remote sensing

MDPI

Technical Note

Accurate Rainfall Prediction Using GNSS PWV Based on Pre-Trained Transformer Model

Wenjie Yin [1], Chen Zhou [1,*], Yuan Tian [1], Hui Qiu [1], Wei Zhang [2], Hua Chen [3], Pan Liu [3], Qile Zhao [4], Jian Kong [5] and Yibin Yao [5]

[1] School of Earth and Space Science Technology, Wuhan University, Wuhan 430072, China; windsoryin@whu.edu.cn (W.Y.); tian_yuan@whu.edu.cn (Y.T.); huiqiu@whu.edu.cn (H.Q.)
[2] Department of Forecasting and Networking, China Meteorological Administration, Beijing 100081, China; zhangwei_nsmc@cma.gov.cn
[3] School of Water Resources and Hydropower Engineering, Wuhan University, Wuhan 430062, China; chua@whu.edu.cn (H.C.); liupan@whu.edu.cn (P.L.)
[4] GNSS Research Center, Wuhan University, Wuhan 430062, China; zhaoql@whu.edu.cn
[5] School of Geodesy and Geomatics, Wuhan University, Wuhan 430062, China; jkong@whu.edu.cn (J.K.); ybyao@whu.edu.cn (Y.Y.)
* Correspondence: chenzhou@whu.edu.cn

Abstract: With an increase in the intensity and frequency of extreme rainfall events, there is a pressing need for accurate rainfall nowcasting applications. In recent years, precipitable water vapor (PWV) data obtained from GNSS observations have been widely used in rainfall prediction. Unlike previous studies mainly focusing on rainfall occurrences, this study proposes a transformer-based model for hourly rainfall prediction, integrating the GNSS PWV and ERA5 meteorological data. The proposed model employs the ProbSparse self-attention to efficiently capture long-range dependencies in time series data, crucial for correlating historical PWV variations with rainfall events. Additionally, the adoption of the DILATE loss function better captures the structural and timing aspects of rainfall prediction. Furthermore, traditional rainfall prediction models are typically trained on datasets specific to one region, which limits their generalization ability due to regional meteorological differences and the scarcity of data in certain areas. Therefore, we adopt a pre-training and fine-tuning strategy using global datasets to mitigate data scarcity in newly deployed GNSS stations, enhancing model adaptability to local conditions. The evaluation results demonstrate satisfactory performance over other methods, with the fine-tuned model achieving an MSE = 3.954, DTW = 0.232, and TDI = 0.101. This approach shows great potential for real-time rainfall nowcasting in a local area, especially with limited data.

Keywords: rainfall prediction; deep learning; transfer learning; Global Navigation Satellite System (GNSS); precipitable water vapor (PWV)

check for updates

Academic Editor: Ismail Gultepe

Received: 31 March 2025
Revised: 2 June 2025
Accepted: 3 June 2025
Published: 12 June 2025

Citation: Yin, W.; Zhou, C.; Tian, Y.; Qiu, H.; Zhang, W.; Chen, H.; Liu, P.; Zhao, Q.; Kong, J.; Yao, Y. Accurate Rainfall Prediction Using GNSS PWV Based on Pre-Trained Transformer Model. *Remote Sens.* **2025**, *17*, 2023. https://doi.org/10.3390/rs17122023

Copyright: © 2025 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

Rainfall, a pivotal meteorological element in the hydrologic cycle, significantly contributes to the occurrences of natural hazards such as floods and landslides [1–4]. Given that, accurate prediction of rainfall is indispensable for disaster prediction and early warning, arousing great attraction for developing reliable and robust rainfall nowcasting methodologies [5–12].

The rainfall results from complex microphysical processes associated with various atmospheric factors, including water vapor, humidity, temperature, etc. Water vapor—one of the most essential elements in the atmosphere—is crucial in controlling the weather

patterns on a worldwide scale [13]. Studies have shown that an increase in the water vapor content could result in more frequent and intense extreme rainfall events in the context of global warming [14–17]. Therefore, accurate measurement of water vapor content is important for monitoring and predicting rainfall. However, traditional sensing techniques, including radiosondes, water vapor radiometers, and satellite-based methods, are unable to meet the increasing need for the continuous, accurate, and high-resolution retrieval of water vapor [18–21]. The Global Navigation Satellite System (GNSS) has a distinct edge over other remote sensing techniques for its higher spatiotemporal resolution and all-weather operation [22].

In GNSS meteorology, precipitable water vapor (PWV) refers to the amount of water vapor stored in a column of unit cross section from the ground to the top of the troposphere. From GNSS observations, the zenith tropospheric delay (ZTD) can be derived, which consists of the zenith hydrostatic delay (ZHD), predominantly caused by the dry part of the atmosphere, and the zenith wet delay (ZWD), caused by water vapor. PWV is retrieved from ZWD via a conversion factor based on an empirical equation [22,23].

Previous studies of GNSS-derived PWV have shown great potential for weather prediction applications and climate studies. In prior literature, strong correlations have been identified between the intense rainfall and PWV variations based on the observation and the formation mechanism of rainfall [6,9,17,24]. One key application is the assimilation of GNSS-derived PWV into the Numerical Weather Prediction process that can enhance short-range forecasts of precipitation and heavy rainfall events [25–27]. Furthermore, various models, including the threshold-based and machine learning approaches, have been developed to predict rainfall events by utilizing the PWV data and other meteorological information [5,6,8,12,28–32]. Threshold-based methods offer a straightforward and simple approach with relatively good prediction accuracy. However, the efficacy depends heavily on selecting the optimal thresholds. To enhance the performance and consistency of the threshold-based model, diverse approaches, such as empirical values [31], time-varied thresholds [8,33], and the percentile method [32,34], have been explored in the previous literature. Despite these efforts, threshold-based methods have the demerits of high false alarm rate (FAR) and dependency on the manually chosen thresholds, making it challenging to build an accurate and robust prediction method that can cover all cases of rainfall.

Although diverse methods have been employed to address the critical challenge of rainfall nowcasting, the intricate and dynamic microphysical processes of rainfall, along with the involvement of numerous meteorological parameters, make machine learning (ML) a promising approach to address this complexity. Liu et al. developed a backpropagation neural network model that attains a probability of detection (POD) greater than 96% with a FAR of approximately 40% [35]. Similarly, Benevides et al. used a nonlinear autoregressive neural network to predict rainfall intensity, achieving a 64% classification score for intense rain events with a 22% FAR [30]. However, most studies focus on predicting the rainfall occurrence rather than forecasting the rainfall time series directly. Zhao et al. presented a 1 h lead-time rainfall forecast model using a support vector machine algorithm with GNSS PWV and meteorological parameters, achieving an average root-mean-square error (RMSE) of approximately 1.35 mm/h [28].

However, current ML-based methods rely heavily on the established GNSS stations with extensive historical data, posing a significant obstacle for newly installed sensors lacking local training datasets. To address this gap, we employ a transformer-based model architecture combined with a pre-training and fine-tuning strategy. This method leverages multi-source inputs to achieve accurate quantitative rainfall estimation even with limited local data. Specifically, our study leverages the global network of International GNSS Service (IGS) stations and ERA5 reanalysis data to develop a pre-trained model.

