2) - E^2(PM_{10})}\sqrt{E(PM_{2.5}^2) - E^2(PM_{2.5})}}\end{aligned} \tag{6}$$

where ρ is the correlation coefficient, cov is the covariance, and σ is the standard deviation.

The correlation between *PM10* and *PM2.5* is 0.7982, so the *PM10* data from January 1, 2017 to December 31, 2021 was selected for this study. The *PM10* data were averaged over 24 data per day before being used.

Precipitation

ERA5 is the fifth generation of the European Centre Weather Forecasting (ECMWF) reanalysis dataset, which mainly used in global digital weather forecasting, air quality analysis, atmospheric composition monitoring, climate monitoring, etc. (Lavers et al. 2022). For this study, ERA5 was utilized to accquire hourly precipitation data from January 1, 2017 to December 31, 2021. Daily precipitation data were derived by summing up the 24-hourly precipitation values. Since ERA5 provides gridded data with a resolution of 0.25°×0.25°, it should be interpolated to the Wuhai station based on longitude, latitude, and precipitation values.

Methods

The missing rate of PWV_{GNSS} data is 6.16% and the missing rate of *PM10* data is 5.94%, so the missing of PWV_{GNSS} and *PM10* data are interpolated by using SSA before being used. Firstly, SSA is performed on the PWV_{GNSS} and *PM10* after removing their linear trends to obtain the main components of both, and the main period of both is determined using fast Fourier transform. Secondly, a fitted *PWV* (PWV_{LSF}) data is obtained through the application of least square spectral analysis to the PWV_{GNSS}. Then, the *PWV* residual (PWV_R) is obtained by subtracting PWV_{LSF} from PWV_{GNSS}, and the PWV_{RPC} is obtained through the application of SSA. Similarly, the *PM10* residual ($PM10_R$) is obtained by subtracting $PM10_{PC}$ from *PM10*. The correlation coefficients were obtained by correlating the PWV_{RPC} with the $PM10_{PC}$ and precipitation.

Singular spectrum analysis

Interpolation In this study, the SSA was used to pre-processing PWV_{GNSS} and *PM10* data, and iterative interpolation was used for missing data (Guo et al. 2018; Jin et al. 2021). The specific steps of interpolation are as follows:

(1) The missing data in the original time series were filled with 0 and marked accordingly, followed by centralizing the completed data.

(2) The singular spectrum decomposition is performed on the new time series data to determine the first i-order reconstruction component (*RC*) as the main component of the original series, and the value at the missing position in the original series is replaced by the value in the first *RC* (noted as *R1*), and this process is repeated until the data residual RMS of the two interpolated data before and after is less than a certain value.

(3) Add the value in the second *RC* (noted as *R2*) to reconstruct the missing part of the data, i.e., the missing data is obtained by linear superposition of *R1* and *R2*, and repeat the above step (2) until *RC* is linearly superimposed to the missing data.

Decomposition In this study, SSA is performed on PWV_{GNSS} and *PM10* to determine the trend and period data of both. The decomposition principle is as follows:

(1) Construct a time-delay matrix with the amount of data N and choose an appropriate window size M, in which M is generally greater than 1 and less than $N/2$. For time series where the period is already known, the window length M is generally taken as a common multiple of the period for better separation of the periodic part. Let $K = N - M + 1$, then the constructed time-delay matrix X can be expressed as:

$$X = [X_1, X_2, \cdots, X_k] = \left(x_{i,j=1}^{L,k}\right) = \begin{bmatrix} x_1 & x_2 & \cdots & x_{N-M+1} \\ x_2 & x_3 & \cdots & x_{N-M+2} \\ \vdots & \vdots & \ddots & \vdots \\ x_M & x_{M+1} & \cdots & x_N \end{bmatrix} \tag{7}$$

where $X_i = (x_i, \cdots, x_{i+M-1})^T, (1 \le i \le K)$. X is a time-delay matrix of order $M \times K$. Each row and column of

it is a subsequence of the initial time series, and it is a Hankel matrix.

(2) Singular values decomposition (SVD), there was a matrix $S = XX^T$. $\lambda_1, \cdots, \lambda_M$ the eigenvalues of S arranged in descending order of magnitude as $\lambda_1 \geq \cdots \geq \lambda_M \geq 0$, and the corresponding eigenvectors are $U_1, \cdots, U_M$, respectively. The matrices X_i have rank 1; such matrices are sometimes called elementary matrices. The SVD of the time-delay matrix X can be expressed as:

$$X = X_1 + X_2 + \cdots + X_d \tag{8}$$

where $d = rank(X)$, $X_k = \sqrt{\lambda_k} U_k V_k^T$, $V_k = X^T U_k / \sqrt{\lambda_k}$ $(k = 1, \cdots, d)$. $\sqrt{\lambda_k}(k = 1, \cdots, d)$ is a singular value of X_k. The new matrix $X^{(r)}(r < d)$ can be constructed by selecting appropriate eigenvalues and corresponding eigenvectors. The collection will be called i-th eigentriple (abbreviated as *ET*) of the SVD.

(3) Diagonal averaging. The purpose of diagonal averaging is to reconvert the primitive matrix X_k obtained from the decomposition in the previous step into a new time series of length N, called the *RC*, and the sum of all *RCs* is equal to the original series. Define the matrix $Z = X_k, z_1, z_2, \cdots, z_N$, as the time series obtained by diagonal averaging of Z. Let $M^* = min(M, K), K^* = \max(M, K)$, if M < K, then $z_{ij}^* = z_{ij}$, otherwise $z_{ij}^* = z_{ji}$, the equation for diagonal averaging can be expressed as:

$$z_i = \begin{cases} \frac{1}{i} \sum_{m=1}^{k} z_{m,i=m+1}^*, 1 \leq i \leq M^* \\ \frac{1}{M^*} \sum_{m=1}^{L^*} z_{m,i=m+1}^*, M^* \leq i \leq K^* \\ \frac{1}{N-i+1} \sum_{m=i-K^*+1}^{N-K^{*+1}} z_{m,i=m+1}^*, K^* < i \leq N \end{cases} \tag{9}$$

Least squares spectrum analysis

The PWV_{LSF} data were obtained using least squares spectrum analysis, and their variation in rate, amplitude, and phase on the time scale was (Shi et al. 2022b):

$$y = a_0 + a_1 t + \sum_{1}^{n} A_i \sin(\omega_i (t - t_0) + \varphi_i) \tag{10}$$

where A is the amplitude, ω is the periodic signal, a_0 is the deviation of the time series, a_1 is the linear trend, φ_i is the initial phase.

Analysis

Analysis of precipitable water vapor data

PWVGNSS time series data analysis

To obtain the linear trend of PWV_{GNSS} (PWV_{GNSSL}) from 2017 to 2021, a least-squares linear fit is initially conducted:

$$PWV_{GNSSL} = 0.37Y + 10.42 \tag{11}$$

where Y represents the year and PWV_{GNSS} shows an increasing trend during 2017–2021 with a growth rate of 0.37 mm/year. It can indicate that the climate in Wuhai area has gradually become wetter in recent years and the water vapor content in the air has increased (Zhu et al. 2022).

