

Article

A Hybrid Deep Learning Algorithm for Tropospheric Zenith Wet Delay Modeling with the Spatiotemporal Variation Considered

Yin Wu, Lu Huang, Wei Feng * and Su Tian

Faculty of Geosciences and Environmental Engineering, Southwest Jiaotong University, Chengdu 611756, China; ywu@my.swjtu.edu.cn (Y.W.); huanglu@swjtu.edu.cn (L.H.); tiansu@my.swjtu.edu.cn (S.T.)

* Correspondence: wfeng@swjtu.edu.cn

Abstract: The tropospheric Zenith Wet Delay (ZWD) is one of the primary sources of error in Global Navigation Satellite Systems (GNSS). Precise ZWD modeling is essential for GNSS positioning and Precipitable Water Vapor (PWV) retrieval. However, the ZWD modeling is challenged due to the high spatiotemporal variability of water vapor, especially in low latitudes and specific climatic regions. Traditional ZWD models make it difficult to accurately fit the nonlinear variations in ZWD in these areas. A hybrid deep learning algorithm is developed for high-precision ZWD modeling, which considers the spatiotemporal characteristics and influencing factors of ZWD. The Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) are combined in the proposed algorithm to make a novel architecture, namely, the hybrid CNN-LSTM (CL) algorithm, combining CNN for local spatial feature extracting and LSTM for complex sequence dependency training. Data from 46 radiosonde sites in South America spanning from 2015 to 2021 are used to develop models of ZWD under three strategies, i.e., model CL-A without surface parameters, model CL-B with surface temperature, and model CL-C introducing surface temperature and water vapor pressure. The modeling accuracy of the proposed models is validated using the data from 46 radiosonde sites in 2022. The results indicate that CL-A demonstrates slightly better accuracy compared to the Global Pressure and Temperature 3 (GPT3) model; CL-B shows a precision increase of 14% compared to the Saastamoinen model, and CL-C exhibits accuracy improvements of 30% and 12% compared to the Saastamoinen and Askne and Nordius (AN) model, respectively. Evaluating the models' generalization capabilities at non-modeled sites in South America, data from six sites in 2022 were used. CL-A shows overall better performance compared to the GPT3 model; CL-B's accuracy is 19% better than the Saastamoinen model, and CL-C's accuracy is enhanced by 33% and 10% compared to the Saastamoinen and AN model, respectively. Additionally, the proposed hybrid algorithm demonstrates a certain degree of improvement in both modeling accuracy and generalization accuracy for the South American region compared to individual CNN and LSTM algorithm.



Citation: Wu, Y.; Huang, L.; Feng, W.; Tian, S. A Hybrid Deep Learning Algorithm for Tropospheric Zenith Wet Delay Modeling with the Spatiotemporal Variation Considered. *Atmosphere* **2024**, *15*, 121. <https://doi.org/10.3390/atmos15010121>

Academic Editors: Li Li and Pengfei Xia

Received: 7 November 2023

Revised: 11 January 2024

Accepted: 16 January 2024

Published: 19 January 2024

Keywords: zenith wet delay (ZWD); convolutional neural network (CNN); long short-term memory (LSTM); deep learning; spatiotemporal characteristics

1. Introduction

Global Navigation Satellite System (GNSS) signals are affected by atmospheric refraction and bending of the propagation path, resulting in propagation delay, which is one of the main sources of GNSS positioning [1]. The troposphere is a non-dispersive medium, and the tropospheric delay is frequency-independent. This delay cannot be weakened by dual-frequency or multi-frequency combinations [2,3]. The tropospheric delay is related to the observed satellite elevation angle, and the Zenith Tropospheric Delay (ZTD) is usually mapped to the slant path of the observed satellite elevation angle by a mapping function. ZTD can be regarded as both a hydrostatic and non-hydrostatic component [4]. The former is called Zenith Hydrostatic Delay (ZHD), which accounts for more than 90% of the total delay, and the latter is the Zenith Wet Delay (ZWD), caused by water vapor in the lower



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/>).

atmosphere, generally only accounting for 10% or less of the total delay [5]. Chen and Liu [6] counted and analyzed the modeling accuracy of 9 ZHD models and 18 ZWD models, and the results showed that the modeling accuracy of the existing models for the ZHD could reach the sub-centimeter level, while the modeling accuracy of the ZWD was poor, up to 10 cm.

Due to the high variation in water vapor in the lower atmosphere in terms of time, space, and altitude, it is difficult to accurately calculate ZWD in general [7,8], which makes it more difficult to estimate ZWD in the process of GNSS positioning solution, thereby prolonging the convergence time of Ambiguity Resolution (AR) [9]. Therefore, accurate modeling of ZWD is a crucial issue that holds great significance for monitoring the atmospheric water vapor content [10]. Prior accurate ZWD can significantly improve the performance of positioning and location-based service. Studies have shown that prior accurate ZWD constraints could shorten the convergence time of Precise Point Positioning (PPP), especially in the up direction [11–13]. Jiang et al. [14] confirmed that the residual caused by large height differences can be weakened by attaching tropospheric constraints, and the fixed rate of AR and coordinate accuracy can be improved. Generally, ZWD can be obtained by the following: 1. Empirical models, such as the University of New Brunswick (UNB) proposed models [15] (including UNB, UNB3, and UNB3m) and the Global Pressure and Temperature (GPT) models [16] (including GPT, GPT2, GPT2w, and GPT3); 2. The meteorological parameter model can achieve centimeter-level ZWD correction accuracy by inputting the surface meteorological parameters. The common models include the Hopfield [17], Saastamoinen [18], Black [19], and Askne and Nordius (AN) [20] models. Among them, the AN model also needs to provide the weighted mean temperature (T_m) and the parameter of the exponential decay vertical trend of the water vapor (λ) as input parameters so it can describe the vertical water vapor variation over the site, and is considered the most accurate formula for ZWD estimation [21]; 3. Measuring the atmospheric profiles by the radiosonde over the site and calculating the ZWD by integrating non-hydrostatic refractive index in each level; 4. ZWD is estimated as an unknown parameter [11]. However, empirical models are difficult to accurately predict short-term or non-trend fluctuations of ZWD in some specific regions (generally between the latitudes of 30° N and 30° S) [22], and the applicability of meteorological parameter models in specific regions is greatly reduced [21]. Although the direct estimation as a parameter is effective, it may decrease the efficiency of an algorithm.

