



Article

Evaluation and Correction of GFS Water Vapor Products over United States Using GPS Data

Hai-Lei Liu ^{1,*} , Xiao-Qing Zhou ¹, Yu-Yang Zhu ², Min-Zheng Duan ^{3,4} , Bing Chen ⁵ and Sheng-Lan Zhang ¹

¹ Key Laboratory of Atmospheric Sounding, Chengdu University of Information Technology, Chengdu 610225, China; 3220308035@stu.cuit.edu.cn (X.-Q.Z.); zslan@cuit.edu.cn (S.-L.Z.)

² Guangxi Meteorological Technical Equipment Center, Guangxi Zhuang Autonomous Region Meteorological Bureau, Nanning 530022, China; 3210307062@stu.cuit.edu.cn

³ Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing 100029, China; dmz@mail.iap.ac.cn

⁴ College of Earth and Planetary Sciences, University of Chinese Academy of Sciences, Beijing 100049, China

⁵ Department of Atmospheric Science, Yunnan University, Kunming 650500, China; chenbing@ynu.edu.cn

* Correspondence: liuhailei@cuit.edu.cn

Abstract: Precipitable water vapor (PWV) is one of the most dynamic components of the atmosphere, playing a critical role in precipitation formation, the hydrological cycle, and climate change. This study used SuomiNet Global Positioning System (GPS) data from April 2021 to June 2023 in the United States to comprehensively evaluate 3 and 6 h Global Forecast System (GFS) PWV products (i.e., PWV_{3h} and PWV_{6h}). There was high consistency between the GFS PWV and GPS PWV data, with correlation coefficients (Rs) higher than 0.98 and a root mean square error (RMSE) of about 0.23 cm. The PWV_{3h} product performed slightly better than PWV_{6h}. PWV tended to be underestimated when PWV > 4 cm, and the degree of underestimation increased with increasing water vapor value. The RMSE showed obvious seasonal and diurnal variations, with the RMSE value in summer (i.e., 0.280 cm) considerably higher than in winter (i.e., 0.158 cm), and nighttime were RMSEs higher than daytime RMSEs. Clear-sky conditions showed smaller RMSEs, while cloudy-sky conditions exhibited a smaller range of monthly RMSEs and higher Rs. PWV demonstrated a clear spatial pattern, with both Rs and RMSEs decreasing with increasing elevation and latitude. Based on these temporal and spatial patterns, Back Propagation neural network and random forest (RF) models were employed, using PWV, Julian day, and geographic information (i.e., latitude, longitude, and elevation) as input data to correct the GFS PWV products. The results indicated that the RF model was more advantageous for water vapor correction, improving overall accuracy by 12.08%. In addition, the accuracy of GFS PWV forecasts during hurricane weather was also evaluated. In this extreme weather, the RMSE of the GFS PWV forecast increased comparably to normal weather, but it remained less than 0.4 cm in most cases.

Keywords: PWV; GFS; SuomiNet GPS; machine learning; evaluation and correction



Citation: Liu, H.-L.; Zhou, X.-Q.; Zhu, Y.-Y.; Duan, M.-Z.; Chen, B.; Zhang, S.-L. Evaluation and Correction of GFS Water Vapor Products over United States Using GPS Data. *Remote Sens.* **2024**, *16*, 3043. <https://doi.org/10.3390/rs16163043>

Academic Editors: Noel Aloysius, Mircea Alexe and Iulian-Horia Holobacă

Received: 30 June 2024

Revised: 11 August 2024

Accepted: 15 August 2024

Published: 19 August 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Precipitable water vapor (PWV) is the total water vapor in a vertical column of a unit cross-sectional area extending from the Earth's surface to the top of the atmosphere [1,2]. As a key physical quantity characterizing atmospheric water vapor content, PWV has important applications in numerical weather prediction, water resource management, agriculture, and ecosystem protection [3,4]. In addition, it is one of the most important greenhouse gases in the atmosphere, influencing the Earth's ecological balance and global climate change [5,6].

In recent decades, the scientific community has made great efforts to accurately characterize PWV in the atmosphere, and various effective techniques have been developed [7], such as satellite remote sensing technology, radiosonde technology, microwave radiometers, and ground-based observation [8–11].

Satellite remote sensing technology can assess atmospheric water vapor content using microwaves or infrared radiation information [12]. Although satellite observations are not constrained by geographical limitations, they have drawbacks [13]. For instance, algorithms for observing water vapor with geostationary meteorological satellites primarily use infrared detection, which is affected by atmospheric disturbances and cloud cover, reducing measurement accuracy and limiting PWV retrieval to clear-sky conditions [14]. Polar-orbiting satellites can detect water vapor using near-infrared, infrared, and microwave wavelengths [15,16]. While microwave detection provides all-weather PWV data, the water vapor information obtained via microwaves from polar-orbiting satellites is affected by the uncertainty in microwave surface emissivity, resulting in more difficult PWV estimation over land and poorer PWV forecast accuracy [17]. Additionally, polar-orbiting satellites can only provide data twice a day, leading to low temporal resolution [18].

Radiosonde technology can provide highly accurate vertical water vapor information by launching weather balloons to measure atmospheric parameters such as temperature, humidity, and pressure at different altitudes, thus calculating the PWV value [19–21]. However, weather balloons are usually only launched twice a day, and equipment sites are sparsely distributed, which limits the sonde technology to providing PWV information only in specific areas [22]. Consequently, it cannot reflect the PWV gradient within a region and is unsuitable for continuous global observations [23]. Microwave radiometer technology estimates PWV values by analyzing the absorption characteristics of microwave radiation in the atmosphere. While this method can ultimately provide atmospheric information through data processing and inversion, it is costly and involves complex data processing [24].

Ground-based observations (e.g., the Global Positioning System, GPS) have extended from traditional positioning applications to investigating both the upper and lower atmosphere [25]. GPS can provide high-resolution continuous measurements of zenith tropospheric delay, from which near-real-time total PWV around ground-based GPS receivers can be derived [26,27]. GPS water vapor measurements have high temporal resolution (ranging from 5 min to 2 h), all-weather capabilities, and relative simplicity, making them reliable reference data (with an accuracy of approximately 2 mm) [28,29]. Owing to the advantages of ground-based GNSS in water vapor monitoring, researchers are developing continuously operating GPS receiver networks for climate research and meteorological forecasting applications [30], including SuomiNet and COCONet in the United States; E-GVAP in Western Europe; and the Crustal Movement Observation Network of China (CMONOC) and the China Meteorological Administration GNSS network (CMAGN) in China [31–33].

In addition, the Global Forecast System (GFS) model runs predictions daily at 00, 06, 12, and 18 UTC, simulating the movement of and changes in atmospheric systems to provide seamless global water vapor data [34]. Furthermore, PWV products from numerical prediction and assimilation models are widely used in weather forecasting and climate change research. However, the factors affecting PWV simulations are extremely complex, covering temporal distribution, surface characteristics, and weather and climate conditions. At present, there are few validations for GFS PWV products. In 2022, A. K. Kishkina et al. [35] compared the integral water vapor content of the atmosphere at 13 observation points of the GFS and GNSS network in the Primorski Krai region of Russia. This study validates the three-hour forecast PWV product (PWV_{3h}) and the six-hour forecast PWV product (PWV_{6h}) provided by the GFS for the contiguous United States from April 2021 to June 2023 using atmospheric PWV data from SuomiNet GPS network observation stations. We attempted to correct the water vapor using two machine learning models, namely the Back Propagation (BP) neural network and random forest (RF) model.

Section 2 describes the datasets (including GPS, the GFS, and auxiliary data), data preprocessing process, correction methods, and statistical metrics in this study. In Section 3, Section 3.1 analyzes the overall performance of the GFS PWV products, and Sections 3.2–3.6 examine PWV errors in terms of seasonal, diurnal, spatial, clear-sky, cloudy-sky, and

extreme weather conditions. GFS correction results for water vapor using BP neural networks and RF models are discussed in Section 3.7. Finally, the conclusions are presented in Section 4.

2. Materials and Methods

2.1. Data

In this study, the SuomiNet GPS PWV data and GFS PWV products (i.e., PWV_{3h} and PWV_{6h}) were primarily used. In addition, auxiliary information such as longitude, latitude, Julian day (JD), and elevation was also used.

2.1.1. SuomiNet GPS Data

The GPS-derived PWV data from SuomiNet were used as reference data to evaluate the retrieved PWV in this study [36]. SuomiNet—named in homage to the meteorological satellite trailblazer and University of Wisconsin Professor Verner Suomi—constitutes a proposed network of GPS receivers strategically positioned at universities and other locations [37,38]. This network aims to furnish real-time measurements of atmospheric precipitable water vapor, along with other geodetic and meteorological information. SuomiNet can observe signals from GPS satellites with millimeter-level accuracy under various weather conditions, providing continuous, accurate, all-weather, real-time PWV data within a regional range [39].

The total tropospheric delay of GPS signals is the sum of the wet delay caused by water vapor and the dry delay induced by atmospheric dry components [40]. GPS PWV is primarily estimated from the zenith wet delay (ZWD), which involves a complex analysis of GPS signal reception delays [41]. This process includes precise pressure measurements; sampling and smoothing different ZWDs; and correcting for satellite orbits and clock information to minimize system errors and enhance the accuracy of PWV estimation [28]. Ultimately, the PWV data calculation method involves the product of the estimated zenith humidity and a conversion factor [38,42]. The conversion factor depends on the weighted mean temperature of the atmosphere and can be inferred from surface temperature measurements or accurate numerical weather models [43]. With further data processing and quality control, GPS PWV estimation offers high spatial and temporal resolution, making it suitable for meteorological monitoring and climate studies.

In this study, SuomiNet GPS data (330 stations) from April 2021 to June 2023 in the United States (25–55°N, 70–130°W) were used. The statistics of GPS stations at different elevations are shown in Table 1. The stations were mainly concentrated below 2 km, accounting for 93.64% of the total. Figure 1 shows the distribution of station locations and the average water vapor at each station within the study area. The eastern region is predominantly plains with lower elevations, while the western region is mainly mountainous with higher elevations. The selected stations were evenly distributed across the study area, ensuring representativeness. Overall, the average PWV at the stations ranged from 0.5 to 4.0 cm. The PWV values were influenced by elevation, with the eastern region generally exhibiting a higher average PWV than the western region.

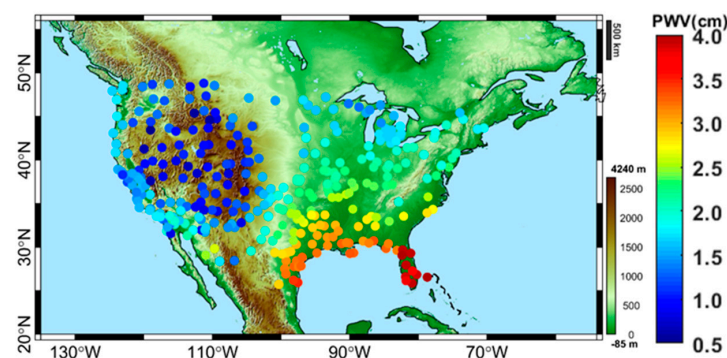


Figure 1. Spatial distribution of total average Global Positioning System precipitable water vapor (GPS PWV) at 330 stations in the continental United States.

Table 1. Information for GPS stations at different altitudes.