By training on public datasets, the pre-trained model captures universal patterns linking PWV, atmospheric parameters, and rainfall dynamics. The pre-training and fine-tuning approach enables rapid fine-tuning with minimal local data, ensuring adaptability to new regions without compromising accuracy [36–38]. The primary contribution of this paper lies in proposing a state-of-the-art transformer-based rainfall forecast model that integrates multi-source data. Furthermore, we demonstrate that the pre-trained model, utilizing ERA5 data as the training dataset, can be effectively applied to real-time nowcasting without requiring extensive local historical datasets. This approach showcases its potential by enabling the rapid deployment and use of the GNSS rainfall prediction technique in regions where historical data is scarce and advanced nowcasting technologies, such as weather radar, are unavailable.

The rest of this paper is organized as follows. In Section 2, we present the description of datasets used in this study, the methodology of PWV retrieval, and the technical details of the model and evaluation metrics. Section 3 presents the results and comprehensive analysis of model performance using pre-training and fine-tuning datasets. The discussion and summary of this study are given in Sections 4 and 5.

2. Materials and Methods

2.1. ERA5 Dataset

ERA5-Land enhanced dataset is produced by the European Centre for Medium-Range Weather Forecasts (ECMWF), providing hourly high-resolution data of surface variables [39]. By numerically integrating the ECMWF land surface model at global scales, the ERA5-Land dataset is able to reproduce the land component of the ERA5 climate reanalysis with a finer spatial resolution of 9 km as opposed to 31 km for ERA5 [40]. Due to the continuity and global coverage of the ERA5-land dataset, we use hourly meteorological variables from ERA5-land data as the pre-training dataset, including temperature, pressure, and rainfall.

2.2. GNSS Dataset

GNSS data, acquired from networks of ground-based GNSS receivers, provide high-precision positioning information that is instrumental in deriving atmospheric parameters such as PWV. In this study, we utilize GNSS observations from two sources, corresponding to the pre-training and fine-tuning phases of model development.

For the pre-training phase, GNSS data are sourced from the IGS troposphere products. These products provide ZTD estimates at a temporal resolution of five minutes, derived from a global network of over 400 GNSS stations. The data are organized into daily files per site, offering a comprehensive and standardized dataset for capturing universal PWV patterns. We integrate ERA5 reanalysis data that provides hourly meteorological parameters, which are spatially and temporally interpolated to align with IGS station locations. We used the four nearest ERA5 grid points to interpolate data to each GNSS station's location. Cubic-spline method is used to interpolate the hourly ERA5 data to match the time step of GNSS observation. This multi-source dataset enables the pre-trained model to learn generalized relationships between PWV, atmospheric conditions, and rainfall across diverse geographic and climatic regions.

For the fine-tuning phase, GNSS observations are collected from a GNSS receiver integrated with co-located meteorological sensor, newly deployed in Wuhan, China (see Figure 1). These sensors provide real-time measurements of ZTD and meteorological parameters, such as temperature, pressure, and rainfall. The dataset spans from June 2023 to September 2024, covering two summer rainy seasons with continuous observations. The fine-tuning dataset is derived from local sensors, capturing the specific atmospheric and

climatic conditions of the region. This dataset enables us to assess the transferability of the pre-trained model and evaluate its accuracy in real-time nowcasting applications under region-specific rainfall patterns and PWV variations.

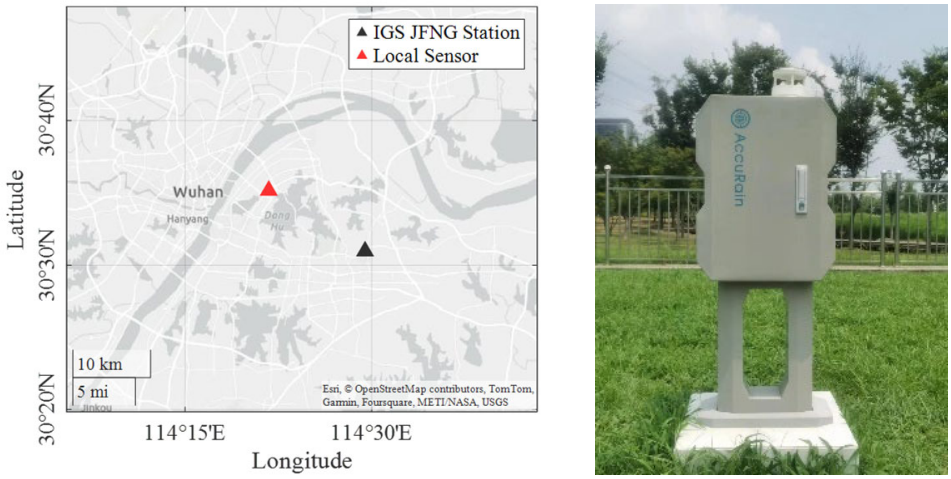

Figure 1. The photographs of the IGS JFNG station and deployed device integrated with GNSS receiver and meteorological sensor.

2.3. PWV Retrieval

GNSS signals experience atmospheric delays as they travel from satellites to ground-based receivers. The total delay is typically projected to the zenith direction, termed the ZTD, which comprises two components: approximately 10% ZWD and 90% ZHD, both derived from GNSS observations.

The ZHD is modeled using the empirical Saastamoinen formula [41]:

$$ZTD = ZHD + ZWD \tag{1}$$

$$\mathrm{ZHD} = \frac{0.0022768 \cdot \mathrm{P}_{re}}{1 - 0.0026 \cdot \cos(\phi) - 0.00028 \cdot \mathrm{H}} \tag{2}$$

where P_{re} represents the surface pressure (hPa), ϕ is the latitude, and H is the station height (km). The ZTD is calculated directly from GNSS data, and the ZWD is subsequently obtained by subtracting the modeled ZHD from the total ZTD.

After ZHD modeling, the wet component ZWD is computed by subtraction of ZHD from ZTD. Then, ZWD is converted to PWV using the following conversion equation [22,23]:

$$\mathrm{PWV} = \Pi \cdot \mathrm{ZWD} \tag{3}$$

$$\Pi = \frac{10^6}{\rho_w \cdot R_v \cdot \left(k_3 / T_\mathrm{m} + k_2'\right)} \tag{4}$$

where k_3 and k_2' are empirical constants with values of $(3.776 \pm 0.014) \times 10^6$ $\mathrm{K^2 \cdot hpa^{-1}}$ and 16.48 $\mathrm{K \cdot hpa^{-1}}$, respectively; R_V is the specific gas constant for water vapor; ρ is the liquid water density; and T_m is the weighted mean temperature in the atmospheric column. In practice, T_m is often computed from the commonly used empirical relations with the observed surface temperature [22,23].

We employed RTKLIB software v2.4.3, an open-source software package developed by the Tokyo University of Marine Science and Technology [42], for real-time GNSS processing using observations collected from a local GNSS receiver. The data processing with RTKLIB software v2.4.3 was configured as follows: (1) The ionospheric Iono-Free LC model and the Estimate ZTD were selected to correct ionospheric and tropospheric delay; (2) the elevation cut-off angle was set to 7°; (3) IGS real-time state space representation (SSR) products

were utilized; (4) four constellations including GPS, Beidou, Galileo, and GLONASS, were incorporated into the processing.

2.4. Model Architecture

Nowadays, transformer models have shown superior performance in many research areas. In time series forecasting problem, transformer-based models can capture long-range dependency better than recurrent neural network (RNN) models and achieve decent performance [43]. In this study, an optimized transformer-based model is adopted and tailored to the rainfall prediction problem. The basic structure of the model is shown in Figure 2.