It is evident that PWV_{GNSS} exhibits a notable periodicity, reaching its peak in July and August each year, with a maximum value of 48.86 mm. Conversely, during the winter and spring seasons, the values typically remain below 10 mm (Fig. 2). The increase in temperature can enhance the atmosphere's ability to store water vapor. Under conditions where there is sufficient water vapor source, the higher the content of atmospheric water vapor, the larger the PWV_{GNSS} value (Shi et al. 2020). Wuhai experiences a characteristic temperate continental climate, characterized by distinct features such as hot and

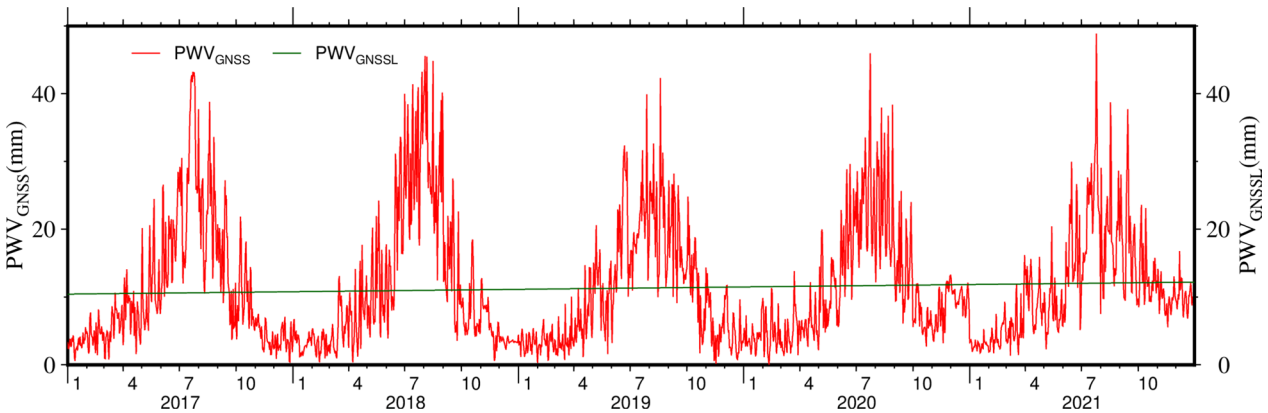


Fig. 2 PWV_{GNSS} and PWV_{GNSSL} data for 2017 to 2021

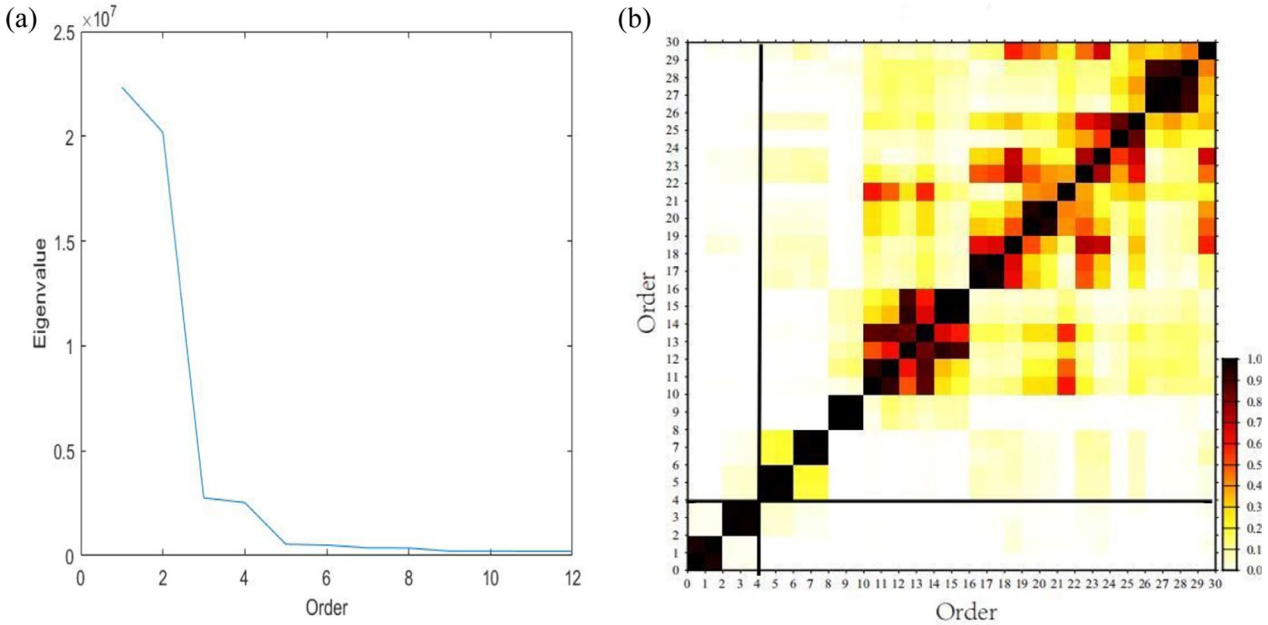


Fig. 3 Order-eigenvalue and w-correlation of reconstruction. **a** Eigenvalues of different orders. **b** w-correlation for the first 30 orders of reconstruction

rainy conditions in summer and autumn, and relatively lower rainfall with drier weather in winter and spring. Consequently, the PWV_{GNSS} values in Wuhai tend to be higher during summer and autumn compared to winter and spring. Additionally, the occurrence of sandstorms is more frequent during the spring season.

PWVGNSS time series data analysis

The decomposition of the PWV_{GNSS} with the linear trend removed using SSA avoids the effect of the linear trend decomposition on each order. PWV_{GNSS} subtracts PWV_{GNSSL} to obtain a set of data without linear trend, and SSA analysis is performed on PWV_{GNSS} data without linear trend. Based on several practical experiences and periods, the window length M was chosen as 730 and the principal components of PWV_{GNSS} were obtained. The eigenvalues of different orders were arranged using the descending order method (Fig. 3).

The eigenvalues obtained after the SSA decomposition of PWV_{GNSS}, the eigenvalues are decreasing in value as the order increases. When the order exceeds 4 times, the eigenvalues tend to level off (Fig. 3a). The larger the value of the eigenvalue, the greater the energy in the signal represented by its corresponding eigenvector, and the smaller the energy vice versa. The correlation between the orders after SSA decomposition is found using the w-correlation method (Hassani 2007), and the decomposition is poor when the order is greater than 4 (Fig. 3b). Therefore, the first four orders (*RC1*, *RC2*, *RC3*, *RC4*) reconstruct the PWV_{GNSS} time series. $RC1+RC2$ form a new periodic term and $RC3+RC4$ form a new periodic term.

From January 1, 2017 to December 31, 2021, the maximum value of PWV_{GNSS} is 48.86 mm. Although there is a clear trend of fluctuation in PWV_{GNSS} from sandstorm season each year, PWV_{GNSS} values are much larger in the non-sandstorm season than in the sandstorm season. After SSA decomposition of the PWV_{GNSS} time series with the linear trend removed, the variance contribution of the first 4th-order reconstructed components reached 69.61%, indicating that the first 4th-order reconstructed time series can characterize this PWV_{GNSS} time series. Analyzing $RC1+RC2$ and $RC3+RC4$ using the fast Fourier transform yields their respective periods (Fig. 4).