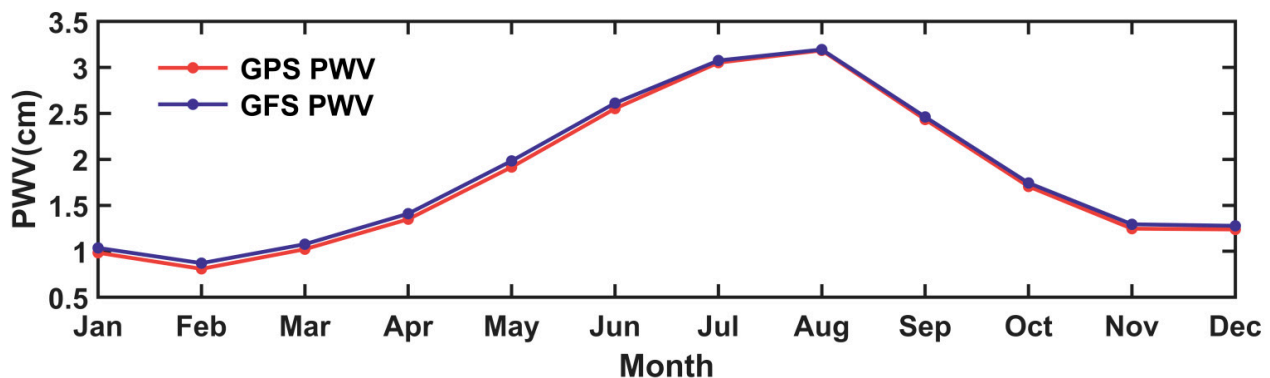
Elevation (km)	Number of Sites	Percent (%)	Average Height (km)
<0.0	3	0.91	−0.03
0.0–1.0	247	74.85	0.28
1.0–2.0	59	17.88	1.49
2.0–3.0	21	6.36	2.28

2.1.2. GFS Data

The GFS is a global numerical weather prediction system maintained by the U.S. National Oceanic Service and Atmospheric Administration (NOAA) that provides forecast data in many resolutions and at different time intervals, with spatial resolutions ranging from 0.25° to 1° (longitude/latitude) [44]. The prediction time steps are 0 to 120 h (1 h intervals) and 120 to 384 h (3 h intervals). GFS forecasts are updated four times per day (i.e., 00, 06, 12, and 18 UTC), providing forecasts for up to 16 days [45]. However, for GFS historical data, only three-hourly and six-hourly data were available for 2021 to 2023. Therefore, three-hourly and six-hourly GFS PWV forecasts at a resolution of 0.25° grid were used in this study [46,47].

The GFS model's 15th upgrade was officially implemented on 12 June 2019. This upgrade introduced several key improvements and new features to enhance the accuracy and precision of weather forecasts [44]. The model's core was upgraded to use the Finite-Volume Cubed-Sphere Dynamical Core (FV3), replacing the previous dynamical core. Additionally, the 15th version of GFS employs a Four-Dimensional Variational (4D-Var) data assimilation method, upgrading the data assimilation system and incorporating a wider range of satellite data, including data from polar-orbiting and geostationary satellites.

Figure 2 presents preliminary statistics of the monthly average PWV from GPS and the GFS (including PWV_{3h} and PWV_{6h}) across all stations from April 2021 to June 2023. There is a high consistency between GPS PWV and GFS PWV data. They show seasonal characteristics, with higher precipitation in summer than in winter. More detailed GFS PWV product validation results are presented in Section 3.

**Figure 2.** Time series of monthly mean GPS and GFS PWV, including three-hour (PWV_{3h}) and six-hour forecasts (PWV_{6h}).

2.2. Methods

2.2.1. Data Processing

We used the nearest-neighbor method to spatially match spatiotemporally equivalent GPS, GFS, and auxiliary data. Since water vapor represents the water content in the entire atmospheric column, it highly depends on elevation. Considering that GFS data are gridded ($0.25^\circ \times 0.25^\circ$), the GFS PWV data were corrected for elevation using GPS and GFS elevation data to ensure validation accuracy [48,49]. The correction equation is as follows:

$$PWV = PWV_0 \times \exp\left(\frac{C\Delta h}{1000}\right) \quad (1)$$

where PWV_0 and PWV are the values of the GFS water vapor products before and after correction, respectively. Δh (unit: m) is the elevation difference between GFS and GPS after interpolation to the same station. C is equal to 0.439 [50].

2.2.2. Neural Networks and Random Forest Model

Neural networks can capture the complex nonlinear relationships between input and output variables and have been widely applied in the meteorology field [51,52]. The random forest (RF) model is highly accurate and robust and uses the integration method to fuse multiple decision trees to obtain an optimal result, benefitting the processing of high-dimensional and large-scale datasets [53,54]. As the sample is input, each decision tree individually performs evaluation and regression. The predictions of all trees are then aggregated to reduce overfitting and improve prediction accuracy.

In this study, multilayer Back Propagation (BP) neural networks and RF models were employed to construct correction models for PWV. The flowcharts of these models are presented in Figures 3 and 4. After comprehensive validation, two correction models for water vapor products were established based on observed patterns. To enhance prediction accuracy, models were developed separately for each month, from January to December. The inputs for each model include GFS PWV, JD, and geographical information (i.e., latitude, longitude, and altitude), with the output being consistent with the GPS PWV. The dataset was partitioned into rows, with odd-numbered rows used to construct the training sets for the neural network and random forest models and even-numbered rows used to construct the validation sets.

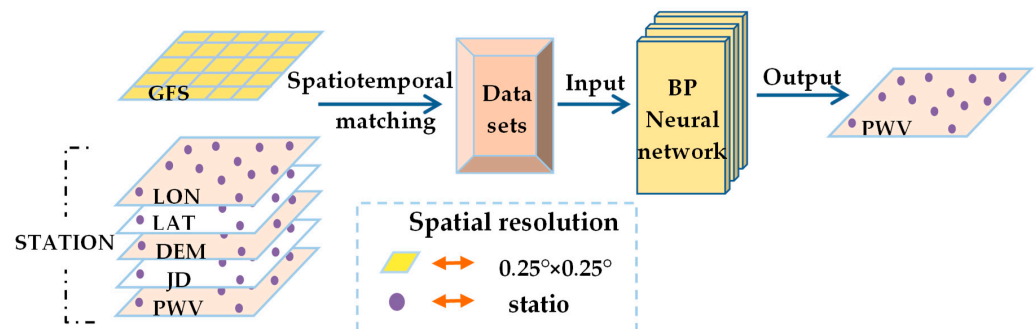


Figure 3. Flowchart of the Back Propagation (BP) neural network PWV estimation model.

In the model construction process, the size of the hidden layer was set from 0 to 80 with an interval of 10 for neural network training and validation. When the hidden layer size reached 60, the optimization percentage of root mean square error (RMSE) began to level off, showing that the model's performance was relatively optimal. Similarly, for the RF model, experiments were conducted with the number of numTrees set at intervals of 10. When the number of decision trees reached 80, the performance stabilized.

2.2.3. Statistical Metrics

The accuracy of the GFS PWV products and the correction results of the models were evaluated using the correlation coefficient (R), RMSE, bias, and standard deviation (STD). The formulas for these four statistical indicators are presented below.

$$R = \frac{\sum_{i=1}^n (PWV_{GFS} - \overline{PWV_{GFS}})(PWV_{GPS} - \overline{PWV_{GPS}})}{\sqrt{\sum_{i=1}^n (PWV_{GFS} - \overline{PWV_{GFS}})^2 \sum_{i=1}^n (PWV_{GPS} - \overline{PWV_{GPS}})^2}} \quad (2)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (PWV_{GFS} - PWV_{GPS})^2} \quad (3)$$

$$Bias = \frac{1}{N} \sum_{i=1}^N (PWV_{GFS} - PWV_{GPS}) \quad (4)$$

$$\text{STD} = \sqrt{\frac{1}{N} \sum_{i=1}^N (PWV_{\text{GFS}} - PWV_{\text{GPS}} - \mu)^2} \quad (5)$$

where PWV_{GFS} is the original PWV from GFS; PWV_{GPS} is the PWV measured at GPS stations; $\overline{PWV_{\text{GFS}}}$ and $\overline{PWV_{\text{GPS}}}$ are the mean values of GFS PWV and GPS PWV, respectively; μ is the mean difference between the two datasets; and N is the sample size.

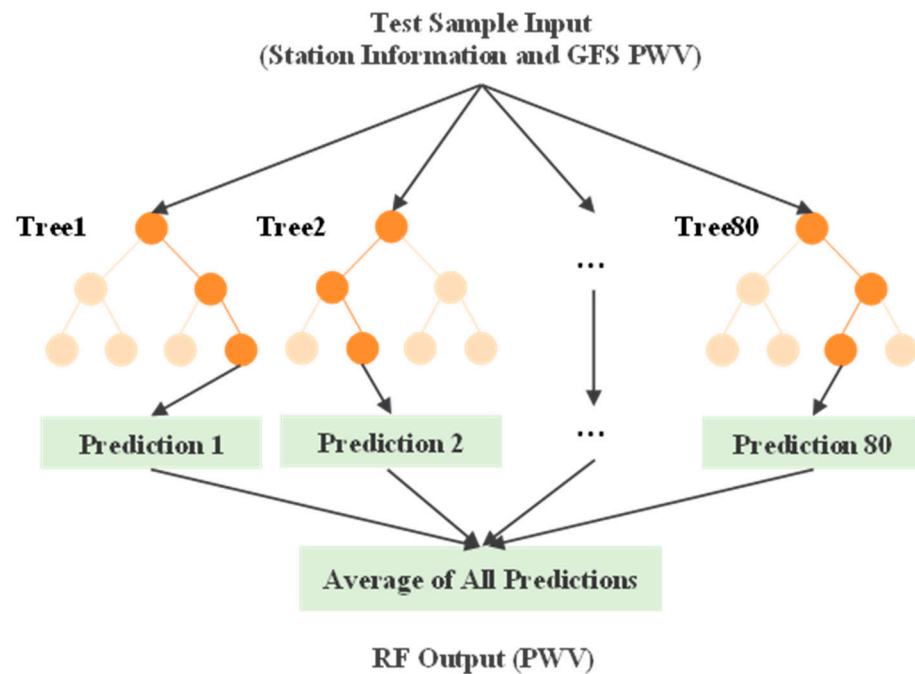


Figure 4. Flowchart of the RF PWV estimation model.

In general, the R primarily assesses the degree of consistency between datasets, although it does not imply absolute consistency. The RMSE indicates the magnitude of differences between GPS and GFS PWV. Positive/negative bias represents an overestimation/underestimation of PWV. The STD is a statistical measure of data dispersion.

3. Results

3.1. Overall Performance of GFS PWV Products

Figure 5 presents two-dimensional histograms of GFS PWV_{3h} (Figure 5a) and PWV_{6h} (Figure 5b) products compared with GPS PWV data. The GFS PWV product and GPS PWV data are in good agreement, with an R above 0.98. The R observed in this study is more precise compared to the GFS water vapor predictions for the Primorski Krai region of Russia (where the average R ranges from 0.9 to 0.97) [35]. This difference may be attributed to variations in topography, climate, latitude, and other regional factors. The RMSE values are 0.227 cm and 0.233 cm, respectively. The densely populated regions of the data are evenly distributed around the fit line. The STD values of the GFS PWV and GPS PWV are small, indicating that the overall difference has not deviated much. In addition, the biases for PWV_{3h} and PWV_{6h} are 0.05 cm and 0.04 cm, both positive. This means that GFS PWV is slightly overestimated, assuming that GPS PWV is unbiased.

Studies on water vapor products from meteorological satellites indicate that errors in water vapor retrieval are related to the amount of water vapor.