In recent years, machine learning (deep learning) algorithms have achieved promising applications in atmospheric modeling thanks to their excellent abilities in information extraction, nonlinear feature modeling, and large data processing [23,24]. Yang et al. [25] used the Back Propagation (BP) neural network to model the residuals of the Hopfield model and Saastamoinen model. They improved the ZWD modeling accuracy of the meteorological parameter model. Li et al. [26] used three machine learning algorithms of the BP neural network, the Radial Basis Function (RBF) neural network, and the Least Squares Support Vector Machine (LSSVM), combined with the fifth generation of the European Centre for Medium-Range Weather Forecasts Reanalysis (ERA5) and GNSS data, to construct the ZTD model of 2020 in North America. The experiments showed that the accuracy and stability of LSSVM and RBF were better than BP, but the LSSVM and RBF could not be applied to large sample data. Based on 118 radiosonde sites in China and surrounding areas, Gao et al. [27] established two T_m models using the Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) neural network, comparing to the BP neural network, confirmed that LSTM had better generalization ability for long-term sequence processing than RNN due to its special memory cell, and the nonlinear fitting ability of LSTM was stronger than BP neural network. Lu et al. [28] established a Precipitable Water Vapor (PWV) model on the West Coast of the United States by fusing Moderate-resolution Imaging Spectroradiometer (MODIS) and ERA5 data using the Convolutional Neural Network (CNN). Compared to the Multilayer Perceptron (MLP) algorithm, one verified that CNN could extract more detailed spatial features in multivariate time series data. Osah et al. [29]

constructed a regional ZTD model based on the location and surface meteorological parameters (pressure, temperature, and water vapor pressure) of four International GNSS Service (IGS) stations in West Africa using the deep learning algorithm. Ding et al. [30] verified the effectiveness of multi-parameters for constructing the ZWD model by the multilayer feedforward neural network.

Tropospheric delay is greatly affected by location, season, and other factors. Current research focuses on improving the performance of ZWD models in low-latitude tropical regions [31]. However, the algorithm's structural characteristics and the adaptability of the atmospheric spatiotemporal characteristics are not considered in most studies. Therefore, taking advantage of the deep learning algorithms, this paper combines the local spatial feature extraction ability of CNN and the ability of LSTM to learn complex sequences with long-term dependencies for ZWD modeling research. The hybrid algorithm is proposed from encoder to decoder, while CNN is an encoder and LSTM is a decoder. The input parameters of each epoch are encoded by the multilayer CNN, and the local spatial features are identified and then compressed into a one-dimensional vector with spatiotemporal information transmitted to the LSTM. The multilayer LSTM receives the spatiotemporal features sequence, then extracts the long sequence of temporal dependencies and inter-site spatial features and decodes the output. The ZWD modeling strategies are considered in different scenarios and explore the influence of the surface meteorological parameters on ZWD regression modeling. This paper evaluates the accuracy of models established by the hybrid algorithm, using both modeling and non-modeled sites in the South American region for the year 2022. The assessment utilizes Root Mean Square Error (RMSE) as a metric and compares the proposed models with empirical and meteorological parameter models. The results indicate that the hybrid algorithm exhibits good spatiotemporal modeling ability in South America and shows a significant enhancement in ZWD modeling accuracy compared to the empirical and meteorological parameter models. This study validates the effectiveness of the hybrid approach in enhancing the modeling precision of ZWD in South America.

The structure of this paper is as follows. Section 2 describes the principle and structure of the proposed hybrid deep learning algorithm. Subsequently, Section 3 introduces the dataset employed for ZWD modeling and the determination of hyperparameters for the hybrid deep learning algorithm. Section 4 is the experimental analysis. By comparing with the numerical integration of radiosonde data, the spatiotemporal modeling accuracy and generalization ability of different models are discussed. The conclusion is given in Section 5.

2. Methodology

2.1. Spatial Feature Extraction of ZWD

ZWD is affected by the location of the site, water vapor in the lower atmosphere, temperature, and other factors varying with time, and the spatiotemporal characteristics of ZWD should be fully considered in the modeling so as to fit the complex nonlinear relationship and improve the ZWD modeling ability [32,33].

ZWD modeling is formulated as a spatiotemporal sequence regression and forecasting problem. Spatiotemporal characteristics are usually affected by many factors: (1) complex spatiotemporal feature changes due to the regional spatiotemporal correlations; (2) different spatiotemporal variation characteristics of ZWD due to the changes in climatic conditions at different sites; (3) local spatiotemporal correlations help capture the changes in ZWD in neighboring regions, while global spatiotemporal correlations are beneficial for extracting overall ZWD variations.

CNN was initially employed in the field of image recognition [34] and has later found widespread application in natural language processing [35] and time series forecasting [36]. CNN exhibits enhanced capability in extracting latent information from multidimensional data. In contrast to the intricate structures and hidden nodes of deep networks, CNN significantly reduces computational costs through weight sharing [37]. A typical CNN

includes convolutional layers, pooling layers, and fully connected layers. The convolutional layer consists of a series of convolutional kernels whose parameters are adjusted through the backpropagation algorithm. Convolution is considered a linear operation, sliding over the input data, multiplying at each position, and summing the results, thus extracting various features from the input data [38]. The pooling layer reduces the computational load of the network by down-sampling the features from the convolutional layer output [39]. The fully connected layer combines all local features into global features [40]. Therefore, the advantages of the convolutional layer and pooling layer are considered, employing them as encoders in the hybrid algorithm to extract local spatial features from the input data. The mathematical expressions of convolution and pooling layers are shown in Equations (1) and (2), respectively, in which the pooling layer selects the max pooling mechanism.

$$Z_{m-n+1}^{(l)}(u) = \sum_{i=1}^m \sum_{j=1}^n x_m^{(l-1)}(i) \cdot k_n^{(l)}(j) + b^{(l)} \tag{1}$$

$$A(i) = \max_{p=1}^{i+q-1} Z(p) \tag{2}$$

where $x_m^{(l-1)}$ and $Z_{m-n+1}^{(l)}$ represent the input and output of the current convolutional layer; $k_n^{(l)}$ and $b^{(l)}$ are the convolution kernel and bias in the current layer, respectively. n , m , and $m - n + 1$ denote the sizes of the convolution kernel, input, and output in the current layer, respectively. A is the pooling layer output. q is the pooling window size.