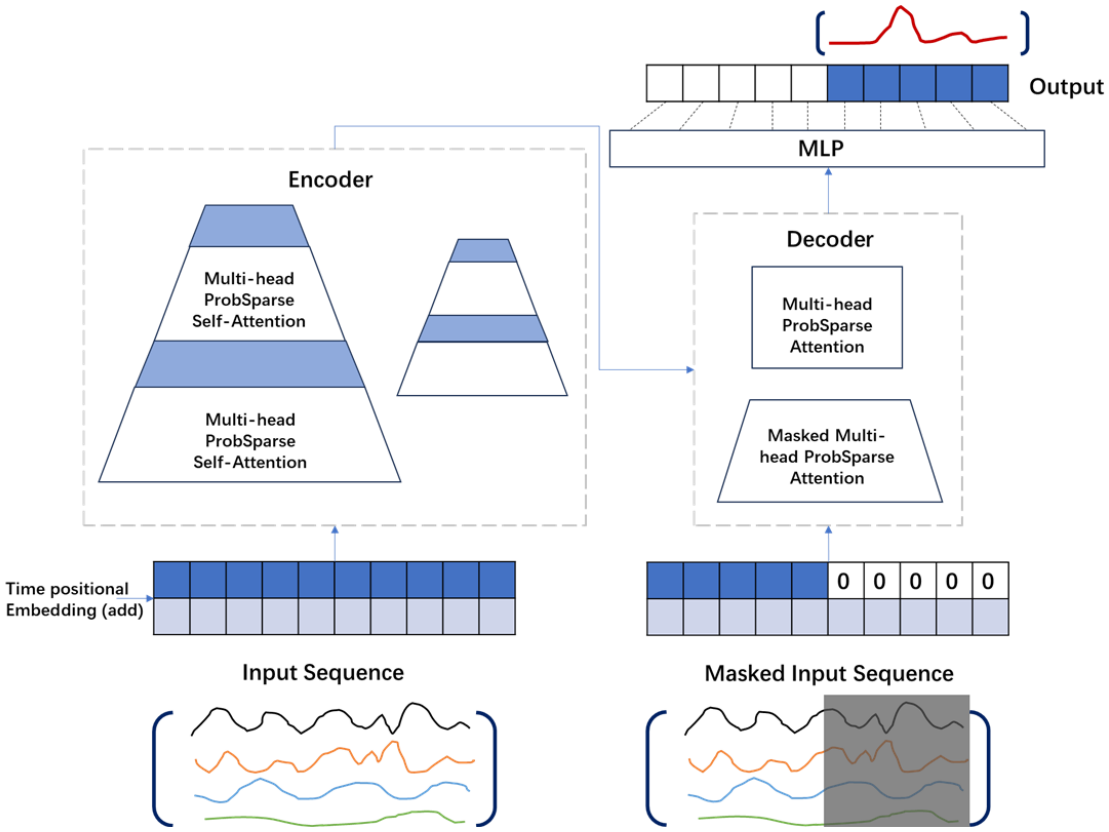

Figure 2. The architecture of the transformer-based prediction model.

2.4.1. Encoder–Decoder Architecture:

The transformer model adopts the encoder–decoder architecture and self-attention mechanism. The inputs of multi-variate time series added with the positional embeddings are received by the encoder to encode input representations, $\mathcal{X}^t = \left\{ \mathrm{x}_1^t, \ldots, \mathrm{x}_{L_x}^t \mid \mathrm{x}_i^t \in \mathbb{R}^{d_x} \right\}$, into hidden states, $\mathcal{H}^t = \left\{ \mathrm{h}_1^t, \ldots, \mathrm{h}_{L_h}^t \right\}$. The decoder computes a new hidden state from the received hidden states $\mathcal{H}^t$ and masked input sequence, from which the decoder decodes the output representations, $\mathcal{Y}^t = \left\{ \mathrm{y}_1^t, \ldots, \mathrm{y}_{L_y}^t \mid \mathrm{y}_i^t \in \mathbb{R}^{d_y} \right\}$, namely the prediction results.

2.4.2. Attention Mechanism

The canonical self-attention [44] is defined based on the tuple inputs, i.e., query, key, and value (Q, K, V), which performs the scaled dot-product as

$$\text{Attention}\ (\mathrm{Q}, \mathrm{K}, \mathrm{V}) = \text{Softmax}\left(\frac{\mathrm{QK}}{\sqrt{d}} \right) \mathrm{V}, \tag{5}$$

where $Q \in \mathbb{R}^{L_Q \times d}$, $K \in \mathbb{R}^{L_K \times d}$, $V \in \mathbb{R}^{L_V \times d}$, and d is the input dimension.

It requires the quadratic times dot-product computation and memory usage, which is the major drawback when enhancing prediction capacity. ProbSparse attention, a variant of the standard self-attention mechanism designed to reduce computational complexity for long sequences [43]. This is particularly advantageous for rainfall prediction, where sudden rainfall events may correlate with historical PWV variations in time series.

$$\text{ProbAtten}(Q, K, V) = \text{Softmax}\left(\frac{\overline{Q}K^\top}{\sqrt{d}}\right)V \tag{6}$$

where $\overline{Q}$ is a sparse matrix of the same size as q, and it only contains the top queries under the sparsity measurement of $\bar{M}(q_i, K)$. The ProbSparse self-attention reduces the calculation and memory usage. Under the multi-head perspective, this attention generates different sparse query–key pairs for each head, which avoids severe information loss in return.

$$\bar{M}(q_i, K) = \max_j \left\{\frac{q_i k_j^\top}{\sqrt{d}}\right\} - \frac{1}{L_K}\sum_{j=1}^{L_K}\frac{q_i k_j^\top}{\sqrt{d}} \tag{7}$$

The core of ProbSparse attention lies in the construction of a sparse query matrix, which contains only these selected active queries. For each query, q_i, in the original query matrix Q, a sparsity measurement, $\bar{M}(q_i, K)$, is computed to estimate its importance. We then select sparse Top-u from these measurements as $\overline{Q}$. Under the assumption $L_Q = L_K = L$, the total ProbSparse self-attention time complexity and space complexity are $\mathcal{O}(L \ln L)$.

2.4.3. Loss Function

A loss function is used to measure the model's output and update the model's parameters. Different loss functions penalize different aspects and can guide the model in different directions during the training process, which in return will affect the model's performance.

Mean squared error (MSE) is a commonly used and straightforward metric for calculating time series forecasting errors [28,43,45]. However, MSE fails to effectively distinguish between predictions with similar error values but different forecasting skills when dealing with sudden changes in time series. When it comes to rainfall prediction, the use of MSE will smoothen the final output because rainfall suddenly rises up from a value of zero. Therefore, we employed the DILATE loss to fit the rainfall prediction task, which is an optimized objective function for training deep learning networks in time series prediction [45]. It consists of a shape term and a temporal term, balanced by a hyperparameter. The shape term is based on the Dynamic Time Warping (DTW) [46], which assesses the structural shape similarity between the predicted and the actual time series. The temporal term is inspired by the Time Distortion Index (TDI) and is designed to penalize temporal distortions between the predicted and the actual time series [47].

2.5. Model Pre-Training and Fine-Tuning

Climate variability is observed globally, with significant differences in rainfall patterns across regions. Traditional rainfall prediction models are typically trained on datasets specific to one region, which limits their generalization ability due to regional meteorological differences and the scarcity of data in certain areas. To address these challenges, we adopt a pre-training and fine-tuning approach.