In this study, PWV values between 0 and 7 cm were divided into 14 PWV bins according to a 0.5 cm interval. The mean bias and STD of each PWV bin are shown in Figure 6. When the PWV is less than 4 cm, the mean bias is smoother, but it still shows a slight positive bias, as demonstrated by the bias statistics in Table 2. When the PWV is greater than 4 cm, GFS PWV is underestimated compared with GPS PWV, and the averaged biases of the PWV bins are -0.186 cm (PWV_{3h}) and -0.241 cm (PWV_{6h}), respectively. The

STD range increases with higher PWV, indicating greater dispersion in differences at higher PWV values. Compared with PWV_{3h} , PWV_{6h} has a larger negative bias of -0.834 cm. The bias of PWV over different bins is similar to the findings of Roman [37], who validated satellite PWV (AIRS) using the SuomiNet network, generally showing an overestimation of low PWV and an underestimation of high PWV. In addition, the R decreases and the RMSE increases with increasing PWV, as shown in Table 2.

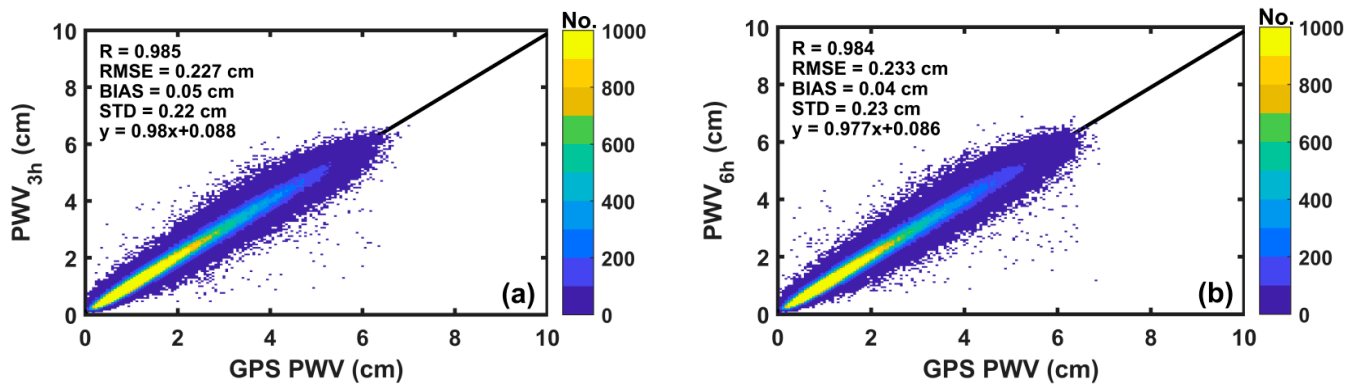


Figure 5. Two-dimensional histogram for GPS PWV versus GFS PWV_{3h} (a) and PWV_{6h} (b) products.

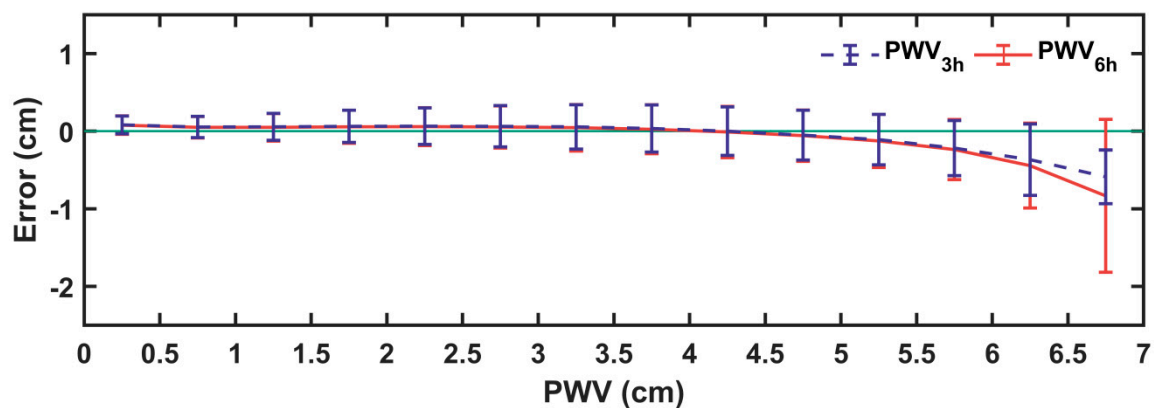


Figure 6. Error statistics diagram of 0.5 cm PWV bin.

Table 2. Summary statistics of statistical indicator changes in GPS PWV.

3 h	Intervals (cm)	Number	R	RMSE (cm)	Bias (cm)
	0–1	203,534	0.850	0.145	0.062
	1–2	204,541	0.839	0.198	0.059
	2–3	125,440	0.754	0.258	0.064
	3–4	86,271	0.683	0.298	0.046
	4–5	48,228	0.627	0.318	−0.020
	5–6	13,454	0.495	0.362	−0.135
	6–7	507	0.177	0.594	−0.379
6 h	Intervals (cm)	Number	R	RMSE (cm)	Bias (cm)
	0–1	195,429	0.845	0.146	0.059
	1–2	196,252	0.833	0.200	0.053
	2–3	119,574	0.744	0.263	0.056
	3–4	82,453	0.668	0.309	0.035
	4–5	47,207	0.614	0.332	−0.030
	5–6	13,583	0.472	0.388	−0.152
	6–7	578	0.036	0.742	−0.462

3.2. Seasonal Variations in GFS PWV Accuracy

Figure 7 presents a boxplot of the monthly time series for GPS and GFS PWV_{3h} and PWV_{6h}. The precipitable water in this region shows clear seasonal variations. The figure shows that from January to March, the precipitable water is relatively low and varies less, with a median of about 1 cm. Precipitable water increases slightly in spring and autumn (April, May, and October), but the range of fluctuations remains relatively small, with a median of about 1.6 cm. In summer, the PWV reaches its peak from June to August, and its median is about 3 cm. The overall trend in Figure 7 is consistent with Figure 8a. Considering that the effect of climate may lead to differences in GFS PWV accuracy in different seasons, statistical metrics are used to calculate the monthly average R, RMSE, and bias for the selected region to study the water vapor values and accuracy issues of GFS PWV_{3h} and PWV_{6h} from January to December. The results are presented in Figure 8.

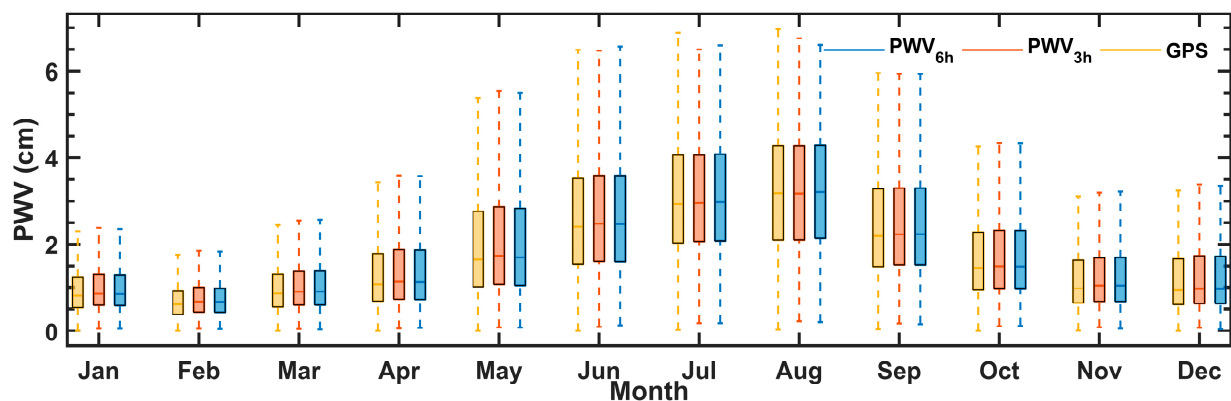


Figure 7. Boxplot of the monthly time series for GPS (yellow), GFS PWV_{3h} (red), and PWV_{6h} (blue).

Overall, the monthly average R values for GFS PWV_{3h} and GFS PWV_{6h} are above 0.97, with good accuracy. The monthly R values for PWV_{3h} are slightly higher than those for PWV_{6h}. The R value is also higher in winter (December, January, and February) than in summer (June, July, and August). However, PWV_{3h} exhibits a larger monthly bias than PWV_{6h}, with both showing the smallest bias in August and the largest in May.

The RMSE varies between 0.14 cm and 0.3 cm, with higher values in summer than in winter, consistent with water vapor variation trends (Figure 8a), increasing and then decreasing. The PWV_{3h} and PWV_{6h} predictions show the lowest RMSEs in February (i.e., 0.143 cm and 0.147 cm) and the highest in July (i.e., 0.282 cm and 0.290 cm). The seasonal variation in RMSE may be attributed to differences in data assimilation processes between the GFS model and GPS data, with the model's assimilation capabilities possibly being challenged by the complex convective activity and water vapor distribution during summer, leading to increased RMSE. The bias distribution does not show an obvious seasonal variation range for these data. Only the PWV_{6h} bias in August is negative, further validating the overall overestimation of the GFS PWV products. The United States is located in the Northern Hemisphere and experiences heavy rainfall during the summer, especially from June to September, often accompanied by hurricane activity. Figure 6 shows that high PWV tends to be underestimated, increasing the underestimated data from June to September. This may be one of the reasons for the small and even negative bias in summer.

3.3. Diurnal Variations in GFS PWV Accuracy

The distribution of and variation in PWV in the atmosphere are influenced by the diurnal cycle, with fluctuations in water vapor content and humidity driven by differences in day and night temperatures, convective activity, cloud evolution, and surface heat exchange. To verify the diurnal accuracy of GFS products, PWV_{3h} (03, 09, 15, and 21 UTC) and PWV_{6h} (06, 12, 18, and 00 UTC) were used as the total dataset, and statistics for R, RMSE, and bias for the eight time points were computed, as shown in Figure 9.