The periodic term consisting of *RC1* and *RC2* with values ranging from − 11.60 mm to 12.4283 mm. The peak of the new periodic term occurs during August each year and tends to decrease year by year, and the valley occurs in January each year (Fig. 4a). The first period is composed of *RC1* and *RC2*, which are similar to two basic functions of the first period. The second period is composed of *RC3* and *RC4*, which are similar to two basic functions of the second period. *RC1* and *RC2* have the same period, as do *RC3* and *RC4*, therefore it is necessary to study $RC1+RC2$ and $RC3+RC4$. The main period of the period term consisting of $RC1+RC2$ calculated by the fast Fourier transform, which has a principal period of 0.9355 year

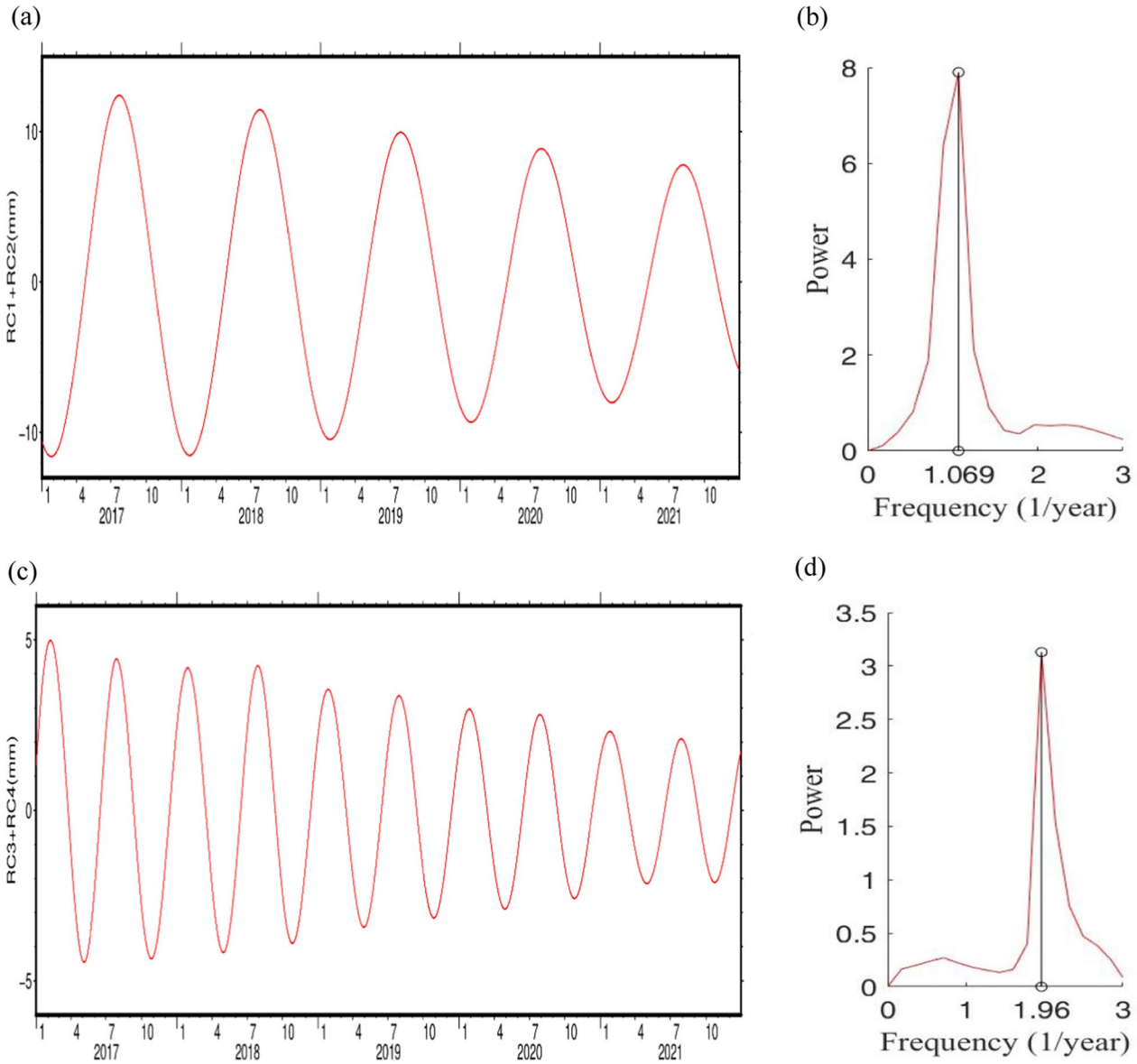


Fig. 4 Reconstruction component and Frequency-Power of PWV_{GNSS}. **a** Reconstruction of component *RC1* + *RC2*. **b** Frequency-Power diagrams of (**a**). **c** Reconstruction of component *RC3* + *RC4*. **d** Frequency-Power diagrams of (**c**)

(Fig. 4b). The periodic term composed of *RC3* and *RC4* with values ranging from − 4.46 mm to 4.99 mm, with peaks occurring in February and August each year and a decreasing trend year by year, and valleys occurring in May and November. The peak of the cycle consisting of *RC3* + *RC4* occurs approximately in the summer and winter of each year (Fig. 4c), a phenomenon due to the influence of PWV_{GNSS} by precipitation and particulate matter, respectively (Shi et al. 2022a, b; Wang et al. 2016). The main period of the period term consisting of *RC3* + *RC4* calculated by the fast Fourier transform, which has a principal period of 0.5102 year (Fig. 4d).

Least squares spectral analysis of PWVGNSS and determination of PWV residuals

In this study, the fitted data PWV_{LSF} was derived through least squares spectral analysis for PWV_{GNSS} data from January 1, 2017 to December 31, 2021. The fitting process involved incorporating two periods, namely 0.9355 years and 0.5102 years (Fig. 5a). The PWV_R is obtained by subtracting PWV_{LSF} from PWV_{GNSS} (Fig. 5b).

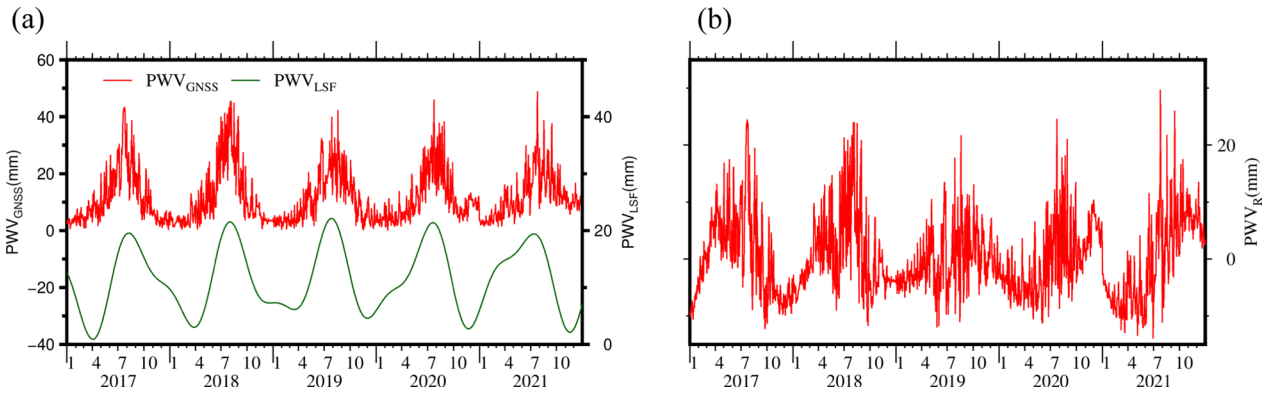


Fig. 5 PWV_{LSF} and PWV_R data. **a** 2017 to 2021 PWV_{LSF} time series data. **b** 2017 to 2021 PWV_R time series data