Pre-training and fine-tuning is a technique where a model, initially trained on a large-scale dataset, is subsequently fine-tuned on the target task [36–38,48]. In the context of rainfall prediction, we first pre-train the model using ERA5 meteorological data and GNSS PWV data provided by IGS stations distributed globally. This global pre-training

dataset encompasses the geographical area of interest worldwide, offering both accessibility and commendable data continuity. The pre-trained model is then fine-tuned on local meteorological data to adapt to the specific weather characteristics of the target region or station. To optimize performance, it is advisable to select an IGS station situated within the region of interest, paired with corresponding ERA5 data, for the pre-training phase. In this investigation, the JFNG station is employed to pre-train the model, which is then fine-tuned using the local dataset.

The steps for implementing this fine-tuning method are as follows: Obtain and pre-process the target dataset to match the input of the pre-trained model; load the pre-trained model weights except the output layer; fine-tune the model on the target dataset using a smaller learning rate through iterative training, which avoids large updates to the pre-trained weights and allows it to train the output layer from scratch. After fine-tuning, evaluate the model's performance on the test dataset to assess its accuracy and generalization ability.

2.6. Automated Machine Learning (AutoML) Method

In this study, we leverage Microsoft's AutoML toolbox, the FLAML library v2.3.4, to compare our proposed transformer-based model with traditional machine learning models, such as XGBoost, LightGBM, etc. [49–51]. FLAML is a lightweight Python library designed for efficient automation of machine learning workflows, which optimizes performance by automatically selecting and tuning a set of models to identify the best performer. The AutoML search strategy begins with a pool of learners, which employs Estimated Cost for Improvement (ECI)-based sampling and randomized direct search to optimize hyperparameters and sample size [49].

The training and test datasets for the AutoML method are identical to those used for the transformer-based model. However, due to the characteristics of traditional machine learning models, feature preprocessing is necessary. To prepare the input time series data, we performed feature engineering by incorporating lagged time features. The models were configured to predict rainfall for the next hour. MSE was selected as the loss function. Model performance was evaluated using cross-validation to ensure robust and reliable results.

2.7. Evaluation Metrics

To better evaluate the model's ability to predict rainfall in different aspects, three selected evaluation metrics are used. The descriptions of three metrics are summarized in Table 1. We use MSE, DTW, and TDI to evaluate the shape similarity and time delay between prediction and target time series.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2 \tag{8}$$

$$DTW^{\Delta}(y,z) := \min_{A \in \mathcal{A}_{n,m}} \langle A, \Delta(y,z) \rangle \tag{9}$$

$$TDI^{\Delta,\Omega}_{\text{dissim}}(y,z) := \langle A^*, \Omega_{\text{dissim}} \rangle \tag{10}$$

where n is the number of data points, Y_i is the actual value, and $\hat{Y}_i$ is the predicted value. Given two d-dimensional time series, $y \in \mathbb{R}^{d \times n}$ and $z \in \mathbb{R}^{d \times m}$ of lengths n and m, DTW looks for an optimal warping path represented by a binary matrix, $A \subset \{0,1\}^{n \times m}$, where $A_{ij} = 1$, if y_i is associated to z_j, and 0 otherwise [45]; $\Delta(y,z)$ is a $n \times m$ pairwise dissimilarity matrix whose general term is typically chosen as the Euclidean distance, $\Delta\left(y,z\right)_{ij} = \left\|y_i - z_j\right\|_2^2$; $A^* = \arg\min_{A \in \mathcal{A}_{n,m}} \langle A, \Delta(y,z) \rangle$ is the DTW optimal path [45]. $\Omega_{\text{dissim}} \in \mathbb{R}^{n \times m}$ is a matrix penalizing the association between y_i and z_j for $i \neq j$.

Table 1. Details of evaluation metrics.

Metric	Description
MSE	Mean square error
DTW	Dynamic Time Warping
TDI	Time Distortion Index

3. Results

After collecting the GNSS and ERA5 meteorological data, a transformer-based model is established, as described above. In this section, detailed descriptions of the experiment carried out in this paper are provided, which are summarized in Figure 3, mainly covering model pre-training and fine-tuning.

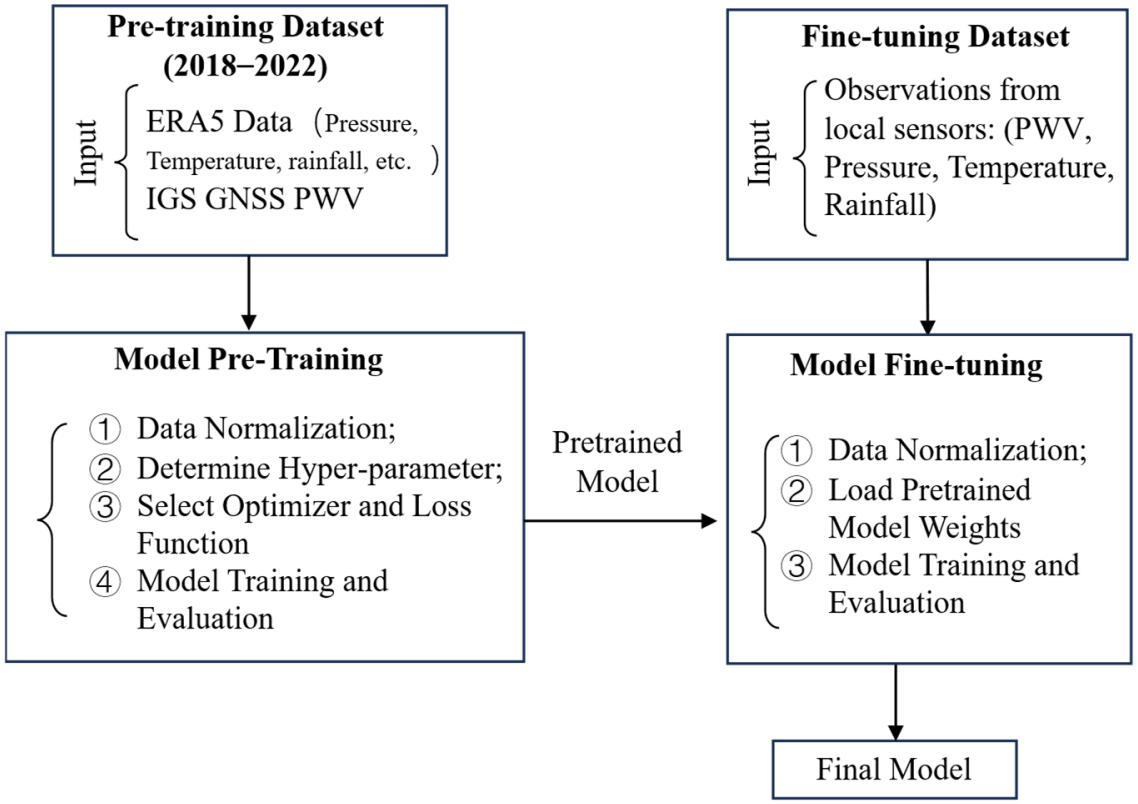

Figure 3. Flowchart of the experiment, including dataset curation, model pre-training, and fine-tuning.

3.1. Experiment Description

Datasets used in this paper mainly consist of the pre-training dataset and the fine-tuning dataset. The pre-training dataset is composed of collected historical PWV time series processed from GNSS observations and ERA5 meteorological data, including temperature, pressure, and rainfall. The total data period spans 5 years, from 2018 to 2022, which is divided into training, validation, and test datasets in a 7:1:2 ratio. In particular, a small section (~10%) of the dataset is used as a validation dataset to ensure the early stopping of the training procedure, aiming to prevent overfitting and improve the generalization ability of models. The fine-tuning dataset consists of observational data collected by the equipment integrated with the GNSS sensor and the meteorological sensor for a period of only 6 months.