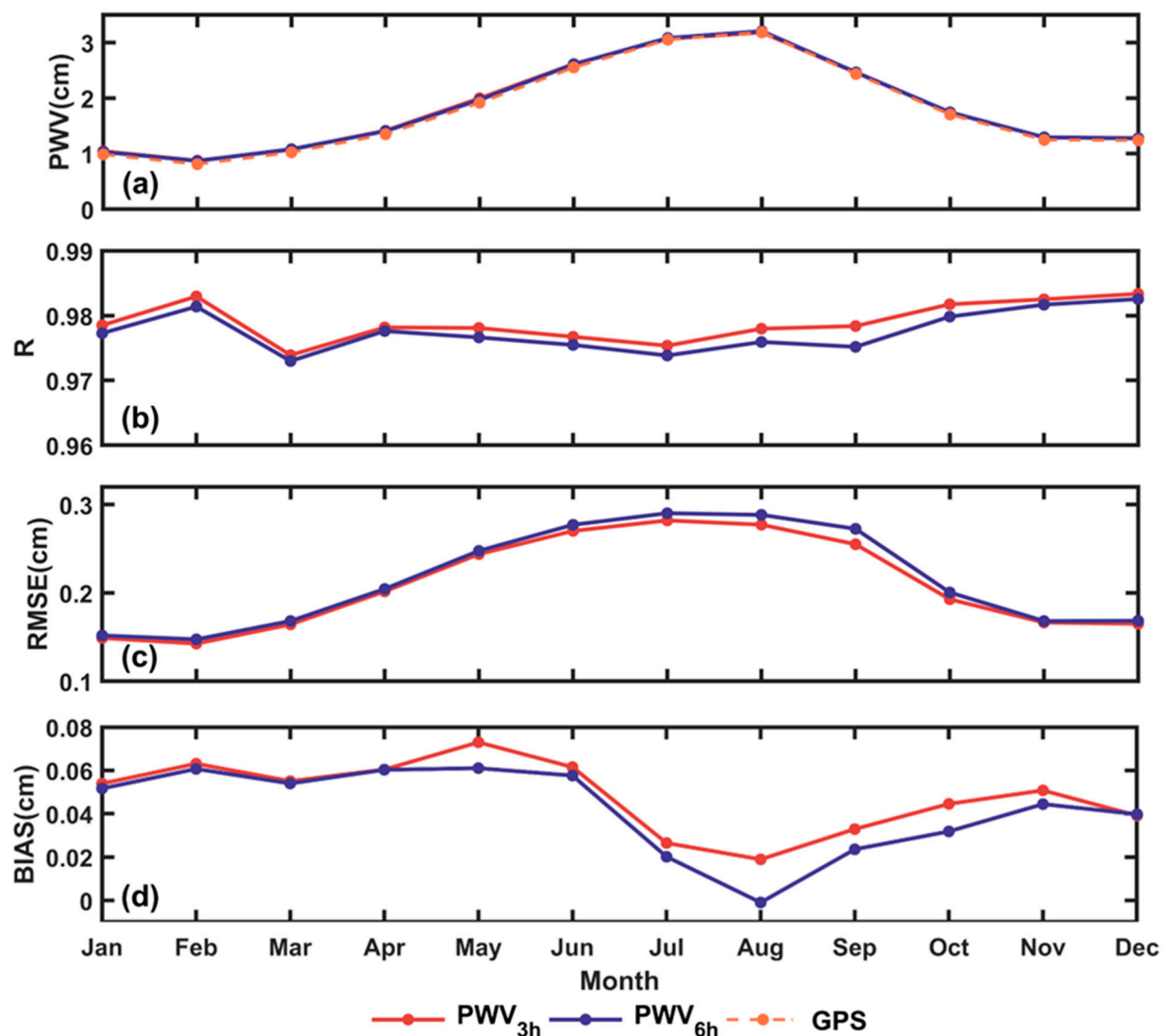


Figure 8. Time series of monthly mean statistics for GPS PWV and GFS PWV (PWV_{3h} and PWV_{6h}) (a) and correlation coefficient (R) (b), root mean square error (RMSE) (c), and bias (d) between GFS PWV (PWV_{3h} and PWV_{6h}) and GPS PWV.

Overall, the three statistical indicators exhibited diurnal variations. At night, the R is lower, and the RMSE and bias are higher, while the opposite can be observed during the daytime. Specifically, there is a high correlation between the eight moments of the 3 h forecast and the 6 h forecast, with Rs all greater than 0.98, indicating a positive correlation between GFS and GPS PWV values. Throughout the day, the maximum R (~0.988) and minimum RMSE (~0.208 cm) occur around UTC 15 in the 3 h forecast (corresponding to a local time range of 7:00 to 10:00 in the validation region). This may be attributed to the relative stability of the atmosphere during this period, leading to slower changes in water vapor content. Subsequently, during the local time range of 13:00 to 16:00 (21 UTC), the increased evaporation in the afternoon likely caused a rapid rise in water vapor content, which the GFS struggled to fully capture, resulting in a slight decrease in the R and an increase in RMSE. Bias values are all positive, ranging from 0.018 cm to 0.063 cm, indicating a slight overestimation of PWV values at each time point. The smallest bias occurs around UTC 18 in the 6 h forecast (corresponding to a local time range of 10:00 to 13:00 in the validation region).

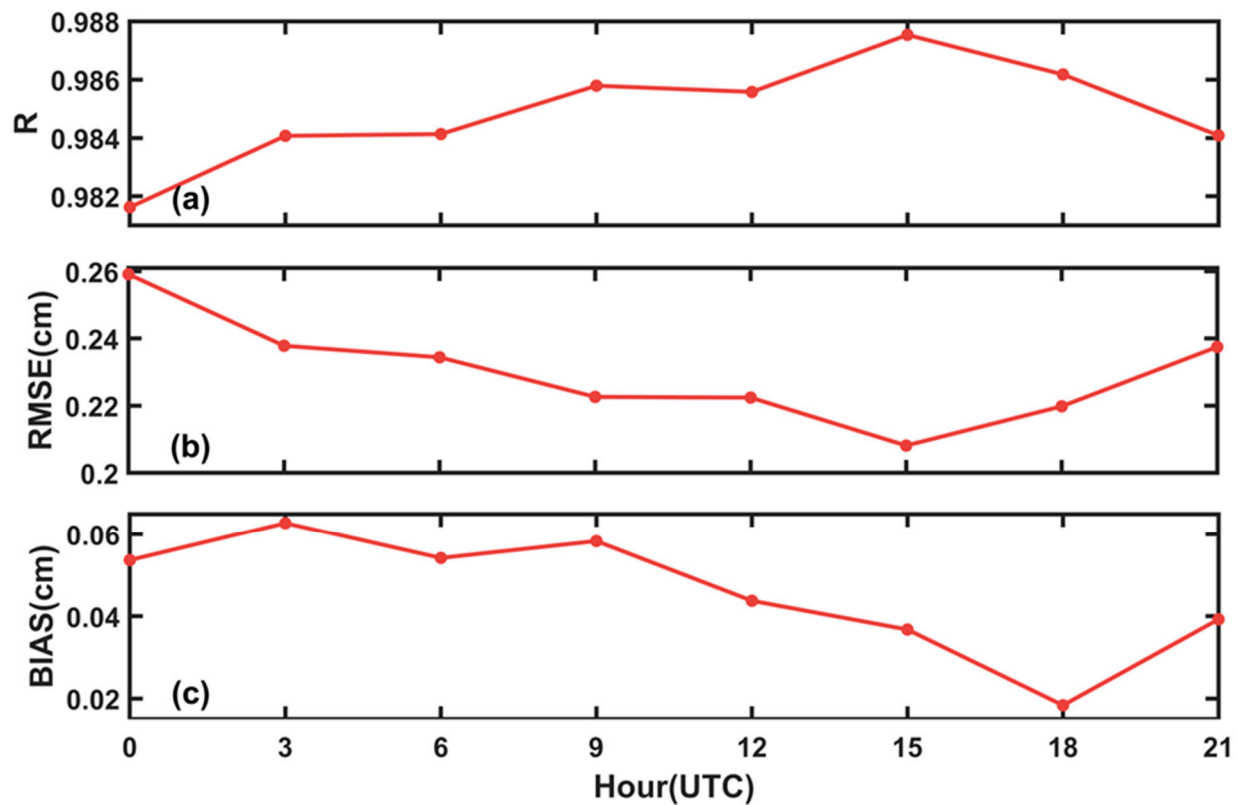


Figure 9. Diurnal variation in R (a), RMSE (b), and bias (c) statistical indicators for GFS PWV_{3h} (03, 09, 15, and 21 UTC) and GFS PWV_{6h} (06, 12, 18, and 00 UTC) compared with GPS.

3.4. Spatial Variations in GFS PWV Accuracy

A station-by-station analysis is conducted to assess the spatial performance of GFS PWV products based on the correlation between the GPS PWV and GFS PWV products. GFS PWV products exhibit different error characteristics across various regions. Figure 9a,b show that GFS PWV products have a high correlation with GPS PWV, with Rs exceeding 0.96 at 91.8% and 89.7% of the stations, respectively. By contrast, very few stations are below 0.94, with 3.9% and 4.2%, respectively. Further comparing the data, 76.7% of the stations show better consistency between PWV_{3h} and GPS PWV than between PWV_{6h} and GPS PWV. Table 3 indicates that elevation appears to have a greater impact on R than latitude. Specifically, lower-elevation stations exhibit higher R values, except for coastal stations, with most having R values > 0.98.

Table 3. Changes in statistics at different elevations.

3 h	DEM (km)	Number	R	RMSE (cm)	Bias (cm)
	−0.041~1	540,950	0.984	0.240	0.056
	1~2	102,807	0.978	0.169	0.023
	2~3	38,218	0.969	0.155	0.021
6 h	DEM (km)	Number	R	RMSE (cm)	Bias (cm)
	−0.041~1	518,701	0.983	0.248	0.049
	1~2	99,576	0.978	0.170	0.018
	2~3	36,799	0.969	0.154	0.021

The statistics show that the RMSE values of the stations are predominantly distributed between 0.1 cm and 0.3 cm, with only about 2.4% having an RMSE > 0.35 cm. The maximum RMSE values are 0.42 cm and 0.45 cm. During seasonal variations, the RMSE correlates with PWV. Figure 10c,d show spatial trends similar to those in Figure 1, demonstrating

that the RMSE is larger at low latitudes with larger mean PWV values. Thus, the RMSE is influenced by the spatial distribution of PWV.

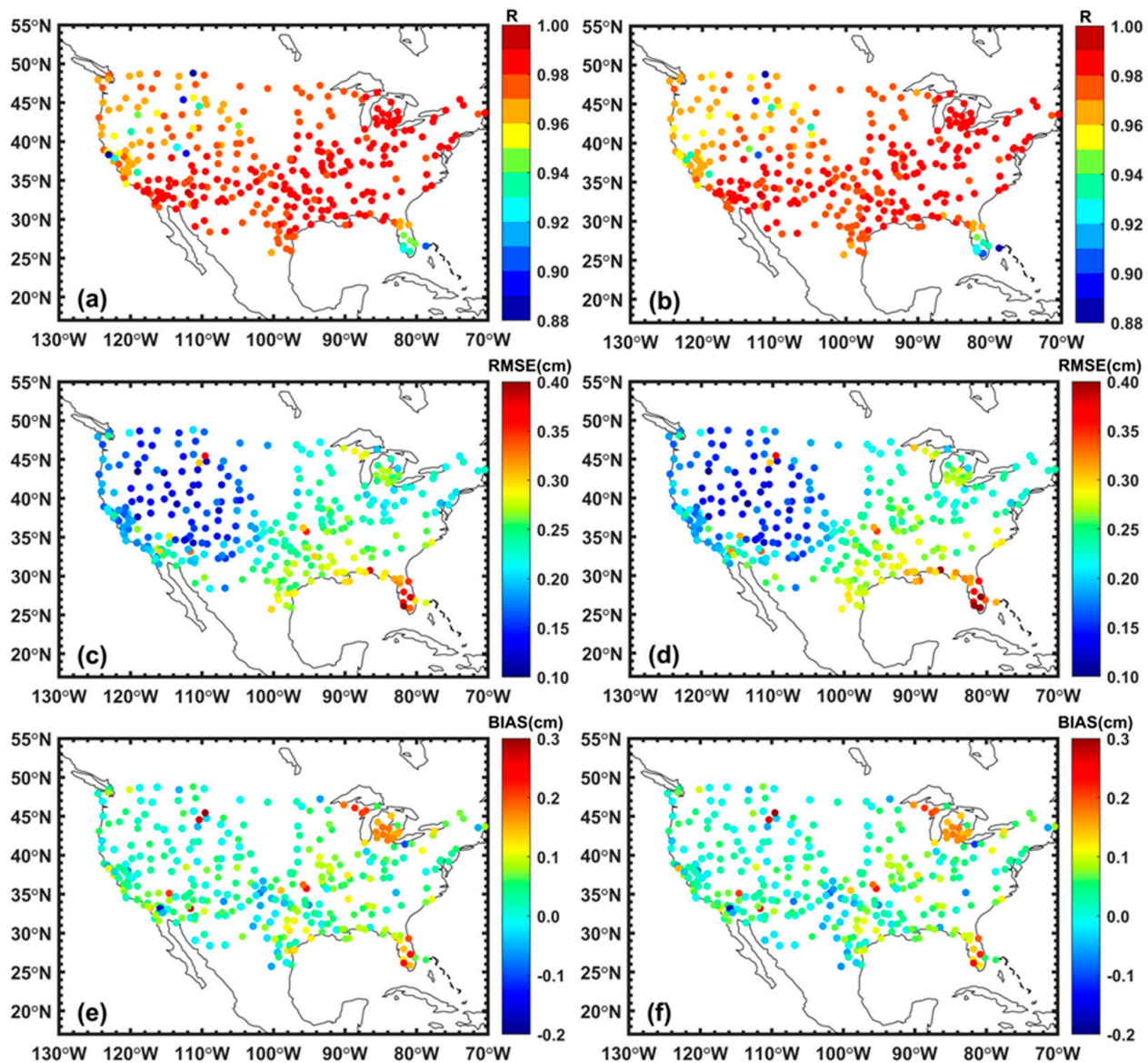


Figure 10. Spatial distribution of R (a,b), RMSE (c,d), and bias (e,f) between GPS PWV and PWV_{3h} and PWV_{6h} .