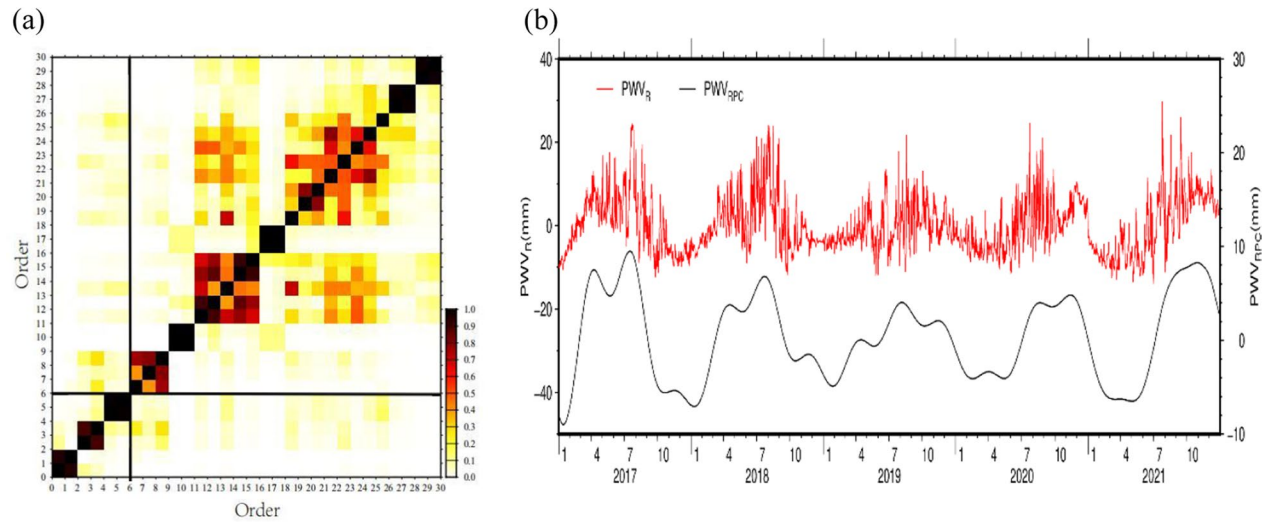


Fig. 6 w-correlation of reconstruction and PWV_{RPC} data. **a** W-correlation of the first 30th order reconstruction components of PWV_R. **b** 2017 to 2021 PWV_R and PWV_{RPC} time series data

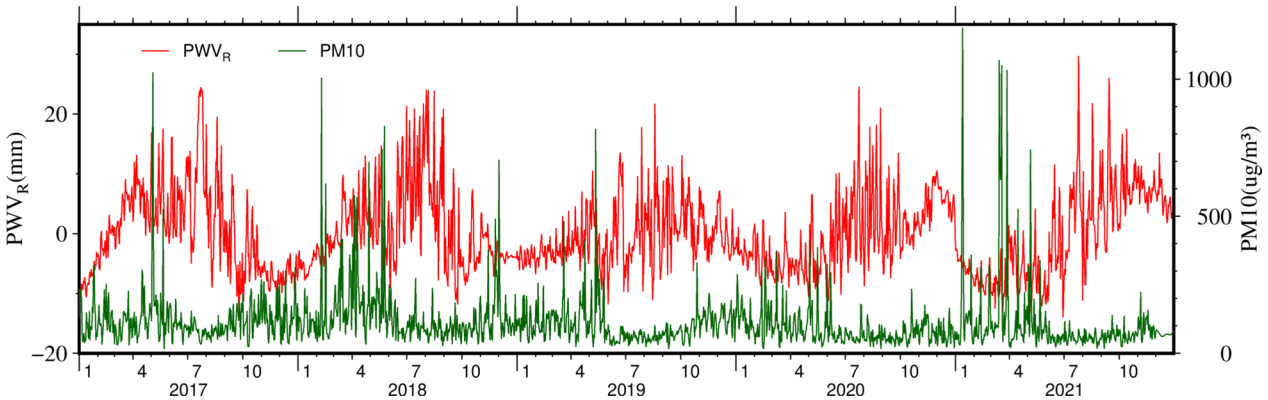


Fig. 7 2017 to 2021 PWV_R and *PM10* time series data

Determination of the principal components of the PWV residuals

SSA was applied to the PWV_R to extract its principal components, with a window length of 880 chosen based on practical usage experience and the period. The correlation between the decomposed components was evaluated using the w-correlation method (Fig. 6a).

After decomposing the PWV_R by SSA, the first 6 orders are regarded as PWV_{RPC}. PWV_{RPC} has three annual peaks during 2017 to 2021 and each of these three peaks is caused by different drivers. The first peak occurs in April each year and is caused by particulate matter, such as *PM10*, brought in by weather such as sandstorms. The second peak occurs in August each year and has the longest duration and highest values compared to the other two peaks. The third peak occurs in December every year, which is because there is a lot of frost and snow in the winter in Wuhai (Fig. 6b).