Figure 1 illustrates the process of local spatial feature extraction by the convolutional and pooling layers for input data. Here, b and m represent the time and feature dimensions of the input data, respectively, while the rest corresponds to Equations (1) and (2). Utilizing the one-dimensional CNN for fixed-time step input data convolution in the time series [36,41], in the feature dimension, as indicated by Equation (1), the feature dimension is reduced to $m - n + 1$ after a single convolution, and it decreases with the stacking of convolution layers. Along the time axis, the convolution kernel performs convolution calculations with a stride of 1 from top to bottom. The pooling layer further compresses the spatial features extracted by the convolutional layer. Due to the fixed-time step in the operations of the convolutional and pooling layers, their outputs do not alter the size of the time dimension. The flattening layer unfolds the output of the pooling layer into a one-dimensional sequence inputted into the LSTM. Based on the above analysis, it is evident that the sequence unfolded by the flattening layer contains both temporal and local spatial information extracted by CNN.

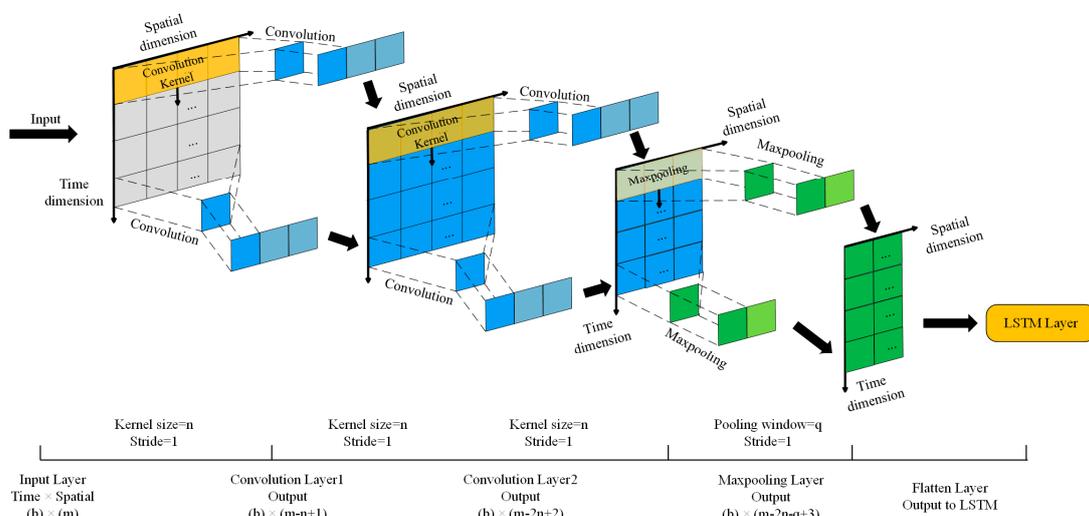


Figure 1. Local spatial features extraction.

2.2. Feature Training for Long Spatiotemporal Sequences

Although CNN can extract data spatial correlation characteristics, its applicability will decrease in a long-term complex sequence [42]. RNN is frequently employed in time series modeling owing to its recurrent structure. However, for long time series, the backpropagation of errors in RNN, computed through the chain rule, can result in the vanishing or exploding gradient problem [43]. Consequently, RNN can effectively learn only short-term dependencies in the data. LSTM is a variation of RNN that addresses this issue by introducing a memory unit structure. The memory unit consists of three components: the input gate; the forget gate; and the output gate [44]. The input gate controls the inputting of external data to the memory cell, while the forget gate and output gate determine whether information is retained or released at each time step, thereby accommodating both long-term and short-term dependencies in the input data. Therefore, this paper introduces LSTM to receive the output sequence from the flattening layer and learns long- and short-term dependency information through the interaction between the gating mechanism and the unit state update module. The mathematical expressions are shown in Equations (3)–(8).

$$f_t = \sigma(W_f \cdot [h_{t-1}, A(t)] + b_f) \quad (3)$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, A(t)] + b_i) \quad (4)$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, A(t)] + b_C) \quad (5)$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (6)$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, A(t)] + b_o) \quad (7)$$

$$h_t = o_t * \tanh(C_t) \quad (8)$$

Equation (3) indicates that the forget gate receives the input $A(t)$ of the current epoch and the hidden state h_{t-1} of the previous epoch, calculating weights for each input section to update C_t . Equation (4) indicates an update to the information entered in the current cell. Equation (5) creates a candidate memory cell \tilde{C}_t using the tanh function. Equation (6) represents the preservation weight for the previous cell state C_{t-1} and the current candidate memory cell \tilde{C}_t in the update of the current cell's state. Equations (7) and (8) indicate the weight of output in the cell and the influence of current information on the hidden state h_t , respectively. W_* and b_* represent the weight and bias matrix of the corresponding module. Among them, $\sigma(\cdot)$ is the sigmoid function, which can be represented using Equation (9), $\tanh(\cdot)$ is the tanh function, which can be represented with Equation (10).

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (9)$$

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (10)$$

As illustrated in Figure 1, after the convolutional and pooling layers, the input data is flattened into a one-dimensional sequence. This sequence contains both the temporal and spatial feature information of the input data. Subsequently, the LSTM layer receives the sequence, through its special memory cell, identifies long and short-term dependencies, modeling the output as illustrated in Figure 2. Cell represents the cell structure of the LSTM, and the pooling layer transmits the spatiotemporal sequence to the LSTM layer. Each cell's input includes the local spatial feature of each site, and the cell state C and hidden state h are calculated by the previous site via the cell. It means that LSTM can identify the short-term characteristics, as shown by the inter-site spatial characteristic in Figure 2. According to Equations (6) and (8), LSTM can selectively retain key information and update C and h between the historical epoch and the current epoch by the gating

mechanism so that C and h can better capture the long-term dependency, which is the temporal characteristic for one site in Figure 2.