The bias results for each station are shown in Figure 10e,f. The PWV_{3h} and PWV_{6h} biases display a similar spatial distribution, with positive biases observed at several north-eastern stations, where water vapor biases range between 0.15 cm and 0.3 cm. The smaller bias values are mostly found at western coastal and high-elevation stations, ranging from -0.18 cm to 0.35 cm. In addition, the bias for approximately 84% of the stations falls within ± 0.1 cm, indicating that GFS PWV products maintain a high level of accuracy across most stations.

3.5. Accuracy Assessment of GFS PWV in Clear and Cloudy Conditions

The literature suggests that the total cloud cover over the United States is approximately 60% [55]. In our study area and time frame, PWV data (including PWV_{3h} and PWV_{6h}) were statistically analyzed, revealing that cloudy-sky conditions account for a larger proportion of the data (about 76.6%). Therefore, it is necessary to study the accuracy

of GFS PWV products under cloudy conditions. Figure 11 presents the statistical results for GFS PWV_{3h} and PWV_{6h} under clear and cloudy conditions. The analysis reveals that the R of the 3 h forecast is slightly higher than that of the 6 h forecast, both under clear and cloudy conditions. However, the correlation coefficients between GFS PWV and GPS PWV are all above 0.98. By contrast, the slope of the fitted curve for the data measured under cloudy conditions is approximately 0.98, which is closer to the 1:1 line. The R for GFS PWV under clear conditions is slightly lower. The bias shows minimal differences across weather conditions and forecast periods at 0.04 cm or 0.05 cm, which is negligible.

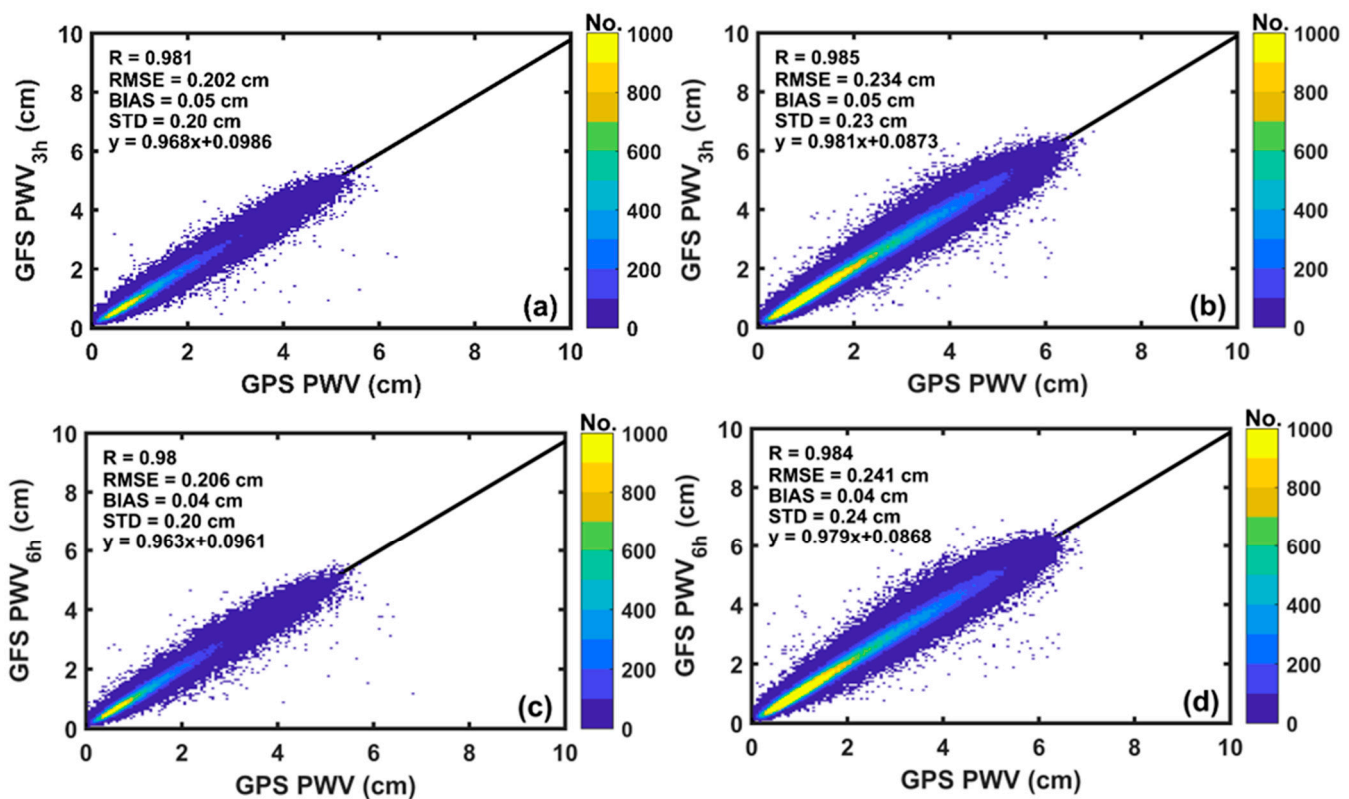


Figure 11. Two-dimensional histograms of GFS PWV_{3h} and GFS PWV_{6h} versus GPS PWV under clear-sky (a,c) and cloudy conditions (b,d).

Cloud cover is a manifestation of water vapor, and changes in clouds affect the distribution and content of water vapor in the atmosphere. Water vapor is relatively low in clear conditions, mostly less than 4 cm. PWV is larger in the presence of clouds and is concentrated in the 0–6 cm range, which may be one of the reasons why the RMSE values in the absence of clouds are smaller than the values in the presence of clouds.

Time series statistics for clear and cloudy skies over the continental United States are illustrated in Figure 12. The monthly average RMSE for clear-sky and cloudy-sky conditions fluctuates between 0.1 and 0.3 cm, showing a trend of increase followed by a decrease. Compared with clear skies, cloudy skies have larger RMSEs for each month. The minimum monthly RMSE for both conditions occurs in February (i.e., 0.129 cm and 0.151 cm, respectively), while the maximum occurs in July (i.e., 0.252 cm and 0.295 cm, respectively).

This indicates a seasonal variation in the RMSE of GFS PWV under clear and cloudy conditions. However, the bias does not show a clear seasonal pattern, with the lowest bias observed in August (i.e., less than 0.011 cm), which is negligible. In two-thirds of the months, the bias between GFS PWV and GPS PWV under cloudy conditions is smaller than under clear conditions. In addition, the R values for cloudy months show a more stable trend, ranging from 0.974 to 0.983, which may be due to the relative stability of meteorological factors when clouds are present.

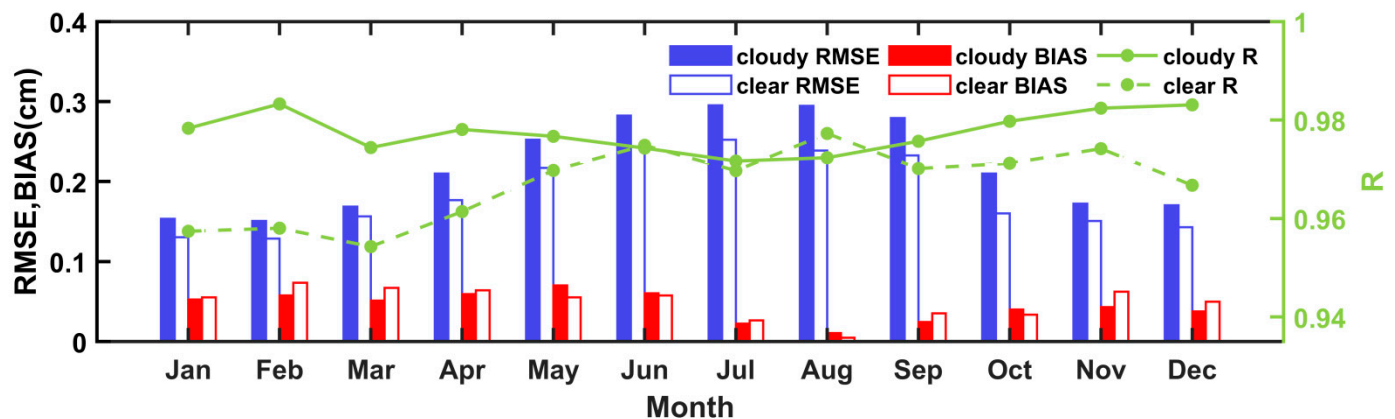


Figure 12. Time series of monthly R, RMSE, and bias between GFS PWV and GPS PWV under clear- and cloudy-sky conditions.

3.6. Accuracy Assessment in Extreme Weather Conditions

Irregular atmospheric motion leads to extreme weather, making it challenging to predict PWV under such conditions. At the same time, accurate PWV forecasting during extreme weather events is crucial for predicting these phenomena.

In this study, extreme weather events (i.e., hurricanes) in the American region were selected to validate GFS PWV. When hurricanes occur, unstable meteorological conditions and the complexity of atmospheric dynamics can damage most stations, resulting in data loss. Only three hurricanes (six stations in total) were selected for validation by screening 2021–2023 United States mainland hurricane events for data. The geographical locations of the hurricanes are marked in Figure 13.

The data from stations affected by the hurricanes were extracted to validate the GFS PWV. The results indicate that the GFS PWV product accuracy slightly decreases, showing differences compared with normal conditions. Generally, precipitation exceeds 4.5 cm during hurricanes, resulting in some PWV deviations. The linear correlation between GFS PWV and GPS PWV weakens. The R is slightly lower under extreme weather conditions than normal conditions, although some data still have an R greater than 0.9. Notably, the PWV range of Figure 13c is essentially maintained at a relatively stable level between 5 and 6 cm, resulting in an R of only 0.304. Specifically, at 03 UTC (Figure 13b,c) and 09 UTC (Figure 13b,d), the GFS PWV_{3h} exhibited a high degree of consistency with GPS observations, with minimal errors. However, at 12 UTC and 15 UTC, the GFS forecasts at certain stations showed significant deviations from the GPS observations (e.g., 15 UTC in Figure 13e). This discrepancy may be related to the GFS model's limited capability to capture atmospheric water vapor at specific time points.