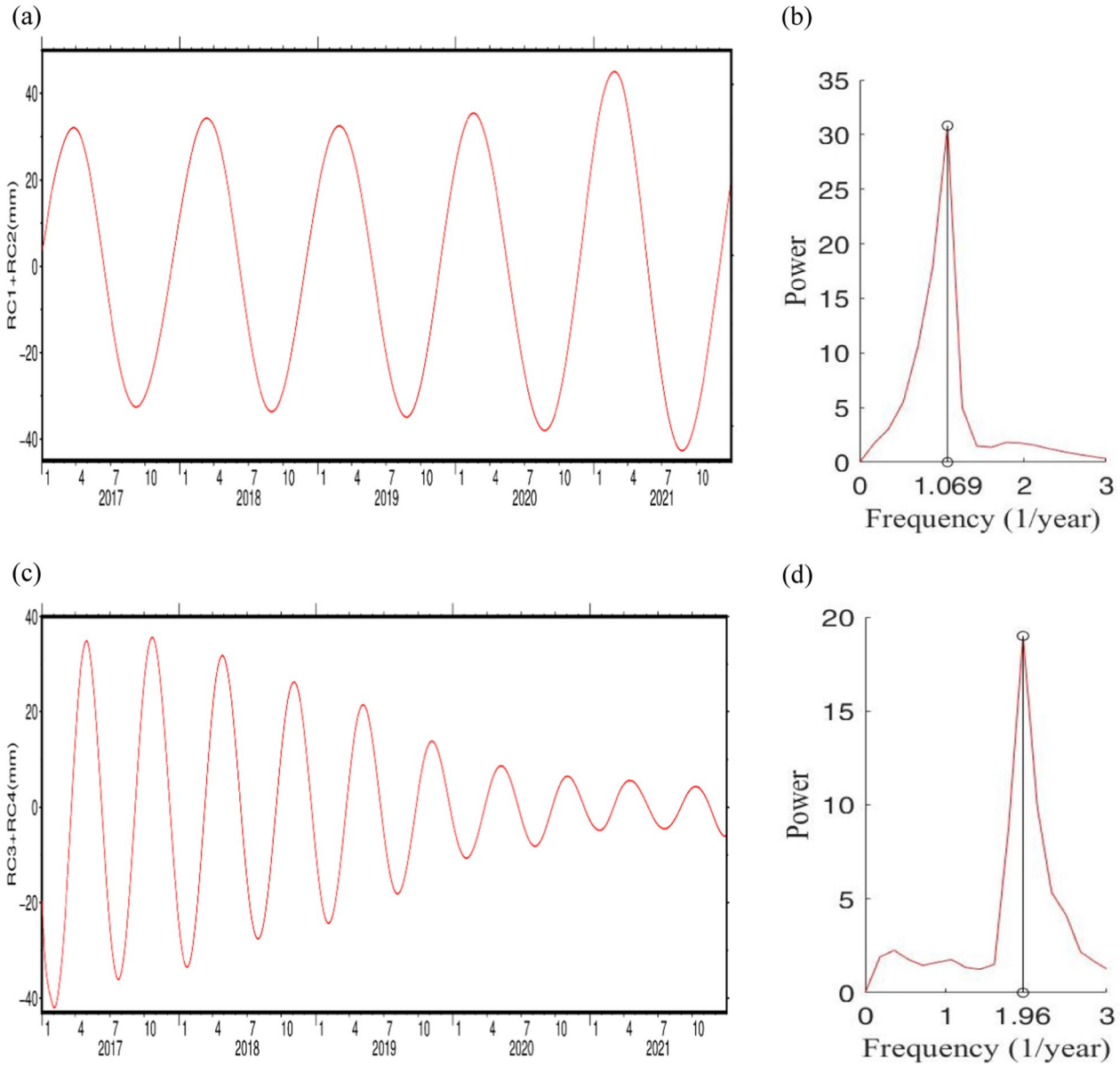


Fig. 8 Reconstruction component and Frequency-Power of *PM10*. **a** Reconstruction of component *RC1* + *RC2*. **b** Frequency-Power diagrams of (**a**). **c** Reconstruction of component *RC3* + *RC4*. **d** Frequency-Power diagrams of (**c**)

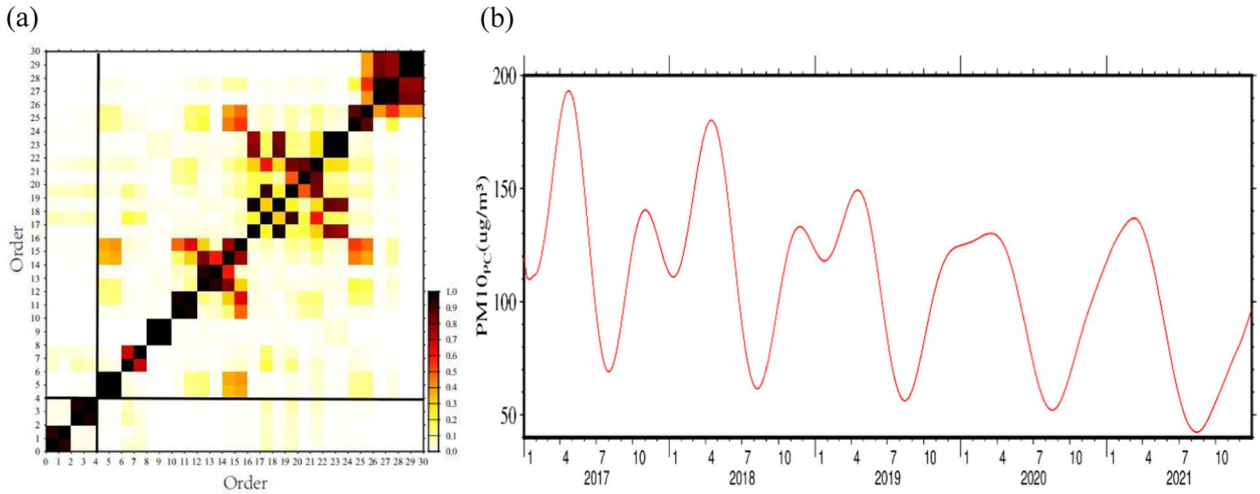


Fig. 9 w-correlation of reconstruction and $PM10_{PC}$ data. **a** w-correlation of the reconstructed components of the first 30 orders of *PM10*. **b** 2017 to 2021 $PM10_{PC}$ time series data

Analysis of *PM10* change pattern

PM10 time series data variation

Since the correlation coefficient between *PM10* and *PM2.5* is 0.7982, the study focused exclusively on *PM10*. From the data, it can be concluded that the average value of *PM10* in the sandstorm season 2017 to 2021 is 142.01 μg/m^3; the average value of 89.64 μg/m^3 in the non-sandstorm season 2017 to 2021. The average values of *PM10* for each year from 2017 to 2021 were 119.18 μg/m^3, 137.29 μg/m^3, 101.77 μg/m^3, 87.69 μg/m^3, and 100.96 μg/m^3, respectively. The highest annual average *PM10* value occurred in 2018, and the annual average *PM10* value decreased after 2018. The highest single-day value of *PM10* was 1187 μg/m^3, which occurred in the spring of 2021 (Fig. 7). The slightly larger annual mean *PM10* in 2021 than in 2020 is due to the transport of particulate matter from the Gobi Desert in southern Mongolia to China in the spring of 2021 due to the influence of the cold cyclone in Mongolia and the cold air behind it. And it caused most of northern China to experience the most extensive, intense, and longest-impacting dust and sand weather in China in the past decade (Zhang et al. 2022b).

Determination of PM10 principal components

To obtain the linear trend of *PM10* ($PM10_L$) from 2017 to 2021, a least-squares linear fit is initially conducted:

$$PM10_L = -10.48Y + 135.54 \tag{12}$$

where Y represents each year, and *PM10* shows a decreasing trend during 2017 to 2021 at a rate of 10.48 μg·m^{-3}/year. It shows that the air quality has been improved in recent years under the implementation of environmental protection policies in Wuhai area (Shi et al. 2007).

In order to avoid the impact of the linear trend on each order, it is necessary to subtract the linear trend from *PM10* before performing the SSA. And the window length M was chosen to be 730 according to the usage experience and period. The correlation between the orders after SSA decomposition is found using the w-correlation method.

The first 4 orders were analyzed using the fast Fourier transform, and the periods of $RC1+RC2$ and $RC3+RC4$ were obtained as 0.9355 and 0.5102 year, respectively (Fig. 8). Thus, the linear trend fitted by Eq. (12) plus the first 4 orders after SSA decomposition form the $PM10_{PC}$ (Fig. 9a). *PM10* demonstrates approximately two peaks and one trough annually throughout the period from 2017 to 2021. The first peak consistently emerges in April, primarily attributable to heightened occurrences of sandstorm weather. The trough occurs in August each year, which is caused by the negative correlation between *PM10* and precipitation (Yang et al. 2021). Rainfall plays a crucial role in cleansing the atmosphere by effectively washing away pollutants and particulate matter, leading to a notable reduction in concentrations of particulate matter. The second peak occurs in December each year, when *PM10* values rise mainly due to haze weather. There is a gap between the second peak and the first peak of each year, as shown by the fact that the value of the second peak is smaller than the value of the first peak. This is because the increase in *PM10* concentration caused by haze is weaker than the increase in *PM10* concentration of sandstorms, but this difference will gradually decrease. This is due to the decreasing frequency and intensity of