Time series datasets often contain missing or invalid data, which can adversely affect model training if not properly handled. To mitigate this effect, we employ a two-fold approach: for short-term missing data points, we apply imputation techniques to estimate

and fill these gaps; for prolonged periods of missing data (e.g., an entire day), we opt for complete exclusion of those temporal segments rather than interpolation. This is because interpolation over extended gaps will generate artificial data that deviate from real patterns, potentially causing the model to learn erroneous knowledge. However, it also introduces another potential issue: direct deletion of segments can lead to temporal data discontinuity. To mitigate this, we have refined our batch data loading process. Previous batch data loading strategies rely on data imputation, which means they do not check data continuity. We now evaluate each time sequence individually during batch preparation. If the data exhibits time gaps or discontinuities, it is excluded from the batch. This strategy effectively mitigates the influence of temporal discontinuities on model training, thereby improving the generalizability and predictive accuracy of our machine learning framework.

During the pre-training stage, data normalization is implemented on the dataset, ensuring consistent scales across input features and improving model convergence. To determine the optimal model hyper-parameters, we designed six schemes to carry out comparison experiments, which are carefully constructed to evaluate the impact of hyper-parameter variations on model performance. The specific hyperparameter configurations and performance comparisons for each scheme are detailed in Tables 2 and 3. These configurations are also informed by the empirical value of best practices from prior research [43]. For instance, the encoder and decoder architectures are configured with three stacked layers, while the batch size is set to 64. The model is optimized using the Adam optimizer, with an initial learning rate of 5×10^{-5}, which is halved after each epoch to facilitate model convergence. The total number of epochs is 50 with proper early stopping. Then, the model is trained on the training dataset and evaluated on the test dataset using the evaluation metrics mentioned above. The optimal model is saved as the pre-trained model. The model is trained on a single 24 GB NVIDIA GeForce RTX 4090 GPU. Pre-training for one experiment typically takes 20 min, while fine-tuning is usually completed within 10 min due to the smaller dataset.

Table 2. Specific model hyper-parameters of 6 schemes.

Schemes \ Parameters	Sequence Length	Label Length	Batch Size	Learning Rate	Encoder Layers	Decoder Layers
Scheme 1	48	48	64	5×10^{-5}	3	3
Scheme 2	48	24	64	5×10^{-5}	3	3
Scheme 3	24	12	64	5×10^{-5}	3	3
Scheme 4	12	6	64	5×10^{-5}	3	3
Scheme 5	72	48	64	5×10^{-5}	3	3
Scheme 6	96	48	64	5×10^{-5}	3	3

Table 3. Model performance comparison of different schemes.

Metric \ Models	Scheme 1	Scheme 2	Scheme 3	Scheme 4	Scheme 5	Scheme 6
MSE	0.368	0.214	0.403	0.263	0.361	0.483
DTW	0.032	0.019	0.071	0.030	0.061	0.097
TDI	0.208	0.206	0.076	0.195	0.078	0.118

3.2. The Application in Real-Time Nowcasting Experiments

To investigate the potential of applying pre-trained models in real-time nowcasting, we transferred the pre-trained model to the scenario and evaluated its performance. We conducted a comparison experiment by deploying an integrated device equipped with a GNSS receiver and meteorological sensor in Wuhan from April to October in 2023 and 2024.

The PWV and meteorological data were collected in real-time processing, without post-processing techniques applied in the global GNSS dataset. The data collected in 2023 served as the training dataset, while the data from 2024 was used as the test dataset. Four models were compared: a pre-trained model directly applied without modification, the same pre-trained model fine-tuned with local data, a model trained from scratch using local data, and an AutoML approach using local data. As is shown in Table 4, the performance was assessed using three metrics, with lower values indicating better performance across all metrics.

Table 4. Comparison between models using different strategies.

Metric \ Models	Pre-Trained Model	Fine-Tuned Model	Model Trained from Scratch	AutoML
MSE	6.212	3.954	4.295	5.869
DTW	1.113	0.232	0.205	2.661
TDI	0.136	0.101	0.171	0.114

The pre-trained model, originally trained on a different dataset, was directly tested on the 2024 test dataset without any adjustments. This approach yielded the highest MSE of 6.212 among four models. These results suggest that the differences between the original training data and the Wuhan-specific dataset result in performance degradation, underscoring the necessity for fine-tuning. The AutoML approach shows an MSE of 5.869, a DTW of 2.661, and a TDI of 0.114. The MSE was higher than both the fine-tuned and scratch-trained models, and the DTW was the highest among all methods, indicating weaker performance in capturing temporal dynamics using simple model structures.

The fine-tuned pre-trained model outperformed all other methods in terms of MSE and TDI, highlighting the advantage of combining general knowledge from pre-training with local data adaptation. The model trained from scratch, while achieving the best DTW, fell short of the fine-tuned model in MSE and TDI, suggesting that pre-training provides a valuable starting point that fine-tuning can enhance. These results demonstrate that fine-tuning a pre-trained model with local GNSS and meteorological data is the most effective strategy for real-time nowcasting in this experiment, offering the best balance between accuracy and temporal consistency. Figure 4 presents a comparison between the measured hourly rainfall, shown as a black line, and the corresponding hourly rainfall predicted by the fine-tuned model, shown as a red line. Figure 4a illustrates this comparison across the entire fine-tuning test period, which spans from June to September 2024. This broader view captures multiple rainfall events of varying intensities, providing an overview of the model's predictive capability against the observed rainfall. Figure 4b offers a detailed perspective of the comparison on a selected shorter interval, from 10 July at 18:00 to 12 July at 00:00. This shows that the model successfully predicted the heavy rainfall event occurring around 00:00 on 11 July, where true rainfall approached 60 mm. While the prediction captured this event, its intensity was substantially less, peaking around 40 mm. In general, the proposed model, which involves fine-tuning a pre-trained model with local GNSS and meteorological data, proved to be the most effective strategy for real-time rainfall nowcasting, showing variable performance in matching the intensity and precise timing of heavy rainfall events.