This validation also calculates the RMSE, showing that the RMSE under extreme weather conditions is slightly larger than under normal conditions. The increased RMSE may be due to a combination of factors, including model performance under extreme conditions, complex atmospheric conditions during hurricanes, and limitations of the data itself. However, during Hurricane LAN on 30 September 2022, the maximum RMSE of PWV at the corresponding station (latitude: 34.345°N longitude: 77.875°W) only reached 0.529 cm. This indicates that the GFS model has relatively good overall prediction accuracy for water vapor products during extreme weather events.

In addition, we validated the GFS water vapor projections within 250 km of the hurricane reach by analyzing data from five hurricane crossings in 2021 and 2022 (shown in Figure 14). The data were categorized and analyzed based on forecast times (00 UTC, 06 UTC, 12 UTC, and 18 UTC) and forecast intervals (3 h and 6 h). Overall, the GFS PWV_{3h} and PWV_{6h} forecasts demonstrated a generally good correlation with GPS observations. Specifically, the forecasts at 00 UTC showed the most concentrated data points with minimal errors for both 3 h and 6 h intervals. The 3 h forecast at 06 UTC had a relative error of less

than 10% in most cases. However, the forecast accuracy at 18 UTC was relatively lower, with a more scattered distribution. This may be related to the corresponding local time of 10:00, when surface temperatures begin to rise, and boundary layer convection intensifies. The 3 h forecast exhibited slightly lower errors compared to the 6 h forecast, with RMSE values of 0.318 cm and 0.322 cm, respectively. The variability in errors over shorter time scales may be attributed to the limitations of the GFS model's physical parameterization in accurately predicting localized water vapor changes. As the forecast interval increases, the model's ability to capture rapid observed changes diminishes, a situation that may be more pronounced during hurricane events.

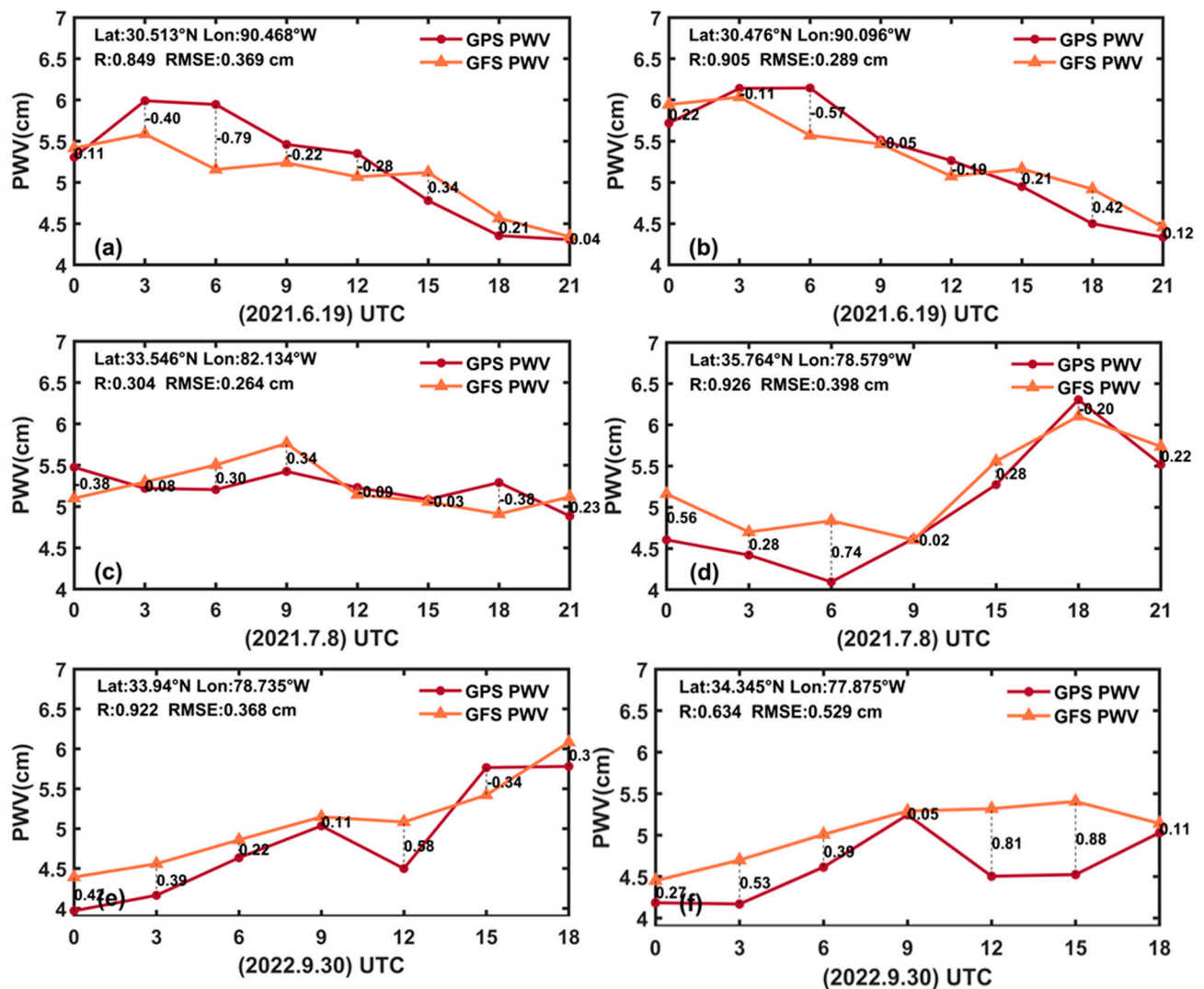


Figure 13. Daily variation in water vapor values and their statistical indicators at different stations for three different hurricanes: CLAUDETTE (a,b), ELSA (c,d), and LAN (e,f). Three-hour forecasts are at 03, 09, 15, and 21 UTC. Six-hour forecasts are at 06, 12, 18, and 00 UTC.

3.7. Correcting GFS PWV

The preceding analysis reveals a correlation between the statistical indicators of the GFS PWV data and the spatial and temporal information. Specifically, the RMSE is higher in summer than in winter, and both the R and RMSE decrease with increasing elevation and latitude. This pattern suggests the possibility of correcting the PWV data. We plan to incorporate geographical location and time information into the model input to enhance accuracy and reliability. Information such as geographic location and JD was added to the model inputs to hopefully improve the accuracy and reliability of GFS predictions.

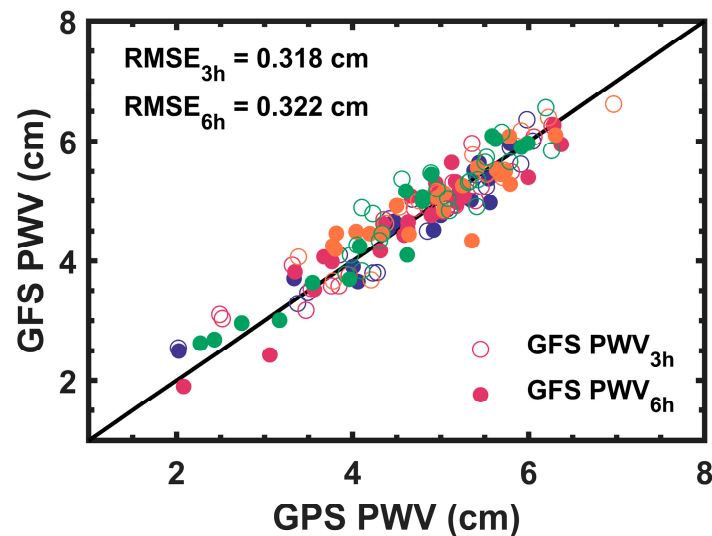


Figure 14. The validation results of GPS station data and GFS PWV_{3h} and PWV_{6h} forecasts within a 250 km radius of the hurricane impact zone during the passage of five hurricanes (CLAUDETTE, ELSA, FRED, IDA, and LAN) in 2021 and 2022. Different colors represent forecast times of 00 UTC (blue), 06 UTC (purple), 12 UTC (orange), and 18 UTC (green), with hollow and solid markers corresponding to PWV_{3h} and PWV_{6h} forecast results, respectively.

GFS PWV data from the continental United States, JD, geographical information (i.e., latitude, longitude, and elevation), and GPS PWV datasets were used to train and validate the BP neural network and RF correction models in our study. We selected odd data rows as training data and even data rows as validation data, building 12 revised models from January to December. The results are shown in Figure 15. The RMSE results indicate improvements in PWV correction for both models. After model correction, the water vapor prediction accuracy improves, and the RMSE decreases. The data in the figure show that the clearest improvement in RMSE occurs in February, with the neural network and random forest revisions resulting in RMSE reductions of 18.82% and 21.79%, respectively. However, the accuracy of both methods is relatively poor in August and September, which may be related to the water vapor value.

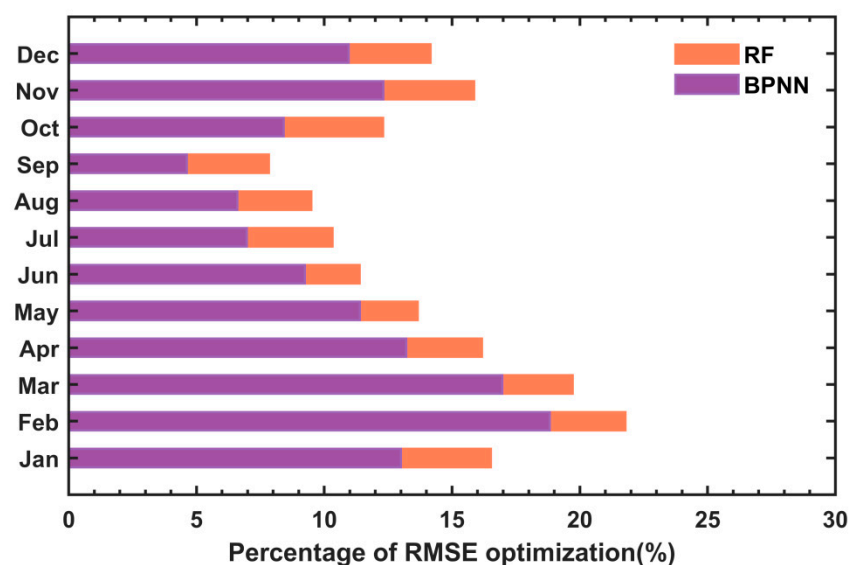


Figure 15. Monthly time series of RMSE-optimized percentages after BP neural network and RF revision.

The optimization percentages are the lowest in September, with improvements of 4.62% and 7.83%, respectively. The data corrected by the RF model show an RMSE improvement of over 10% in ten months. Overall, the RF model outperforms the neural network model, as the RMSE optimization percentages for each month are greater with the RF model than with the neural network model. This suggests that the RF is more effective than the neural network in water vapor revision.

4. Discussion

PWV is a key parameter in meteorological and climate research. Currently, most studies focus on other meteorological factors, such as temperature and humidity, and there are few validation studies on GFS PWV. Evaluating the accuracy and reliability of GFS PWV products is especially necessary owing to the important role of water vapor in weather forecasting and climate research, especially in short-term weather forecasting and solar energy applications.

In numerical forecasting, it is common for forecast errors to increase over time. However, the increasing magnitude of and differences in error across different time scales require detailed assessment and research. This is essential for improving forecast models, optimizing prediction methods, and practical applications. In this study, we evaluated the three-hour and six-hour GFS water vapor forecast products.