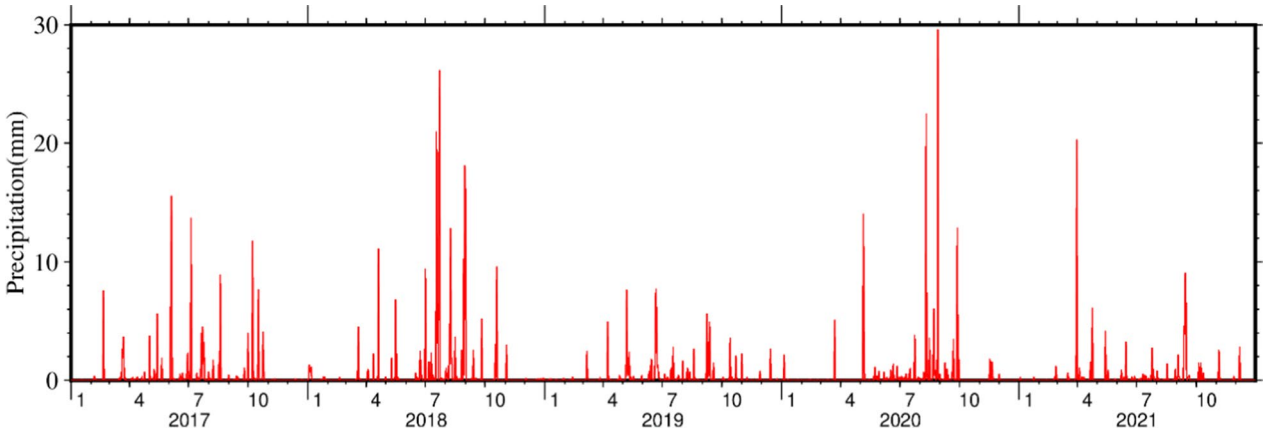


Fig. 10 Precipitation time series data for 2017 to 2021

Table 1 Correlation coefficients between PWV_{RPC} with $PM10_{PC}$ and precipitation in the sandstorm and non-sandstorm seasons

Sandstorm season			Non-sandstorm season		
Time	**Correlation coefficient**		**Time**	**Correlation coefficient**	
February—April 2017	$PM10_{PC}$	0.9497	May 2017—January 2018	$PM10_{PC}$	−0.0271
	Precipitation	−0.0042		Precipitation	0.1266
February—April 2018	$PM10_{PC}$	0.9994	May 2018—January 2019	$PM10_{PC}$	−0.4418
	Precipitation	0.1581		Precipitation	0.2306
February—April 2019	$PM10_{PC}$	0.9959	May 2019—February 2020	$PM10_{PC}$	−0.8101
	Precipitation	0.0979		Precipitation	0.0336
February–March 2020	$PM10_{PC}$	0.6017	April 2020—February 2021	$PM10_{PC}$	−0.7312
	Precipitation	0.1846		Precipitation	0.0879
March–April 2021	$PM10_{PC}$	0.9290	May-December 2021	$PM10_{PC}$	−0.5521
	Precipitation	0.0028		Precipitation	0.0703

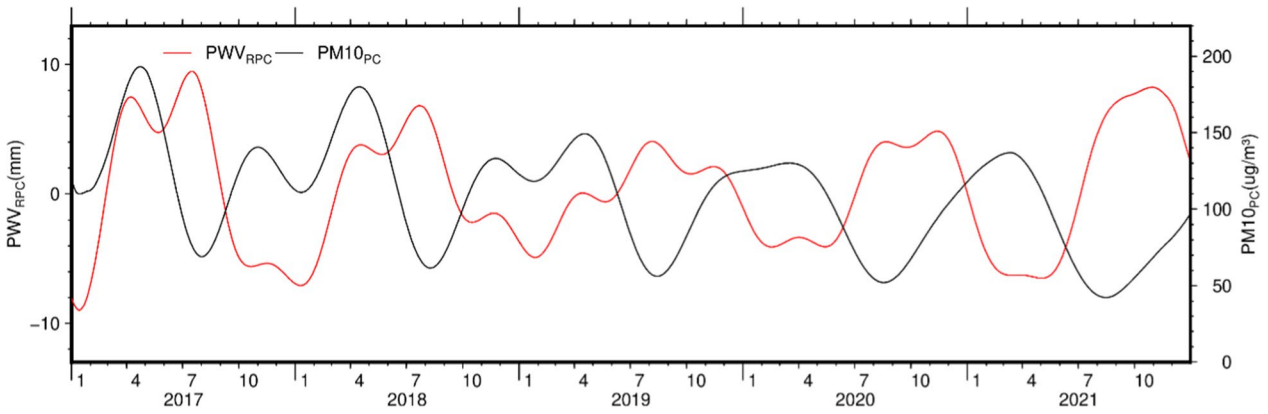


Fig. 11 2017 to 2021 PWV_{RPC} and $PM10_{PC}$ time series data

sandstorms during the sandstorm season as the quality of our environment continues to improve. The second peak in 2020 and the first peak in 2021 form a new peak (Fig. 9b). This is because the spring of 2021, when the Wuhai suffered from the largest and most intense as well as the longest duration sandstorm weather in the last decade (Zhang et al. 2022a, b), so it caused the two periods to merge into one.

Analysis of the variation pattern of precipitation data

Precipitation is also an important cause of PWV_{GNSS} variation, and PWV_{GNSS} shows a significant correlation with precipitation over short time periods (Shi et al. 2022a). From the data, it can be concluded that the average value of precipitation in the non-sandstorm season 2017–2021 is 123.61 mm; the average value of precipitation in the sandstorm season 2017 to 2021 is 18.58 mm. The number of days with daily precipitation exceeding 1 mm in each year of the sandstorm season from 2017 to 2021 were: 6 days, 7 days, 2 days, 1 day and 8 days, respectively. Precipitation in Wuhai is mainly concentrated in May to September, with the highest daily precipitation occurring in July or August each year, and less from October to April of the following year (Fig. 10).

Correlation analysis of factors in the sandstorm season

In order to investigate the impact of *PM10* and precipitation on PWV_{GNSS} during the sandstorm season, the correlation between PWV_{RPC} and $PM10_{PC}$, precipitation was calculated separately. Subsequently, the influence of *PM10* and precipitation on PWV_{GNSS} was analyzed. From February to April is the sandstorm season in Wuhai, but there are some years when the season occurs late or ends early. Therefore, correlation analysis should be performed based on the actual time of the sandstorm season.

The correlation coefficients between PWV_{RPC} and $PM10_{PC}$ are greater than 0.6 in the sandstorm season. The correlation coefficient between PWV_{RPC} and precipitation is less than 0.2 in the sandstorm season. It indicates that *PM10* influences the variation of PWV_{GNSS} during the sandstorm season. Since there is less precipitation during the sandstorm season, the precipitation produces little change in PWV_{GNSS}. Due to the long time span of the non-sandstorm season, the factors influencing of PWV_{GNSS} are more complex than those of the sandstorm season (Table 1), so the correlation coefficients of PWV_{RPC} with $PM10_{PC}$ and precipitation are not significant (Fig. 11).

Results and discussion

The *PWV* obtained based on GNSS and *PM10* were analyzed. The results showed that PWV_{GNSS} shows an increasing trend and two main periods of 0.9355 and 0.5102 year from 2017 to 2021. The PWV_{GNSS} is rising at a rate of 0.37 mm/year, which indicates that the climate in Wuhai has become progressively wetter in recent years. *PM10* shows a decreasing trend from 2017 to 2021 with a rate of -10.48 $\mu g \cdot m^{-3}$/year, and there are two main periods of *PM10* with 0.9355 and 0.5102 years respectively.

This study analyzed the relationship between the PWV_{RPC} with the $PM10_{PC}$ and precipitation during the sandstorm season and non-sandstorm season. The results show that the correlation between PWV_{RPC} and $PM10_{PC}$ during the sandstorm season is significant, with correlation coefficients greater than 0.6. Thus, *PM10* is a factor in the influence of sandstorm season on the *PWV* determining from GNSS data.