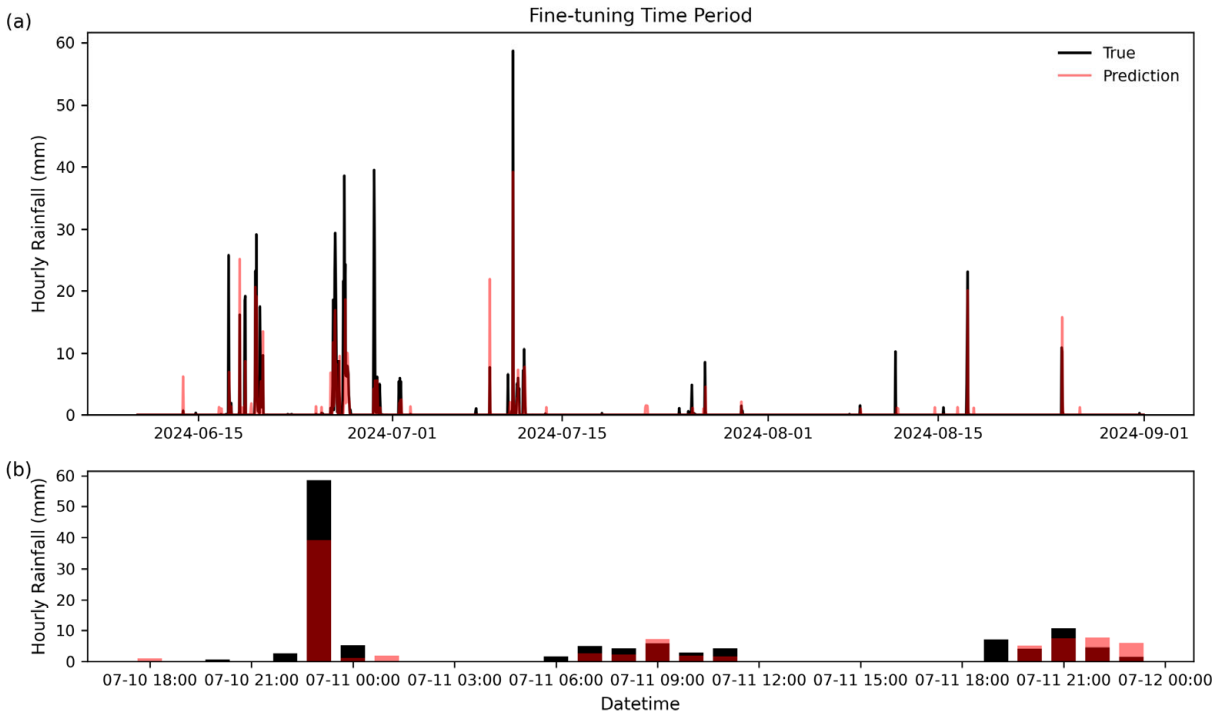

Figure 4. (**a**) Time series comparison of rainfall predictions from the fine-tuned model with observations over the entire fine-tuning period; (**b**) the same comparison for the selected period of 10 July to 12 July 2024.

4. Discussion

The results presented in Section 3 well demonstrate the satisfactory performance of the proposed transformer-based model for rainfall nowcasting, leveraging the GNSS PWV and ERA5 meteorological data. The transformer-based model employed in this study outperforms traditional approaches, as evidenced by the superior performance metrics in Tables 3 and 4. Unlike the RNN-based models, which struggle with vanishing gradients over long input sequences, the ProbSparse attention mechanism used in this model can effectively capture long-range dependencies in time series with reduced computational complexity [43]. This is particularly advantageous for rainfall prediction, where sudden occurrences of rainfall are possibly correlated with PWV variations in historical time series [6,17,29,52]. Unlike MSE loss, which smooths predictions and fails to differentiate subtle forecasting skills, the adoption of the DILATE loss function better captures the structural and timing aspects of rainfall events [45]. We also use RMSE, DTW, and TDI to quantitatively assess the model's prediction performance in both shape and temporal aspects. We carry out different schemes to determine the optimal pre-trained model. Scheme 2 demonstrates strong overall accuracy, evidenced by its lower RMSE and DTW scores, indicating better average point-wise accuracy. But its TDI score is less competitive than that of Scheme 3. This highlights a critical trade-off: Scheme 2, despite its general strengths, may be less suitable for applications where the precise timing of events—such as the exact onset critical for flood warnings—needs to be predicted. This detailed understanding of metric interdependencies is crucial for selecting the optimal model configuration based on specific nowcasting requirements and application priorities. These results also highlight the sensitivity of transformer models to sequence length, a known challenge due to the quadratic complexity of self-attention [43]. Furthermore, the fine-tuned model's performance (MSE = 3.954, DTW = 0.232, TDI = 0.101) significantly surpasses that of the pre-trained model and the AutoML model, underscoring the transformer's ability to adapt to local conditions while leveraging global patterns.

A key contribution of this study is the adopted pre-training and fine-tuning method to address one practical challenge in rainfall prediction—the limitation of data scarcity in the newly deployed GNSS stations restrains the adaptability of previous methods. Furthermore, the effective transfer of prediction models to new regions is an understated difficulty. To address these limitations, we strategically selected the globally available ERA5 dataset for pre-training our proposed transformer model, valued for its comprehensive coverage and wide array of meteorological variables [39]. By leveraging ERA5 in conjunction with PWV data from the global IGS data service, we can construct a foundational pre-trained model using readily accessible global datasets. This approach enables the model to be subsequently and efficiently fine-tuned for specific local scenarios. It mitigates the issue of sparse station data by pre-training the model on large datasets and fine-tuning it on smaller target datasets, reducing the computational resources required. This approach enhances the proposed model's generalization capability, allowing it to capture common features across different regions while improving its ability to adapt to local weather conditions [36,37,53].

In this study, multi-source meteorological data provide a robust foundation for rainfall prediction. PWV, as evidenced by previous research, is strongly correlated with rainfall occurrences [5,6,8,17,28,29,31]. The ERA5 dataset provides high-resolution temperature, pressure, and rainfall variables that complement the global GNSS observation stations. The integration of multi-source meteorological data enhances the model's ability to capture the complicated relationship among data, outperforming the single-parameter methods [5,28,35].

Despite these advancements, there remain several limitations. Discrepancy between the data distribution of the post-processing dataset and real-time observation could introduce a lot of biases, which need careful consideration [36,38,53]. The 6-month fine-tuning dataset may not fully represent seasonal variability, potentially reducing accuracy during unrepresented periods. In future works, we plan to deploy more observation devices and gather more local data across multiple seasons and years to validate the model's robustness and generalizability. Also, the proposed method needs extensive evaluation across different geographical zones to assess its adaptability to diverse climates.

5. Conclusions

A transformer-based rainfall prediction model using GNSS-derived PWV and common meteorological parameters is proposed in this study. Unlike previous GNSS-PWV studies that focus on the prediction of rainfall occurrences, the proposed model predicts rainfall time series quantitatively. Furthermore, this proposed transformer-based approach achieves a satisfactory performance when adapted to real-time rainfall nowcasting using newly deployed sensors, indicating that the use of a global dataset as a pre-training dataset can be successfully employed in real-time nowcasting. It demonstrates the potential of a pre-training and fine-tuning strategy to predict rainfall in areas where historical dataset is limited. The overall results suggest that deep learning networks can contribute to advancing the rainfall prediction performance using GNSS-PWV and the meteorological parameters. However, more observation data across multiple seasons and years are needed to validate the model's robustness and generalizability in future works.

Author Contributions: Conceptualization, W.Y. and C.Z.; methodology, W.Y. and Y.T.; software, Y.T. and W.Y.; validation, H.Q. and W.Z.; data curation, W.Y.; writing—original draft preparation, W.Y.; writing—review and editing, Q.Z., J.K., H.C. and C.Z.; visualization, W.Y.; supervision, P.L. and C.Z.; funding acquisition, H.C., P.L., Y.Y. and C.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Key Research and Development Program of China under Grant 2023YFC3209101.

Data Availability Statement: IGRA radiosonde dataset can be obtained from the National Oceanic and Atmospheric Administration (NOAA) (https://www.ncei.noaa.gov/products/weather-balloon/integrated-global-radiosonde-archive, accessed on 12 March 2025). ERA5 meteorological data are obtained from the European Center for Medium-Range Weather Forecasting (https://cds.climate.copernicus.eu/, accessed on 12 March 2025). IGS data can be accessed from IGS data center from Wuhan University (http://www.igs.gnsswhu.cn/index.php, accessed on 12 March 2025).

Acknowledgments: The authors would like to thank the Hong Kong Observatory for sharing the meteorological data. The authors would also like to thank the Wuhan CeduFuture company for providing data collected from local sensors (https://www.accurain.cn/home).