To improve the accuracy of GFS PWV data, we used machine learning methods for correction. By considering the influencing factors during validation, we constructed a multivariate regression model to correct the GFS PWV data. The error between the corrected GFS data and GPS data was significantly reduced, meaningfully improving the accuracy of atmospheric water vapor content estimation. In the future, the spatial and temporal continuity of GFS water vapor can be further improved by using a multi-source data fusion method that combines GPS data with data from radio-sounding techniques and microwave radiometers.

Water vapor is closely related to atmospheric instability. An environment with sufficient water vapor is often accompanied by high convective available potential energy. Accurately monitoring and forecasting atmospheric water vapor content can effectively predict severe convective weather. During our extreme weather validation, many GPS station data were missing owing to hurricane impacts. In the future, we can continue this work if more data become available. In addition, future research could focus on PWV products under other severe weather conditions, such as hailstorms and heavy rainfall. Additionally, this study analyzed the differences in GFS PWV products under clear and cloudy conditions. Future research could investigate error variations under different cloud types and forecast accuracy under varying precipitation conditions.

5. Conclusions

In this study, the accuracy of GFS PWV products—including PWV_{3h} and PWV_{6h} under various temporal, spatial, and special weather conditions—was validated using PWV data from 330 GPS stations in the continental United States (25–55°N, 70–130°W) collected from April 2021 to June 2023. Based on the spatial and temporal dependence of the statistical indicators, JD and geographic information (i.e., latitude, longitude, and elevation) were used as inputs in the model. BP neural networks and RF were used to construct the model and correct GFS PWV.

GFS PWV products exhibited high-level consistency with GPS PWV, with an R above 0.98. The performance of PWV_{3h} is slightly better than that of PWV_{6h} , with an RMSE of only 0.227 cm for PWV_{3h} . The GFS PWV products show a slight overestimation, which is more pronounced for the monthly variations. Additionally, a 5 mm PWV bin error statistical analysis indicates that this overestimation and underestimation are related to the amount of PWV. Specifically, GFS PWV is underestimated relative to GPS PWV for PWV values greater than 4 cm, and the underestimation becomes more pronounced as the water

vapor value increases. Compared with PWV_{3h} , the PWV_{6h} difference with water vapor clearly changes, and the dispersion of error values is larger.

Our evaluation of seasonal variations indicates that the RMSE exhibits strong seasonal dependency, with higher RMSE values in summer than in winter. Monthly R values are relatively low in summer and high in winter. There is no clear seasonal variation in monthly bias, which is only affected during the summer months owing to higher precipitation, with the PWV_{6h} product predicting a negative bias value for PWV in August. Overall, heavy rainfall in summer increases the difficulty of accurately predicting GFS PWV. In terms of diurnal variation, the consistency during the day is better than at night, with slightly higher RMSE and bias values observed at night. This suggests that GFS predictions are more accurate during the day.

Owing to stable meteorological factors in cloudy weather, the R of GFS PWV under cloudy conditions is slightly better than under clear conditions, with a fluctuation of only 0.009 under cloudy conditions compared with 0.023 under clear conditions. PWV values are relatively low in clear conditions (i.e., mostly less than 6 cm), whereas they are higher under cloudy conditions, resulting in an RMSE of 0.238 cm under cloudy conditions compared with 0.204 cm under clear conditions. This study validated GFS PWV products under extreme weather events. During hurricanes, characterized by unstable meteorological conditions and high precipitation, relatively low R values and higher RMSE values can be observed. However, the maximum R still exceeds 0.9, and the RMSE is around 0.5 cm. This indicates that GFS PWV products maintain a certain level of accuracy during hurricane events.

Spatially, elevation, latitude, and longitude variations impact statistical indicators. The R and RMSE values are high in the east ($R > 0.98$ and $RMSE > 0.2$ cm at most stations) and low in the west, with elevation playing a more important role than latitude in the differences observed between the eastern and western regions. Temporally and spatially, the RMSE is influenced by water vapor. Comparing the spatial variation in the RMSE with the annual average PWV map reveals that the spatial variability of the RMSE is similar to that of annual mean water vapor, which varies with latitude. Overall, the RMSE is larger in low-latitude regions with higher average PWV values and smaller in high-latitude regions.

Finally, the spatiotemporal analysis indicates that statistical indicators exhibit certain regularities with time, elevation, and latitude/longitude, demonstrating the feasibility of revising GFS PWV products. In this study, water vapor estimation was modeled using multilayer BP neural networks and the RF model. Experimental validation demonstrated that the corrected water vapor results improved, with the RF correction model exhibiting clear enhancements. The overall RMSE accuracy improved by 12.08%, with the RMSE in February decreasing by 21.79%.

Author Contributions: Conceptualization, H.-L.L.; methodology, H.-L.L.; software, X.-Q.Z.; validation, H.-L.L. and X.-Q.Z.; formal analysis, Y.-Y.Z.; investigation, Y.-Y.Z.; resources, Y.-Y.Z. and M.-Z.D.; data curation, H.-L.L.; writing—original draft preparation, H.-L.L. and X.-Q.Z.; writing—review and editing, H.-L.L. and X.-Q.Z.; visualization, S.-L.Z.; supervision, B.C.; project administration, H.-L.L.; funding acquisition, H.-L.L. All authors have read and agreed to the published version of the manuscript.

Funding: This study was supported by the National Natural Science Foundation of China, grant numbers 42030107 and 42175150.

Data Availability Statement: GPS PWV products are available at <https://www.unidata.ucar.edu/data/suominet/> (last access: 29 June 2024). GFS PWV products are available at <https://www.nco.ncep.noaa.gov/pmb/products/gfs> (last access: 29 June 2024). United States hurricane data are available at <https://www.nhc.noaa.gov/> (last access: 29 June 2024).

Acknowledgments: The authors thank the GPS and GFS teams at the National Hurricane Center and the Central Pacific Hurricane Center for providing the data used in this work.

Conflicts of Interest: The authors declare no conflicts of interest.

References

- Shi, F.L.; Xin, J.Y.; Yang, L.K.; Cong, Z.Y.; Liu, R.X.; Ma, Y.N.; Wang, Y.S.; Lu, X.F.; Zhao, L. The first validation of the precipitable water vapor of multisensor satellites over the typical regions in China. *Remote Sens. Environ.* **2018**, *206*, 107–122. [\[CrossRef\]](#)
- Senkal, O.; Yildiz, B.Y.; Sahin, M.; Pestemalci, V. Precipitable water modelling using artificial neural network in Cukurova region. *Environ. Monit. Assess.* **2012**, *184*, 141–147. [\[CrossRef\]](#) [\[PubMed\]](#)
- Gong, Y.Z.; Liu, Z.Z. Evaluating the Accuracy of Jason-3 Water Vapor Product Using PWV Data from Global Radiosonde and GNSS Stations. *IEEE Trans. Geosci. Remote Sens.* **2021**, *59*, 4008–4017. [\[CrossRef\]](#)
- Zveryaev, I.I.; Allan, R.P. Water vapor variability in the tropics and its links to dynamics and precipitation. *J. Geophys. Res. Atmos.* **2005**, *110*, 17. [\[CrossRef\]](#)
- Martins, V.S.; Lyapustin, A.; Wang, Y.J.; Giles, D.M.; Smirnov, A.; Slutsker, I.; Korkin, S. Global validation of columnar water vapor derived from EOS MODIS-MAIAC algorithm against the ground-based AERONET observations. *Atmos. Res.* **2019**, *225*, 181–192. [\[CrossRef\]](#)
- Vey, S.; Dietrich, R.; Rülke, A.; Fritsche, M.; Steigenberger, P.; Rothacher, M. Validation of Precipitable Water Vapor within the NCEP/DOE Reanalysis Using Global GPS Observations from One Decade. *J. Clim.* **2010**, *23*, 1675–1695. [\[CrossRef\]](#)
- Schneider, M.; Romero, P.M.; Hase, F.; Blumenstock, T.; Cuevas, E.; Ramos, R. Continuous quality assessment of atmospheric water vapour measurement techniques: FTIR, Cimel, MFRSR, GPS, and Vaisala RS92. *Atmos. Meas. Tech.* **2010**, *3*, 323–338. [\[CrossRef\]](#)
- Alshawaf, F.; Fuhrmann, T.; Knöpfler, A.; Luo, X.; Mayer, M.; Hinz, S.; Heck, B. Accurate Estimation of Atmospheric Water Vapor Using GNSS Observations and Surface Meteorological Data. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 3764–3771. [\[CrossRef\]](#)
- Tobin, D.C.; Revercomb, H.E.; Knuteson, R.O.; Lesht, B.M.; Strow, L.L.; Hannon, S.E.; Feltz, W.F.; Moy, L.A.; Fetzer, E.J.; Cress, T.S. Atmospheric Radiation Measurement site atmospheric state best estimates for Atmospheric Infrared Sounder temperature and water vapor retrieval validation. *J. Geophys. Res. Atmos.* **2006**, *111*, 18. [\[CrossRef\]](#)
- Mehta, S.; Singh, S.; Mitra, A.; Ghosh, S.K.; Raha, S. Diurnal Variation of the Lower Tropospheric Water Vapor Observed Using Microwave Radiometer Over Darjeeling (27.05°N, 88.26°E). *J. Indian Soc. Remote Sens.* **2019**, *47*, 619–628. [\[CrossRef\]](#)
- Chen, B.Y.; Liu, Z.Z. Global water vapor variability and trend from the latest 36year (1979 to 2014) data of ECMWF and NCEP reanalyses, radiosonde, GPS, and microwave satellite. *J. Geophys. Res. Atmos.* **2016**, *121*, 11442–11462. [\[CrossRef\]](#)
- Smith, W.L. Atmospheric soundings from satellites—False expectation or the key to improved weather prediction. *Q. J. R. Meteorol. Soc.* **2010**, *117*, 267–297.
- Wang, Y.Q.; Shi, J.C.; Wang, H.; Feng, W.L.; Wang, Y.J. Physical statistical algorithm for precipitable water vapor inversion on land surface based on multi-source remotely sensed data. *Sci. China Earth Sci.* **2015**, *58*, 2340–2352. [\[CrossRef\]](#)
- Wang, Y.Z.; Liu, H.L.; Zhang, Y.; Duan, M.Z.; Tang, S.H.; Deng, X.B. Validation of FY-4A AGRI layer precipitable water products using radiosonde data. *Atmos. Res.* **2021**, *253*, 14. [\[CrossRef\]](#)
- Gong, S.Q.; Hagan, D.F.T.; Wu, X.Y.; Wang, G.J. Spatio-temporal analysis of precipitable water vapour over northwest china utilizing MERSI/FY-3A products. *Int. J. Remote Sens.* **2018**, *39*, 3094–3110. [\[CrossRef\]](#)
- Gao, B.-C.; Kaufman, Y.J. Water vapor retrievals using Moderate Resolution Imaging Spectroradiometer (MODIS) near-infrared channels. *J. Geophys. Res. Atmos.* **2003**, *108*, 2002JD003023. [\[CrossRef\]](#)
- Cadeddu, M.P.; Liljegren, J.C.; Turner, D.D. The Atmospheric radiation measurement (ARM) program network of microwave radiometers: Instrumentation, data, and retrievals. *Atmos. Meas. Tech.* **2013**, *6*, 2359–2372. [\[CrossRef\]](#)
- Zhao, Q.Z.; Zhang, X.Y.; Wu, K.; Liu, Y.; Li, Z.F.; Shi, Y. Comprehensive Precipitable Water Vapor Retrieval and Application Platform Based on Various Water Vapor Detection Techniques. *Remote Sens.* **2022**, *14*, 2507. [\[CrossRef\]](#)
- Zhang, Y.L.; Cai, C.S.; Chen, B.Y.; Dai, W.J. Consistency Evaluation of Precipitable Water Vapor Derived from ERA5, ERA-Interim, GNSS, and Radiosondes Over China. *Radio Sci.* **2019**, *54*, 561–571. [\[CrossRef\]](#)
- Suparta, W.; Alhasa, K.M. Estimation of Atmospheric Water Vapor from ANFIS Technique and Its Validation with GPS Data. *J. Infotel* **2019**, *11*, 8. [\[CrossRef\]](#)
- Liu, H.L.; Tang, S.H.; Zhang, S.L.; Hu, J.Y. Evaluation of MODIS water vapour products over China using radiosonde data. *Int. J. Remote Sens.* **2015**, *36*, 680–690. [\[CrossRef\]](#)
- Tan, J.S.; Chen, B.Y.; Wang, W.; Yu, W.K.; Dai, W.J. Evaluating Precipitable Water Vapor Products from Fengyun-4A Meteorological Satellite Using Radiosonde, GNSS, and ERA5 Data. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 12. [\[CrossRef\]](#)
- González, A.; Expósito, F.J.; Pérez, J.C.; Díaz, J.P.; Taima, D. Verification of precipitable water vapour in high-resolution WRF simulations over a mountainous archipelago. *Q. J. R. Meteorol. Soc.* **2013**, *139*, 2119–2133. [\[CrossRef\]](#)
- Weng, F.; Zou, X.; Wang, X.; Yang, S.; Goldberg, M.D. Introduction to Suomi national polar-orbiting partnership advanced technology microwave sounder for numerical weather prediction and tropical cyclone applications. *J. Geophys. Res. Atmos.* **2012**, *117*, 14. [\[CrossRef\]](#)
- Song, D.S.; Grejner-Brzezinska, D.A. Remote sensing of atmospheric water vapor variation from GPS measurements during a severe weather event. *Earth Planets Space* **2009**, *61*, 1117–1125. [\[CrossRef\]](#)
- Pérez-Ramírez, D.; Whiteman, D.N.; Smirnov, A.; Lyamani, H.; Holben, B.N.; Pinker, R.; Andrade, M.; Alados-Arboledas, L. Evaluation of AERONET precipitable water vapor versus microwave radiometry, GPS, and radiosondes at ARM sites. *J. Geophys. Res. Atmos.* **2014**, *119*, 9596–9613. [\[CrossRef\]](#)