Through their research, Wang and Guo have discovered that the changes of *PWV* determined from GNSS data influenced by particulate matter. Shi have discovered that a clear correspondence between PWV_{GNSS} and precipitation in the short term (Guo et al. 2021a, 2020; Wang et al. 2016; Shi et al. 2022a, b). This paper indicates that during the sandstorm season, the PWV_{RPC} (PWV_R represent the variations in PWV_{GNSS} influenced by other factors) has a significant correlation with *PM10*, while the correlation with precipitation is not significant. The lack of significant correlation with precipitation is primarily due to the paper focus on the period from February to April, during which there is less precipitation and frequent occurrences of sandstorm.

Compared to the expensive cost of building a meteorological station, the cost of establishing GNSS stations is relatively low. In the future, multiple GNSS stations can be set up along the sandstorm transport route in northwest China based on the clear correspondence between PWV_{GNSS} and particulate matter during the sandstorm season. GNSS technology can be utilized to detect the occurrence of sandstorms. This can further determine the actual path and arrival time of the sandstorm and issue warnings to mitigate the various hazards caused by sandstorms.

Acknowledgements

We thank the product service platform of China Earthquake Administration for providing PWV_{GNSS} data. We would like to thank China National Environmental Monitoring Center (CNEMC) for providing *PM10* data, and fifth generation of the European Centre Weather Forecasting (ECMWF) for providing precipitation data. This study is partially supported by the National Natural Science Foundation of China (grant No. 42274006), and the Autonomous and Controllable Special Project for Surveying and Mapping of China (grant No. 816-517).

Author contributions

JG, SH, and XL designed the research. SH, and XJ developed the algorithm. SH, FZ, and MZ analyzed data. SH wrote the manuscript with the contributions from JG. All authors were involved in discussions throughout the development.

Funding

The National Natural Science Foundation of China (Grant No. 42274006) and the Autonomous and Controllable Special Project for Surveying and Mapping of China (Grant No. 816-517) for supporting this study.

Availability of data and materials

The PWV_{GNSS} data are available at http://www.cgps.ac.cn. The *PM10* data are available at http://www.cnemc.cn/sssj. The precipitation data are available at https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels?tab=form.

Declarations

Competing interests
All authors have no competing interests.

Author details
[1]College of Geodesy and Geomatics, Shandong University of Science and Technology, Qingdao 266590, China. [2]Institude of Oceanographic Instrument, Qilu University of Technology (Shandong Academy of Sciences), Qingdao 266100, China.

Received: 29 May 2023 Accepted: 6 August 2023
Published online: 21 August 2023

References
Alghamdi MA, Almazroui M, Shamy M, Redal MA, Alkhalaf AK, Hussein MA, Khode ML (2015) Characterization and elemental composition of atmospheric aerosol loads during springtime dust storm in western Saudi Arabia. Aerosol Air Qual Res 15(2):440–453. https://doi.org/10.4209/aaqr.2014.06.0110

Bevis M, Businger S, Herring TA, Rocken C, Anthes RA, Ware RH (1992) GPS meteorology: remote sensing of atmospheric water vapor using the global positioning system. J Geophys Res 97(D14):15787–15801. https://doi.org/10.1029/92jd01517

Bevis M, Businger S, Chiswell S, Herring TA, Anthes RA, Rocken C, Ware RH (1994) GPS meteorology: mapping zenith wet delays onto precipitable water. J Appl Meteorol Clim. https://doi.org/10.1175/1520-0450(1994)033%3c0379:GMMZWD%3e2.0.CO;2

Chang L, Gao GP, Jin SG, He XF, Xiao RY, Guo LX (2015) Calibration and evaluation of precipitable water vapor from MODIS infrared observations at night. IEEE Trans Geosci Remote Sens 53(5):2612–2620. https://doi.org/10.1109/TGRS.2014.2363089

Chen Q, Dam TV, Sneeuw N, Collilieux X, Weigelt M, Rebischung P (2013) Singular spectrum analysis for modeling seasonal signals from GPS time series. J Geodyn 72:25–35. https://doi.org/10.1016/j.jog.2013.05.005

David N, Gao HO (2016) Using cellular communication networks to detect air pollution. Environ Sci Technol 50(17):9442–9451. https://doi.org/10.1021/acs.est.6b00681

Guo JY, Gao WZ, Yu HJ, Liu X, Kong QL, Chen XD (2018) Gravity tides extracted from relative gravimetric data with singular spectrum analysis. Chin J Geophys 61(10):3889–3902. https://doi.org/10.6038/cjg2018L0460

Guo M, Zhang HW, Xia PF (2020) A method for predicting short-time changes in fine particulate matter (PM2.5) mass concentration based on the global navigation satellite system zenith tropospheric delay. Meteorol App 27(1):1350–4827. https://doi.org/10.1002/met.1866

Guo JY, Hou R, Zhou MS, Jin X, Li GW (2021a) Detection of particulate matter changes caused by 2020 California wildfires based on GNSS and radiosonde station. Remote Sens 13(22):4557. https://doi.org/10.3390/rs13224557

Guo M, Xia PF, Li PJ, Zhang HW (2021b) Global navigation satellite system precipitable water vapour combined with other atmospheric factors to predict the short-term change of PM2.5 mass concentration. Meteorol Zeitschrift 30(5):429–444. https://doi.org/10.1127/metz/2021/1061

Gurbuz G (2021) On variations of the decadal precipitable water vapor (PWV) over Turkey. Adv Space Res 68(1):292–300. https://doi.org/10.1016/j.asr.2021.03.010

Gurbuz G, Bayik GD (2023) Impact of sand and dust storms on tropospheric parameter estimation by GPS. Environ Monit Assess 195(2):332. https://doi.org/10.1007/s10661-023-10956-w

Han J, Dai H, Gu ZL (2021) Sandstorms and desertification in Mongolia, an example of future climate events: a review. Environ Chem Lett 19(6):4063–4073. https://doi.org/10.1007/s10311-021-01285-w

Hassani H (2007) Singular spectrum analysis: methodology and comparison. J Data Sci 5(2):239–257. https://doi.org/10.6339/JDS.2007.05(2).396

Huo W, Zhi XF, Yang LM, Ali M, Zhou CL, Yang F, Yang XH, Meng L, He Q (2019) Research progress on several problems of desert meteorology. Trans Atmos Sci 42(3):469–480. https://doi.org/10.13878/j.cnki.dqkxxb.20180629001

Jin X, Liu X, Guo JY, Shen Y (2021) Analysis and prediction of polar motion using MSSA method. Earth Planets Space 73(1):1–13. https://doi.org/10.1186/s40623-021-01477-2

Kiser D, Metcalf WJ, Elhanan G, Schnieder B, Schlauch K, Joros A, Petersen C, Grzymski J (2020) Particulate matter and emergency visits for asthma: a time-series study of their association in the presence and absence of wildfire smoke in Reno, Nevada, 2013–2018. Environ Health 19(1):92. https://doi.org/10.1186/s12940-020-00646-2

Lau L, He J (2017) Investigation into the effect of atmospheric particulate matter (PM2.5 and PM10) concentrations on GPS signals. Sensors 17(3):508. https://doi.org/10.3390/s17030508

Lavers DA, Simmons A, Vamborg F, Rodwell MJ (2022) An evaluation of ERA5 precipitation for climate monitoring. Q J R Meteorol Soc 148(748):3124–3137. https://doi.org/10.1002/qj.4351

Liu YX, Chen YQ, Liu JN (2000) Remote sensing of water vapor content using ground-based GPS data. Geo-Spatial Inf Sci 3:64–68. https://doi.org/10.1007/BF02826612

Liu X, Wang Y, Huang J, Yu TL, Jiang NH, Yang J, Zhan W (2022) Assessment and calibration of FY-4A AGRI total precipitable water products based on CMONOC. Atmos Res 6(1):7–15. https://doi.org/10.1016/j.atmosres.2022.106096