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

Abbreviations

The following abbreviations are used in this manuscript:

GNSS	Global Navigation Satellite System
PWV	Precipitable water vapor
ZTD	Zenith tropospheric delay
ZHD	Zenith hydrostatic delay
ZWD	Zenith wet delay
MSE	Root-mean-square error
DTW	Dynamic Time Warping
TDI	Time Distortion Index
ML	Machine learning
AutoML	Automated Machine Learning
IGS	International GNSS Service

References

1. Giorgi, F.; Raffaele, F.; Coppola, E. The Response of Precipitation Characteristics to Global Warming from Climate Projections. *Earth Syst. Dynam.* **2019**, *10*, 73–89. [CrossRef]
2. Deng, L.; Feng, J.; Zhao, Y.; Bao, X.; Huang, W.; Hu, H.; Duan, Y. The Remote Effect of Binary Typhoon Infa and Cempaka on the "21.7" Heavy Rainfall in Henan Province, China. *J. Geophys. Res. Atmos.* **2022**, *127*, e2021JD036260. [CrossRef]
3. Trenberth, K.E.; Dai, A.; Rasmussen, R.M.; Parsons, D.B. The Changing Character of Precipitation. *Bull. Am. Meteorol. Soc.* **2003**, *84*, 1205–1218. [CrossRef]
4. Breugem, A.J.; Wesseling, J.G.; Oostindie, K.; Ritsema, C.J. Meteorological Aspects of Heavy Precipitation in Relation to Floods–An Overview. *Earth-Sci. Rev.* **2020**, *204*, 103171. [CrossRef]
5. Liu, Y.; Yao, Y.; Zhao, Q. Real-Time Rainfall Nowcast Model by Combining CAPE and GNSS Observations. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 1–9. [CrossRef]
6. Li, H.; Wang, X.; Choy, S.; Jiang, C.; Wu, S.; Zhang, J.; Qiu, C.; Zhou, K.; Li, L.; Fu, E.; et al. Detecting Heavy Rainfall Using Anomaly-Based Percentile Thresholds of Predictors Derived from GNSS-PWV. *Atmos. Res.* **2022**, *265*, 105912. [CrossRef]
7. Zhang, Y.; Long, M.; Chen, K.; Xing, L.; Jin, R.; Jordan, M.I.; Wang, J. Skilful Nowcasting of Extreme Precipitation with NowcastNet. *Nature* **2023**, *619*, 526–532. [CrossRef] [PubMed]
8. Li, H.; Wang, X.; Wu, S.; Zhang, K.; Chen, X.; Zhang, J.; Qiu, C.; Zhang, S.; Li, L. An Improved Model for Detecting Heavy Precipitation Using GNSS-Derived Zenith Total Delay Measurements. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 5392–5405. [CrossRef]
9. Ramezani Ziarani, M.; Bookhagen, B.; Schmidt, T.; Wickert, J.; De La Torre, A.; Deng, Z.; Calori, A. A Model for the Relationship between Rainfall, GNSS-Derived Integrated Water Vapour, and CAPE in the Eastern Central Andes. *Remote Sens.* **2021**, *13*, 3788. [CrossRef]
10. Zhao, Q.; Zhang, X.; Wu, K.; Liu, Y.; Li, Z.; Shi, Y. Comprehensive Precipitable Water Vapor Retrieval and Application Platform Based on Various Water Vapor Detection Techniques. *Remote Sens.* **2022**, *14*, 2507. [CrossRef]
11. Graffigna, V.; Hernández-Pajares, M.; Azpilicueta, F.; Gende, M. Comprehensive Study on the Tropospheric Wet Delay and Horizontal Gradients during a Severe Weather Event. *Remote Sens.* **2022**, *14*, 888. [CrossRef]

12. Li, L.; Zhang, K.; Wu, S.; Li, H.; Wang, X.; Hu, A.; Li, W.; Fu, E.; Zhang, M.; Shen, Z. An Improved Method for Rainfall Forecast Based on GNSS-PWV. *Remote Sens.* **2022**, *14*, 4280. [CrossRef]
13. Schneider, T.; O'Gorman, P.A.; Levine, X.J. Water Vapor and the Dynamics of Climate Changes. *Rev. Geophys.* **2010**, *48*, RG3001. [CrossRef]
14. Du, H.; Donat, M.G.; Zong, S.; Alexander, L.V.; Manzanas, R.; Kruger, A.; Choi, G.; Salinger, J.; He, H.S.; Li, M.-H.; et al. Extreme Precipitation on Consecutive Days Occurs More Often in a Warming Climate. *Bull. Am. Meteorol. Soc.* **2022**, *103*, E1130–E1145. [CrossRef]
15. Valenzuela, R.A.; Garreaud, R.D. Extreme Daily Rainfall in Central-Southern Chile and Its Relationship with Low-Level Horizontal Water Vapor Fluxes. *J. Hydrometeorol.* **2019**, *20*, 1829–1850. [CrossRef]
16. Van Baelen, J.; Reverdy, M.; Tridon, F.; Labbouz, L.; Dick, G.; Bender, M.; Hagen, M. On the Relationship between Water Vapour Field Evolution and the Life Cycle of Precipitation Systems: Evolution of Water Vapour and Precipitation Systems. *Q.J.R. Meteorol. Soc.* **2011**, *137*, 204–223. [CrossRef]
17. Kunkel, K.E.; Stevens, S.E.; Stevens, L.E.; Karl, T.R. Observed Climatological Relationships of Extreme Daily Precipitation Events with Precipitable Water and Vertical Velocity in the Contiguous United States. *Geophys. Res. Lett.* **2020**, *47*, e2019GL086721. [CrossRef]
18. Liu, Z.; Wong, M.S.; Nichol, J.; Chan, P.W. A Multi-sensor Study of Water Vapour from Radiosonde, MODIS and AERONET: A Case Study of Hong Kong. *Intl. J. Climatol.* **2013**, *33*, 109–120. [CrossRef]
19. Niell, A.E.; Coster, A.J.; Solheim, F.S.; Mendes, V.B.; Toor, P.C.; Langley, R.B.; Upham, C.A. Comparison of Measurements of Atmospheric Wet Delay by Radiosonde, Water Vapor Radiometer, GPS, and VLBI. *Int. J. Atmos. Ocean. Technol.* **2001**, *18*, 830–850. [CrossRef]
20. Ferreira, A.P.; Nieto, R.; Gimeno, L. Completeness of Radiosonde Humidity Observations Based on the Integrated Global Radiosonde Archive. *Earth Syst. Sci. Data* **2019**, *11*, 603–627. [CrossRef]
21. Liu, H.; Tang, S.; Zhang, S.; Hu, J. Evaluation of MODIS Water Vapour Products over China Using Radiosonde Data. *Int. J. Remote Sens.* **2015**, *36*, 680–690. [CrossRef]
22. 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]
23. Bevis, M.; Businger, S.; Chiswell, S.; Herring, T.A.; Anthes, R.A.; Rocken, C.; Ware, R.H. GPS Meteorology: Mapping Zenith Wet Delays onto Precipitable Water. *J. Appl. Meteor.* **1994**, *33*, 379–386. [CrossRef]
24. Naha Biswas, A.; Lee, Y.H.; Tao Yeo, W.; Low, W.C.; Heh, D.Y.; Manandhar, S. Statistical Analysis of Atmospheric Delay Gradient and Rainfall Prediction in a Tropical Region. *Remote Sens.* **2024**, *16*, 4165. [CrossRef]
25. Rohm, W.; Guzikowski, J.; Wilgan, K.; Kryza, M. 4DVAR Assimilation of GNSS Zenith Path Delays and Precipitable Water into a Numerical Weather Prediction Model WRF. *Atmos. Meas. Tech.* **2019**, *12*, 345–361. [CrossRef]
26. De Pondeca, M.; Zou, X. A Case Study of the Variational Assimilation of GPS Zenith Delay Observations into a Mesoscale Model. *J. Appl. Meteorol.* **2001**, *40*, 1559–1576. [CrossRef]
27. Kawabata, T.; Kuroda, T.; Seko, H.; Saito, K. A Cloud-Resolving 4DVAR Assimilation Experiment for a Local Heavy Rainfall Event in the Tokyo Metropolitan Area. *Mon. Weather. Rev.* **2011**, *139*, 1911–1931. [CrossRef]
28. Zhao, Q.; Liu, Y.; Yao, W.; Yao, Y. Hourly Rainfall Forecast Model Using Supervised Learning Algorithm. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 1–9. [CrossRef]
29. Wu, F.; Zhang, K.; Zhao, J.; Jin, Y.; Li, D. Linear and Nonlinear GNSS PWV Features for Heavy Rainfall Forecasting. *Adv. Space Res.* **2023**, *72*, 2170–2184. [CrossRef]
30. Benevides, P.; Catalao, J.; Nico, G. Neural Network Approach to Forecast Hourly Intense Rainfall Using GNSS Precipitable Water Vapor and Meteorological Sensors. *Remote Sens.* **2019**, *11*, 966. [CrossRef]
31. Benevides, P.; Catalao, J.; Miranda, P.M.A. On the Inclusion of GPS Precipitable Water Vapour in the Nowcasting of Rainfall. *Nat. Hazards Earth Syst. Sci.* **2015**, *15*, 2605–2616. [CrossRef]
32. Liu, Y.; Yao, Y.; Zhao, Q.; Li, Z. Stratified Rainfall Forecast Method Using GNSS Observations. *Atmos. Res.* **2022**, *280*, 106421. [CrossRef]
33. Li, H.; Wang, X.; Wu, S.; Zhang, K.; Fu, E.; Xu, Y.; Qiu, C.; Zhang, J.; Li, L. A New Method for Determining an Optimal Diurnal Threshold of GNSS Precipitable Water Vapor for Precipitation Forecasting. *Remote Sens.* **2021**, *13*, 1390. [CrossRef]
34. Zhao, Q.; Liu, Y.; Ma, X.; Yao, W.; Yao, Y.; Li, X. An Improved Rainfall Forecasting Model Based on GNSS Observations. *IEEE Trans. Geosci. Remote Sens.* **2020**, *58*, 4891–4900. [CrossRef]
35. Liu, Y.; Zhao, Q.; Yao, W.; Ma, X.; Yao, Y.; Liu, L. Short-Term Rainfall Forecast Model Based on the Improved BP–NN Algorithm. *Sci. Rep.* **2019**, *9*, 19751. [CrossRef] [PubMed]
36. Almonacid-Olleros, G.; Almonacid, G.; Gil, D.; Medina-Quero, J. Evaluation of Transfer Learning and Fine-Tuning to Nowcast Energy Generation of Photovoltaic Systems in Different Climates. *Sustainability* **2022**, *14*, 3092. [CrossRef]