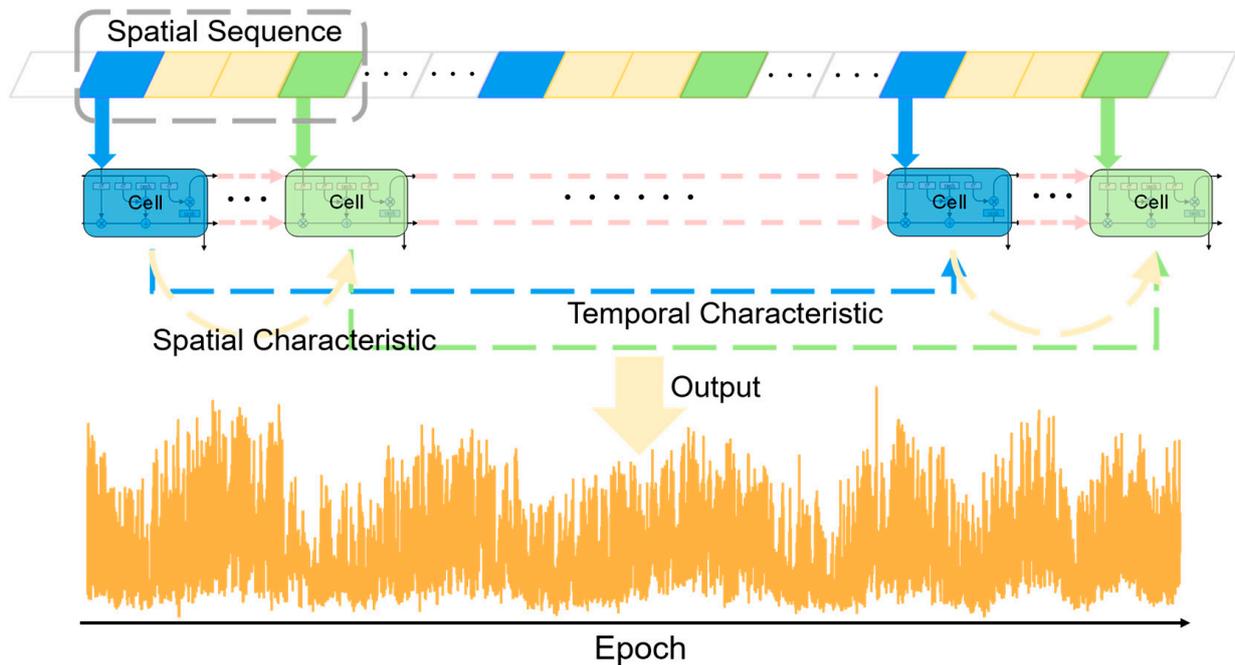


Figure 2. Recognition of spatiotemporal characteristic sequence.

2.3. Combination of CNN and LSTM

Based on the above analysis, the hybrid algorithm of CNN-LSTM (CL) is proposed, and the hybrid architecture is shown in Figure 3, where DOY represents the day of the year; Lat is latitude; Lon is longitude; H is orthometric height; T_s denotes surface temperature, and es indicates surface water vapor pressure. X_A , X_B , and X_C correspond to different training samples. The input layer receives the information of various parameter features; then, the convolutional layer and the pooling layer are used to extract local spatial features. The convolutional layer plays a similar filtering role through convolution operation to recognize the important features of the input parameters [45]. The pooling layer compresses the features extracted by convolution, which are then transformed into a one-dimensional sequence containing both temporal information and spatial features through a flattening layer. They act as the data encoder in the entire network. The max pooling mechanism of the pooling layer retains the essence of data [39]. The LSTM layer trains the spatiotemporal correlation between spatiotemporal sequences through its special memory cell and fits the nonlinear relationship between the input parameters and the ZWD outputs.

The proposed algorithm leverages a serial coupling of CNN and LSTM, allowing for the integration of spatial features extracted by CNN with the temporal modeling capabilities of LSTM [46]. The encoder uses the stacking of two sets of convolutional layers to extract multiple sets of convolution results. The first convolutional layer extracts the basic features of the input data, while the second layer further captures composite features of these basic features [47]. Considering the feature dimensions of the input data, the convolutional kernel size is fixed at 3 with a stride of 1, without zero padding, adding a batch normalization layer before the activation layer to mitigate the vanishing gradient problem. The pooling layer receives the extracted features by convolutional layer and compresses them into spatiotemporal feature vectors through a pooling window of size 2 with a stride of 1. The decoder is designed as a two-layer sequence structure, which produces an output for each epoch. The information of C and h transfer between adjacent cells capture and process the long-term and spatial correlations in the spatiotemporal sequence. Finally, observing the output of a single epoch, the overfitting phenomenon is

restrained by adding a dropout layer to abandon part of neuronal activation. This achieves a hybrid CL algorithm that combines the advantages of CNN for spatial feature extraction and LSTM for long sequence learning.