Paun M, Marghescu I, Tamas R (2015) A Software Radio approach for locating unintentional ozone-generating sources. IEEE International Black Sea Conference on Communications & Networking. IEEE 2015. https://doi.org/10.1109/BlackSeaCom.2015.7185096

Saastamoinen J (1972) Atmospheric correction for the troposphere and stratosphere in radio ranging satellites. Use Artif Satell Geod 15:247–251. https://doi.org/10.1029/GM015p0247

Shen Y, Guo JY, Liu X, Kong QL, Guo LX, Li W (2018) Long-term prediction of polar motion using a combined SSA and ARMA model. J Geod 92(3):333–343. https://doi.org/10.1007/s00190-017-1065-3

Shi YF, Shen YP, Kang E, Li DL, Ding YJ, Zhang GW, Hu RJ (2007) Recent and future climate change in Northwest China. Clim Change 80:379–393. https://doi.org/10.1007/s10584-006-9121-7

Shi C, Zhang WX, Cao YC, Lou YD, Liang H, Fan L, Satirapod C, Trakolkul C (2020) Atmospheric water vapor climatological over Indo-China region based on Beidou/GNSS and relationships with precipitation. AGCS 49(9):1112–1119. https://doi.org/10.11947/j.AGCS.2020.20200339

Shi C, Zhou LH, Fan L, Zhang WX, Cao YC, Wang C, Xiao F et al (2022a) Analysis of "21·7" extreme rainstorm process in Henan Province using BeiDou/GNSS observation. Chin J Geophys 65(1):186–196. https://doi.org/10.6038/cjg2022P0706

Shi T, Liu X, Mu DP, Guo JY, Xing YP (2022b) Reconstructing gap data between GRACE and GRACE-FO based on multi-layer perceptron and analyzing terrestrial water storage changes in the Yellow River basin. Chinese Journal of Geophysics 65(7):2448-2463. https://doi.org/10.6038/cjg2022P0291

Tao Y, Liu C, Liu CY, Zhao XW, Hu HJ (2021) Empirical Wavelet Transform Method for GNSS Coordinate Series Denoising. Journal of Geovisualization and Spatial Analysis 5(1):9. https://doi.org/10.1007/s41651-021-00078-7

Vautard R, Yiou P, Ghil M (1992) Singular-spectrum analysis: a toolkit for short, noisy chaotic signals. Physica D 158(1–4):95–126. https://doi.org/10.1016/0167-2789(92)90103-t

Wang Y, Liu YP, Li JB, Liu LT (2016) The correlation between the variation of PM2.5/PM10 and precipitable water vapor based on GPS and radiosonde. Geomat Inf Sci Wuhan Univ 41(12):1626–1631. https://doi.org/10.13203/j.whugis20140628

Watanabe M, Noma H, Kurai J, Kato K, Sano H, Tatsukawa T, Nakazaki H, Yamasaki A, Shimizu E (2016) Association between pulmonary function and daily levels of sand dust particles assessed by light detection and ranging in schoolchildren in western Japan: A panel study. Allergol Int 65(1):56–61. https://doi.org/10.1016/j.alit.2015.07.005

Wen HF, Dang YM, Li LW (2020) Short-Term PM2.5 Concentration Prediction by Combining GNSS and Meteorological Factors. IEEE Access 8:115202–115216. https://doi.org/10.1109/ACCESS.2020.3003580

Wyatt MG, Kravtsov S, Tsonis AA (2012) Atlantic multidecadal oscillation and northern hemisphere's climate variability. Clim Dyn 38(5):929–949. https://doi.org/10.1007/s00382-011-1071-8

Yang F (2022) Research on the key technologies in water vapor retrieval using ground-based GNSS. AGCS 51(3):470. https://doi.org/10.11947/j.AGCS.2022.20200524

Yang XJ, Zhao XY, Wu FS, Zhang ZM, Xue P, Chen Z, Wang WF, Zhang GB (2021) Relationship of PM10 concentration in Mogao Grottoes to meteorological elements. J Desert Res 41(6):54–64. https://doi.org/10.7522/j.issn.1000-694X.2021.00089

Yao YB, Zhu S, Yue SQ (2012) A global applicable, season-specific model for estimating the weighted mean temperature of the atmosphere. J Geod 86:1125–1135. https://doi.org/10.1007/s00190-012-0568-1

Yao L, Lu N, Yue XF, Du J, Yang CD (2015) Comparison of hourly PM2.5 observations between urban and suburban areas in Beijing, China. IJERPH 12(10):12264–12276. https://doi.org/10.3390/ijerph121012264

Yu JS, Tan K, Zhang CH, Zhao B, Wang DZ, Li Q (2019) Present-day crustal movement of the Chinese mainland based on Global Navigation Satellite System data from 1998 to 2018. Adv Space Res 63(2):840–856. https://doi.org/10.1016/j.asr.2018.10.001

Zabalza J, Ren JC, Wang Z, Marshall S, Wang J (2014) Singular spectrum analysis for effective feature extraction in hyperspectral imaging. IEEE Geosci Remote Sens Lett 11(11):1886–1890. https://doi.org/10.1109/LGRS.2014.2312754

Zhai P, Eskridge RE (1996) Analyses of inhomogeneities in radiosonde temperature and humidity time series. J Clim 9(4):884–894. https://doi.org/10.1175/1520-0442(1996)009%3c0884:AOIIRT%3e2.0.CO;2

Zhang S, Xu H, Lan JH, Goldsmith Y, Torfstein A, Zhang GL, Song YP, Zhou KE, Tan L, Xu S, Xu XM, Enzel Y (2021) Dust storms in northern China during the last 500 years. Sci China Earth Sci 64:813–824. https://doi.org/10.1007/s11430-020-9730-2

Zhang L, Song GY, Fan F, Wu H, Zhou YH, Zhou ZH, Zhang X (2022a) Diagnosis of sandstorm weather process and analysis of sand pollution transportation in northern China from 14th to 16th. Acta Sci Circumstantiae 42(9):351–363. https://doi.org/10.13671/j.hjkxxb.2021.0452

Zhang WY, Zhang SB, Zheng NS, Zhang QS, Ding N (2022b) Study on the retrieval of 3D atmospheric water vapor distribution using GNSS data and RS multi-source data. Chin J Geophys 65(6):1951–1964. https://doi.org/10.6038/cjg2022P0078

Zhao B, Huang Y, Zhang CH, Wang W, Tan K, Du RL (2015) Crustal deformation on the Chinese mainland during 1998–2004 based on GPS data. Geod Geodyn 6(1):7–15. https://doi.org/10.1016/j.geog.2014.12.006

Zhou MS, Guo JY, Liu X, Hou R, Jin X (2022) Analysis of GNSS-derived tropospheric zenith non-hydrostatic delay anomaly during sandstorms in Northern China on 15th March 2021. Remote Sens 14(18):4678. https://doi.org/10.3390/rs14184678

Zhu J, Li Y, Wang T, Chang ZQ, Yu W, Han YF, Wang YB, Liu YY (2020) An improved atmospheric phase delay correction method in spaceborne repeat-track INSAR monitoring. J Geod 40(11):1164–1169. https://doi.org/10.14075/j.jgg.2020.11.012

Zhu B, Zhang Q, Yang JH, Li CH (2022) Response of potential evapotranspiration to warming and wetting in Northwest China. Atmosphere 13(2):353. https://doi.org/10.3390/atmos13020353

Publisher's Note