27. Vaquero-Martínez, J.; Antón, M.; de Galisteo, J.P.O.; Cachorro, V.E.; Costa, M.J.; Román, R.; Bennouna, Y.S. Validation of MODIS integrated water vapor product against reference GPS data at the Iberian Peninsula. *Int. J. Appl. Earth Obs. Geoinf.* **2017**, *63*, 214–221. [\[CrossRef\]](#)
28. Gui, K.; Che, H.Z.; Chen, Q.L.; Zeng, Z.L.; Liu, H.Z.; Wang, Y.Q.; Zheng, Y.; Sun, T.Z.; Liao, T.T.; Wang, H.; et al. Evaluation of radiosonde, MODIS-NIR-Clear, and AERONET precipitable water vapor using IGS ground-based GPS measurements over China. *Atmos. Res.* **2017**, *197*, 461–473. [\[CrossRef\]](#)
29. Zhang, W.X.; Lou, Y.D.; Haase, J.S.; Zhang, R.; Zheng, G.; Huang, J.F.; Shi, C.; Liu, J.N. The Use of Ground-Based GPS Precipitable Water Measurements over China to Assess Radiosonde and ERA-Interim Moisture Trends and Errors from 1999 to 2015. *J. Clim.* **2017**, *30*, 7643–7667. [\[CrossRef\]](#)
30. Vaquero-Martínez, J.; Antón, M.; de Galisteo, J.P.O.; Román, R.; Cachorro, V.E.; Mateos, D. Comparison of integrated water vapor from GNSS and radiosounding at four GRUAN stations. *Sci. Total Environ.* **2019**, *648*, 1639–1648. [\[CrossRef\]](#)
31. Dousa, J.; Bennitt, G.V. Estimation and evaluation of hourly updated global GPS Zenith Total Delays over ten months. *GPS Solut.* **2013**, *17*, 453–464. [\[CrossRef\]](#)
32. Zhao, Q.Z.; Yao, Y.B.; Yao, W.Q.; Zhang, S.B. GNSS-derived PWV and comparison with radiosonde and ECMWF ERA-Interim data over mainland China. *J. Atmos. Sol. Terr. Phys.* **2019**, *182*, 85–92. [\[CrossRef\]](#)
33. Li, J.M.; Qiang, W.W.; Zheng, C.W.; Su, B.; Razzak, F.; Wen, J.R.; Xiong, H. Modeling Multiple Views via Implicitly Preserving Global Consistency and Local Complementarity. *IEEE Trans. Knowl. Data Eng.* **2023**, *35*, 7220–7238. [\[CrossRef\]](#)
34. Liu, H.; Zhou, Q.; Zhang, S.; Deng, X. Estimation of Summer Air Temperature over China Using Himawari-8 AHI and Numerical Weather Prediction Data. *J. Infotel* **2019**, *2019*, 1–10. [\[CrossRef\]](#)
35. Kishkina, A.K.; Shestakov, N.V.; Bugaets, A.N.; Gonchukov, L.V.; Sokolov, O.V. Comparison of the Integral Water Vapor Content of the Atmosphere by Data of the Global Forecast System (GFS) and GNSS Observations (Primorski Krai, Russia). *Water Resour.* **2022**, *49*, 1082–1092. [\[CrossRef\]](#)
36. Roman, J.A.; Knuteson, R.O.; Ackerman, S.A.; Tobin, D.C.; Revercomb, H.E. Assessment of Regional Global Climate Model Water Vapor Bias and Trends Using Precipitable Water Vapor (PWV) Observations from a Network of Global Positioning Satellite (GPS) Receivers in the US Great Plains and Midwest. *J. Clim.* **2012**, *25*, 5471–5493. [\[CrossRef\]](#)
37. Roman, J.; Knuteson, R.; August, T.; Hultberg, T.; Ackerman, S.; Revercomb, H. A global assessment of NASA AIRS v6 and EUMETSAT IASI v6 precipitable water vapor using ground-based GPS SuomiNet stations. *J. Geophys. Res. Atmos.* **2016**, *121*, 8925–8948. [\[CrossRef\]](#)
38. Ware, R.H.; Fulker, D.W. SuomiNet: A Real-Time National GPS Network for Atmospheric Research and Education. *Bull. Am. Meteorol. Soc.* **2000**, *81*, 677–694. [\[CrossRef\]](#)
39. Srivastava, A. Application of GPS PWV for rainfall detection using ERA5 datasets over the Indian IGS locations. *J. Earth Syst. Sci.* **2024**, *133*, 16. [\[CrossRef\]](#)
40. Rocken, C.; Ware, R.; Van Hove, T.; Solheim, F.; Alber, C.; Johnson, J.; Bevis, M.; Businger, S. Sensing atmospheric water vapor with the global positioning system. *Geophys. Res. Lett.* **1993**, *20*, 2631–2634. [\[CrossRef\]](#)
41. Bevis, M.; Businger, S.; Herring, T.A.; Rocken, C.; Anthes, R.A.; Ware, R.H. GPS meteorology: Remote sensing of atmospheric water vapor using the global positioning system. *J. Geophys. Res.* **1992**, *97*, 15787–15801. [\[CrossRef\]](#)
42. Li, Z.; Muller, J.-P.; Cross, P. Comparison of precipitable water vapor derived from radiosonde, GPS, and Moderate-Resolution Imaging Spectroradiometer measurements. *J. Geophys. Res.* **2003**, *108*, 2003JD003372. [\[CrossRef\]](#)
43. Bevis, M.; Businger, S.; Chiswell, S.; Herring, T.A.; Ware, R.H. GPS Meteorology: Mapping Zenith Wet Delays onto Precipitable Water. *J. Appl. Meteor.* **1994**, *33*, 379–386. [\[CrossRef\]](#)
44. Yue, H.W.; Gebremichael, M.; Nourani, V. Evaluation of Global Forecast System (GFS) Medium-Range Precipitation Forecasts in the Nile River Basin. *J. Hydrometeorol.* **2022**, *23*, 101–116. [\[CrossRef\]](#)
45. Zhou, X.Q.; Juang, H.M.H. A model instability issue in the National Centers for Environmental Prediction Global Forecast System version 16 and potential solutions. *Geosci. Model Dev.* **2023**, *16*, 3263–3274. [\[CrossRef\]](#)
46. Chen, B.Y.; Yu, W.K.; Wang, W.; Zhang, Z.T.; Dai, W.J. A Global Assessment of Precipitable Water Vapor Derived From GNSS Zenith Tropospheric Delays With ERA5, NCEP FNL, and NCEP GFS Products. *Earth Space Sci.* **2021**, *8*, 22. [\[CrossRef\]](#)
47. Hamill, T.M.; Hagedorn, R.; Whitaker, J.S. Probabilistic forecast calibration using ECMWF and GFS ensemble reforecasts. Part II: Precipitation. *Mon. Weather Rev.* **2008**, *136*, 2620–2632. [\[CrossRef\]](#)
48. Wang, Y.; Yang, K.; Pan, Z.Y.; Qin, J.; Chen, D.L.; Lin, C.G.; Chen, Y.Y.; Tang, W.; Han, M.; Lu, N.; et al. Evaluation of Precipitable Water Vapor from Four Satellite Products and Four Reanalysis Datasets against GPS Measurements on the Southern Tibetan Plateau. *J. Clim.* **2017**, *30*, 5699–5713. [\[CrossRef\]](#)
49. Jiang, J.; Zhou, T.J.; Zhang, W.X. Evaluation of Satellite and Reanalysis Precipitable Water Vapor Data Sets Against Radiosonde Observations in Central Asia. *Earth Space Sci.* **2019**, *6*, 1129–1148. [\[CrossRef\]](#)
50. Leckner, B. The spectral distribution of solar radiation at the earth's surface—Elements of a model. *Sol. Energy* **1978**, *20*, 143–150. [\[CrossRef\]](#)
51. Zheng, D.Y.; Hu, W.S.; Wang, J.; Zhu, M.C. Research on regional zenith tropospheric delay based on neural network technology. *Surv. Rev.* **2015**, *47*, 286–295. [\[CrossRef\]](#)
52. Ding, M.H. A second generation of the neural network model for predicting weighted mean temperature. *GPS Solut.* **2020**, *24*, 6. [\[CrossRef\]](#)

53. Zhu, W.D.; Li, Y.Q.; Luan, K.F.; Qiu, Z.E.; He, N.Y.; Zhu, X.L.; Zou, Z.Y. Forest Canopy Height Retrieval and Analysis Using Random Forest Model with Multi-Source Remote Sensing Integration. *Sustainability* **2024**, *16*, 1735. [[CrossRef](#)]
54. Huang, Y.; Bao, Y.S.; Petropoulos, G.P.; Lu, Q.F.; Huo, Y.F.; Wang, F. Precipitation Estimation Using FY-4B/AGRI Satellite Data Based on Random Forest. *Remote Sens.* **2024**, *16*, 1267. [[CrossRef](#)]
55. Free, M.; Sun, B.M. Time-Varying Biases in US Total Cloud Cover Data. *J. Atmos. Ocean. Technol.* **2013**, *30*, 2838–2849. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.