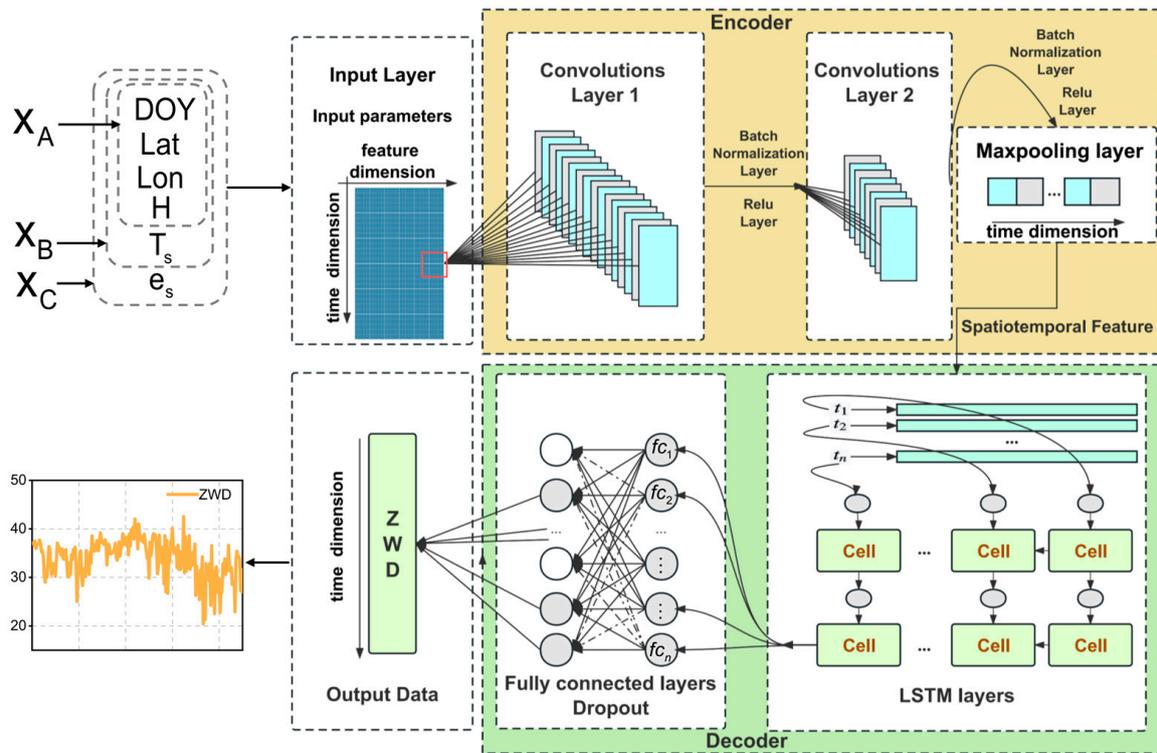


Figure 3. Structure of the hybrid algorithm.

In addition, Mean Square Error (MSE) is employed as the loss function, and the backpropagation algorithm is utilized to update network parameters, minimizing the MSE between the predicted value \bar{y} and reference value y . The mathematical expression of the loss function is as follows. N is the total number of samples.

$$Loss(\bar{y}, y) = \frac{1}{N} \sum_{i=1}^N (y - \bar{y})^2 \tag{11}$$

3. Case Study: South American ZWD Modeling

3.1. Data Sources and ZWD Extraction

Radiosonde is one of the most accurate means of measuring the ZWD [48], and the radiosonde data in this research is collected from the Integrated Global Radiosonde Archive (IGRA, <https://www.ncei.noaa.gov/pub/data/igra/>, (accessed on 15 October 2023)). IGRA covers high-quality meteorological observations from more than 1500 sites since 1960. These observations provide vertical profiles of tropospheric meteorology, including air temperature, total pressure, relative humidity, altitude, and more. The data are recorded twice a day at 00:00 and 12:00 (UTC) [49]. This study focuses on South America, a region that spans both sides of the equator and exhibits complex climatic and topographical conditions. The presence of the Amazon Rain Forest in Northern South America contributes significantly to the water vapor, potentially causing notable deviations in the ZWD modeling of the surrounding area [50]. In order to establish the ZWD model in South America, 46 radiosonde sites were selected for modeling from 13° N to 54° S in latitude and 31° W to 82° W in longitude in 2015–2021. Six radiosonde sites in 2022 were used to verify the generalization ability of the modeling. Figure 4 shows the distribution of selected radiosonde sites, with red dots for sites participating in training for modeling

and performance evaluation, and blue dots for non-modeled sites that are not involved in training but only used to evaluate the generalization ability of the model. In addition, the radiosonde data were preprocessed with quality control before extracting ZWD. The data were removed when two adjacent isobaric layers exceeded 200 hPa or the top pressure layer was greater than 300 hPa [21].

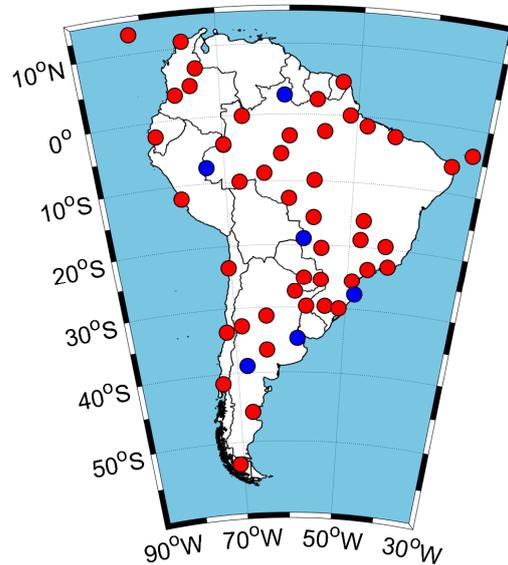


Figure 4. Distribution of radiosonde sites in South America (Red dots indicate sites involved in training and evaluating modeling performance, while blue dots indicate sites not modeled and used solely to evaluate generalization ability).

ZWD is considered to be the integral of the non-hydrostatic refractive index [4]. According to the atmospheric vertical profiles obtained from the radiosonde data, the refractive index of the zenith direction of the radiosonde site can be estimated, and then the ZWD can be obtained after numerical integration. The ZWD obtained from the radiosonde is used as a reference value. The equation is expressed as follows [51]:

$$ZWD = 10^{-6} \int_{H_s}^{Top} N_w(h) dh = 10^{-6} \int_{H_s}^{Top} [k'_2 \cdot \frac{e}{T} + k_3 \cdot \frac{e}{T^2}] \cdot Z_w^{-1} dh \quad (12)$$

where k'_2 and k_3 are atmospheric refractivity coefficients [52]. e and T indicate the water vapor pressure [hPa] and temperature [K] of each isobaric layer above the site, respectively. Z_w is the compressibility factor for atmospheric water vapor. H_s and Top represent the site altitude and the top of the lower atmosphere, respectively. They can be obtained by converting the geopotential height of the IGRA data. The conversion formulas are as follows [53]:

$$H(h, \varphi) = \frac{\gamma_{45} \cdot R(\varphi) \cdot h}{\gamma(\varphi) \cdot R(\varphi) - \gamma_{45} \cdot h} \quad (13)$$

$$\gamma(\varphi) = 9.780325 \cdot \left[\frac{1 + 000193185 \cdot \sin(\varphi)^2}{1 - 0.00669435 \cdot \sin(\varphi)^2} \right]^{1/2} \quad (14)$$

$$R(\varphi) = \frac{6378.137}{1.006803 - 0.006706 \cdot \sin(\varphi)^2} \quad (15)$$

where H and h represent orthometric height [m] and geopotential height [m], respectively. $\gamma(\varphi)$ and $R(\varphi)$ are the gravitational acceleration and radius of curvature of the earth at latitude φ . γ_{45} is the normal gravity at latitude 45° with a value of 9.80665 m/s^2 . Finally, the

conversion between the orthometric height and the geodetic height is completed through the Earth Gravitational Model 2008 (EGM2008) [54].