37. Rasp, S.; Thuerey, N. Data-Driven Medium-Range Weather Prediction with a Resnet Pretrained on Climate Simulations: A New Model for WeatherBench. *J. Adv. Model Earth Syst.* **2021**, *13*, e2020MS002405. [CrossRef]
38. Bodnar, C.; Bruinsma, W.P.; Lucic, A.; Stanley, M.; Allen, A.; Brandstetter, J.; Garvan, P.; Riechert, M.; Weyn, J.A.; Dong, H.; et al. A Foundation Model for the Earth System. *Nature* **2025**, *641*, 1180–1187. [CrossRef]
39. Muñoz Sabater, J. *ERA5-Land Hourly Data from 1950 to Present*; Copernicus Climate Change Service (C3S) Climate Data Store (CDS): Reading, UK, 2019. [CrossRef]
40. Muñoz-Sabater, J.; Dutra, E.; Agustí-Panareda, A.; Albergel, C.; Arduini, G.; Balsamo, G.; Boussetta, S.; Choulga, M.; Harrigan, S.; Hersbach, H.; et al. ERA5-Land: A State-of-the-Art Global Reanalysis Dataset for Land Applications. *Earth Syst. Sci. Data* **2021**, *13*, 4349–4383. [CrossRef]
41. Saastamoinen, J. Atmospheric Correction for the Troposphere and Stratosphere in Radio Ranging Satellites. In *Geophysical Monograph Series*; Henriksen, S.W., Mancini, A., Chovitz, B.H., Eds.; American Geophysical Union: Washington, DC, USA, 2013; pp. 247–251. ISBN 978-1-118-66364-6.
42. Takasu, T.; Yasuda, A. Development of the Low-Cost RTK-GPS Receiver with an Open Source Program Package RTKLIB. In Proceedings of the International Symposium on GPS/GNSS 2009, Jeju, Republic of Korea, 4–6 November 2009.
43. Zhou, H.; Zhang, S.; Peng, J.; Zhang, S.; Li, J.; Xiong, H.; Zhang, W. Informer: Beyond Efficient Transformer for Long Sequence Time-Series Forecasting. In Proceedings of the AAAI Conference on Artificial Intelligence, 2–9 February 2021; AAAI Press: Palo Alto, CA, USA, 2021; pp. 11106–11115.
44. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention Is All You Need. In Proceedings of the Advances in Neural Information Processing Systems; Guyon, I., Luxburg, U.V., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., Garnett, R., Eds.; Curran Associates, Inc.: Red Hook, NY, USA, 2017; Volume 30.
45. Le Guen, V.; Thome, N. Deep Time Series Forecasting with Shape and Temporal Criteria. *IEEE Trans. Pattern Anal. Mach. Intell.* **2023**, *45*, 342–355. [CrossRef]
46. Sakoe, H.; Chiba, S. Dynamic Programming Algorithm Optimization for Spoken Word Recognition. *IEEE Trans. Acoust. Speech Signal Process.* **1978**, *26*, 43–49. [CrossRef]
47. Vallance, L.; Charbonnier, B.; Paul, N.; Dubost, S.; Blanc, P. Towards a Standardized Procedure to Assess Solar Forecast Accuracy: A New Ramp and Time Alignment Metric. *Sol. Energy* **2017**, *150*, 408–422. [CrossRef]
48. Shi, J.; Shirali, A.; Jin, B.; Zhou, S.; Hu, W.; Rangaraj, R.; Wang, S.; Han, J.; Wang, Z.; Lall, U.; et al. Deep Learning and Foundation Models for Weather Prediction: A Survey 2025. *arXiv* **2025**, arXiv:2501.06907.
49. Wang, C.; Wu, Q.; Weimer, M.; Zhu, E. FLAML: A Fast and Lightweight AutoML Library. In Proceedings of the Machine Learning and Systems, Stanford, CA, USA, 31 March–2 April 2019. [CrossRef]
50. Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016.
51. Meng, Q. LightGBM: A Highly Efficient Gradient Boosting Decision Tree. In Proceedings of the Neural Information Processing Systems, Long Beach, CA, USA, 4–9 December 2017.
52. Kunkel, K.E.; Karl, T.R.; Squires, M.F.; Yin, X.; Stegall, S.T.; Easterling, D.R. Precipitation Extremes: Trends and Relationships with Average Precipitation and Precipitable Water in the Contiguous United States. *J. Appl. Meteorol. Climatol.* **2020**, *59*, 125–142. [CrossRef]
53. Poelzl, M.; Kern, R.; Kecorius, S.; Lovrić, M. Exploration of Transfer Learning Techniques for the Prediction of PM10. *Sci. Rep.* **2025**, *15*, 2919. [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.