ZWD is integrated by 46 radiosonde data in 2019, and the spatiotemporal changes with the seasons are plotted, as shown in Figure 5. Seasons and corresponding months are shown in Table 1. As shown in Figure 5, ZWD varies with spatial location and has an obvious negative correlation with latitude. ZWD generally shows a trend of being bigger in the east and north and smaller in the west and south of South America. The reason for this phenomenon is that the western terrain distribution is higher, and the lower the temperature and water vapor pressure, the smaller the ZWD. The northern part is mainly affected by the tropical rainforest climate, with high temperatures and rainfall throughout the year, and the ZWD is bigger in summer and autumn. ZWD is significantly affected by the subtropical monsoon climate and the temperate continental climate seasonal changes in the southeastern region and southern region, respectively. It presents the characteristics of larger ZWD in summer and smaller ZWD in winter. Therefore, ZWD is affected by the climate and geographic regions in South America, displaying distinct spatiotemporal characteristics throughout the year.

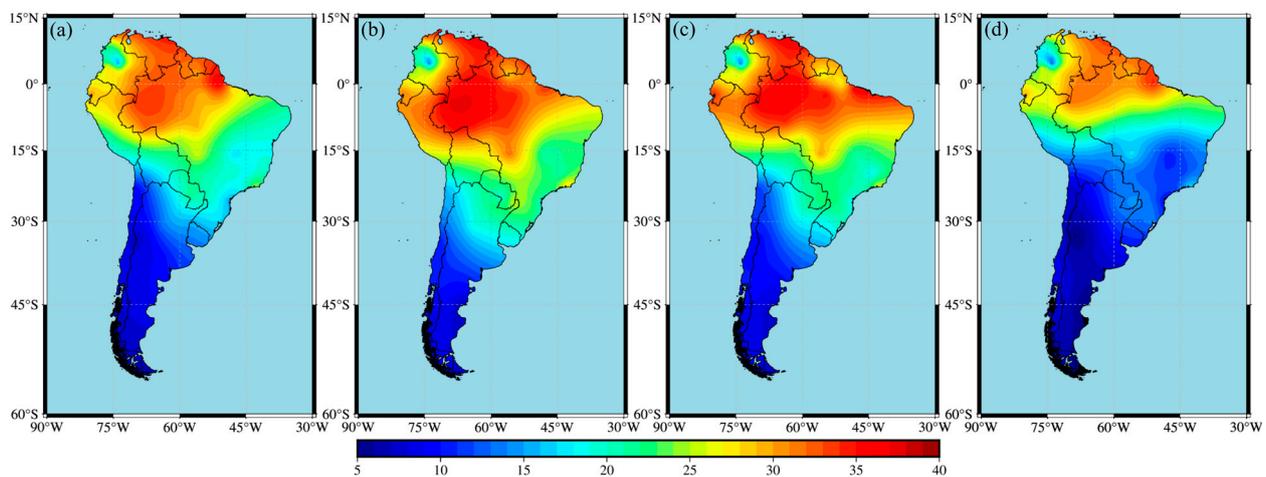


Figure 5. Spatiotemporal variation in 2019 ZWD in South America with the seasons. (a) Spring. (b) Summer. (c) Autumn. (d) Winter.

Table 1. Seasonal Correspondence in Hemispheres.

Season	Corresponding Months in Northern Hemisphere	Corresponding Months in Southern Hemisphere
Spring	March, April, May	September, October, November
Summer	June, July, August	December, January, February
Autumn	September, October, November	March, April, May
Winter	December, January, February	June, July, August

3.2. ZWD Modeling and Impact Factors

The proposed ZWD modeling approach takes the spatiotemporal distribution of ZWD and its factors into account, as shown in Figure 6, mainly including three parts. The first part is data quality control, extracting ZWD from the numerical integration of atmospheric profiles and correlation analysis to determine the relevant impact factors of ZWD modeling. Normalization is conducted to eliminate dimensional discrepancies among various features. Secondly, three strategies are determined according to different scenarios, and the combined ZWD models based on CL are established. The 7-year modeling impact factors are used as an input, and the ZWD obtained from the radiosonde is an output. Cross-validation is conducted to determine the hyperparameters of the combined ZWD models. Simultaneously, the ZWD models are established by the same operation using the CNN

algorithm and LSTM algorithm to verify the effectiveness of the hybrid algorithm. Finally, the combined ZWD models are applied to calculate the ZWD of 2022 and evaluate the performance by comparing them to the traditional models (empirical and meteorological parameter models).

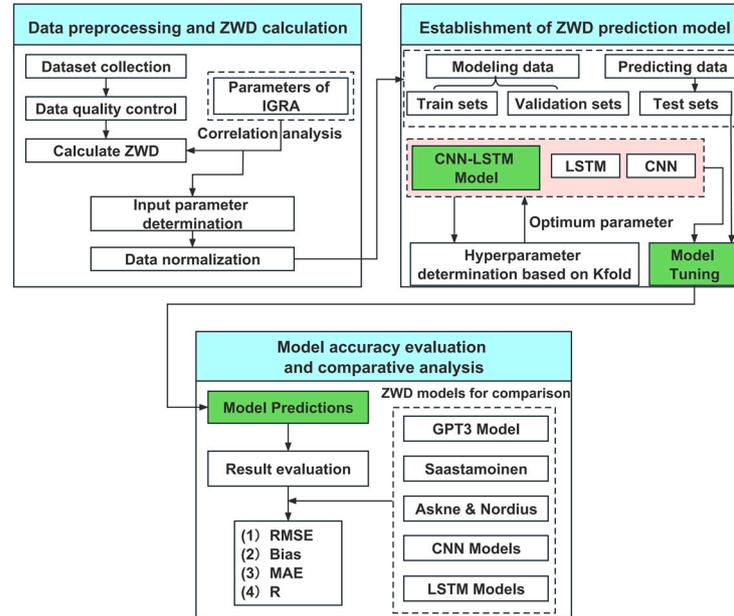


Figure 6. Hybrid ZWD modeling process.

Previous studies on ZWD modeling show that the location, surface temperature, and surface water vapor pressure of the site are factors affecting ZWD [21,29,55]. Correlation analysis is used to calculate the correlation coefficient between different parameters, serving as the basis for selecting the modeling impact factor, as shown in Figure 7.

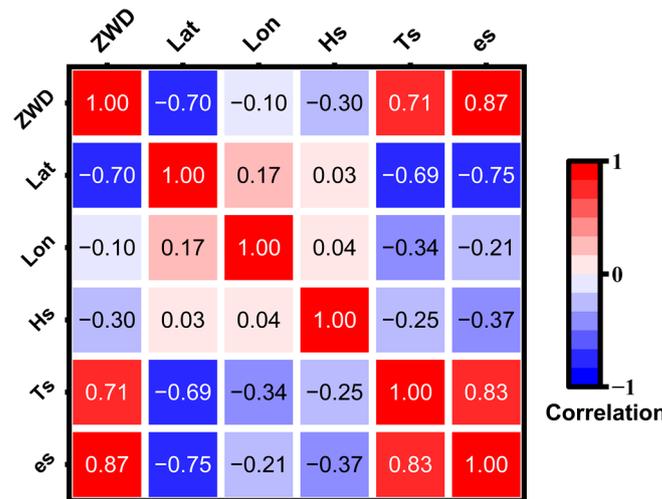


Figure 7. Pearson correlation coefficient of modeling parameters.

As seen in Figure 7, ZWD exhibits a clear positive correlation with surface meteorological parameters (T_s and e_s), while it shows a negative correlation with latitude and altitude. The correlation with longitude is relatively low. Among these factors, e_s has the highest correlation coefficient of 0.87. Similarly, T_s and e_s exhibit a significant correlation with latitude and a relatively weak correlation with longitude while showing a negative correlation with altitude.

Three modeling strategies are designed to explore the influence of surface meteorological parameters on ZWD regression modeling according to the difficulty of obtaining surface meteorological parameters and the correlation coefficient. These strategies include (A) strategies without surface meteorological parameters, (B) strategies using surface temperature, and (C) strategies using surface temperature and water vapor pressure. The output of the model is the high-precision ZWD value calculated by integrating from radiosonde data. Three spatiotemporal distribution models of regional ZWD are established by the proposed hybrid algorithm, using radiosonde data from 2015 to 2021. Table 2 presents the different modeling strategies used in this paper, along with the required parameters and their respective function expressions. The GPT3 model is calculated by the AN model using the meteorological parameters from the Numerical Weather Model (NWM) through a spherical harmonic function inputting the location of the site and DOY.

Table 2. Parameters and formulations for different modeling methods.

Name	Input Parameters	Formulation
CL-A	DOY, Lat, Lon, H _s	$F_A(DOY, Lat, Lon, H_s)$
GPT3		/
CL-B	DOY, Lat, Lon, H _s , T _s	$F_B(DOY, Lat, Lon, H_s, T_s)$
Saastamoinen	Lat, H _s , T _s , e _s	$\frac{0.002277 \cdot (\frac{1255}{T_s} + 0.05) \cdot e_s}{1 - 0.00266 \cdot \cos(2 \cdot Lat) - 0.00028 \cdot H_s}$
CL-C	DOY, Lat, Lon, H _s , T _s , e _s	$F_C(DOY, Lat, Lon, H_s, T_s, e_s)$
Askne and Nordius	e _s , T _m , λ	$10^{-6} \cdot (k'_2 + \frac{k_3}{T_m}) \cdot \frac{R_d}{(\lambda + 1) \cdot g_m} \cdot e_s$

Due to the different magnitude order between the above parameters, direct modeling using the original data may ignore the impact of smaller orders of magnitude of data. Original data will be standardized in advance, using the maximum and minimum standardization method:

$$X' = \frac{2(X - X_{\min})}{X_{\max} - X_{\min}} - 1 \tag{16}$$

where X' is the normalized value; X is the original value, and X_{max} and X_{min} represent the maximum and minimum values, respectively, X' ∈ [−1, 1].

3.3. Determination of Hyperparameters

Deep learning primarily focuses on uncovering the connections between inputs and outputs through hidden layers. In the case of the CL hybrid algorithm, crucial hyperparameters impacting the algorithm's learning and generalization include the number of convolution kernels in the CNN layer and the number of hidden neurons in the LSTM layer [27,56]. To attain optimal modeling and generalization ability, 10 structures are designed in the CNN layer with the number of convolution kernels ranging from 4 to 40 and a step size of 4. Simultaneously, in the LSTM layer, 10 structures are designed with the number of hidden neurons ranging from 5 to 50 and a step size of 5 to identify the optimal hyperparameters in ZWD modeling. Meanwhile, the performance of CNN and LSTM are compared in the same conditions. The max epoch is set to 100, and the mini-batch is set to 128. Figure 8 shows the effect of the number of hyperparameters on the model RMSE. The global performance of the model under different hyperparameters is analyzed by grid-search and 10-fold cross-validation. The 10-fold cross-validation uses 90% of the input data as training samples and 10% as validation samples, iterating this process ten times to ensure that the entire input dataset is utilized for training. Model optimization is achieved by minimizing the loss function through the backpropagation algorithm. Therefore, in the early stages of model optimization, the RMSE of both training and validation samples decreases with an increase in hyperparameters. When the RMSE curve of the validation samples converges or shows an inflection point, it can be considered that the

model parameters are appropriately optimized. Otherwise, overfitting or underfitting may occur.

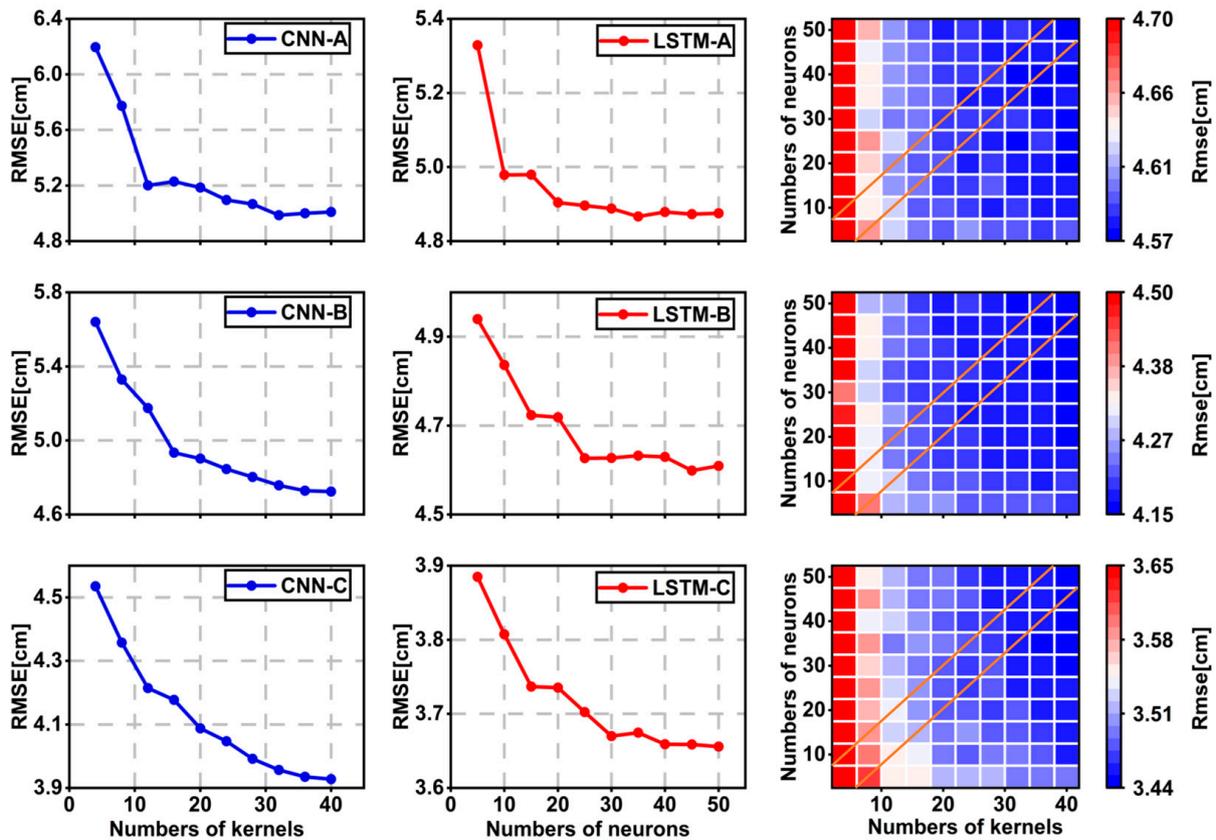


Figure 8. RMSE impacts of different hyperparameters on modeling.

As shown in Figure 8, the RMSE of the model is affected by the number of hyperparameters, and the RMSE of different algorithms shows a downward trend with the increase in their respective hyperparameters. Furthermore, with the hyperparameters number increasing, the hybrid algorithm exhibits better accuracy compared to CNN and LSTM algorithms. The number selection of the convolution kernel depends on the inflection point appearing or flattening in the RMSE curve, so the convolution kernel numbers of CNN-A, CNN-B, and CNN-C are set to 32, 36, and 36, respectively. Similarly, for LSTM, the number of hidden neurons of A/B/C are set to 35/45/40, respectively. In the hybrid algorithm, considering the influence of the two hyperparameters, the accuracy change in the diagonal direction is used as the basis for the optimal hyperparameters. The RMSE decreases with the increase in the number of hyperparameters in the diagonal direction, so the hyperparameters of A/B/C are set to (28,35), (24,30), and (32,40), respectively. In addition, the optimizer algorithm is used, and the overfitting phenomenon is further suppressed by adding a dropout layer to improve the generalization ability of the model, with the dropout value set to 0.1. The proposed hybrid deep learning algorithm, as well as the CNN and LSTM algorithms depicted in Figure 8, are all implemented using the Matlab Neural Network Toolbox. Table 3 provides the additional parameter settings for the proposed algorithm. Table 4 summarizes the average pre-training time, prediction time, and changes in learnable parameters for each algorithm under three different strategies.

Table 3. Parameter setting of the proposed algorithm.

Parameter Type	Initial Value	Search Range	
Input Size	Dependent on feature dimensions	/	
Kernel Size	3	/	
Pooling Size	2	/	
Stride	1	/	
Padding	no	/	
Learning Rate	0.01	0.001–0.01	MaxEpochs = 100; Minibatchsize = 128; Optimizer = ‘Adam’
Learning Rate Drop Period	10	$0.1-0.5 \times \text{MaxEpochs}$	
Learning Rate Drop Factor	0.1	0.1–0.9	
Dropout	0.1	0.1–0.5	
Output Size	1	/	

Table 4. Average pre-training time, prediction time, and parameter changes for different algorithms under three strategies.

Algorithm	Training Time	Parameter Numbers Change	Prediction Time
CNN	About 8 min	289–361	About 3 s
LSTM	About 11 min	5700–9400	About 4 s
CNN-LSTM	About 45 min	33,600–69,700	About 5 s

Table 4 illustrates that the hybrid deep learning algorithm demonstrates a more intricate network structure and a notable increase in learnable parameters compared to the individual CNN and LSTM algorithms. Consequently, additional time is needed to update the network parameters during the pre-training phase. However, after completing pre-training and establishing the optimal model coefficients between input and output, the proposed hybrid algorithm does not encounter prolonged prediction times on the test set despite its complex network structure.

Taking CL-C corresponding training samples from Table 2 as an example, we extracted the intermediate layer output of the CNN in the proposed hybrid algorithm to validate its feature extraction capabilities. Figure 9 illustrates the Pearson correlation coefficients calculated between the output of the two CNN intermediate layers (referred to as CONV1 and CONV2, respectively) and the training samples. From Figure 9, it can be observed that the feature dimension of the training samples is 6. After the first convolutional layer, the dimension is reduced to 4, demonstrating that the convolutional layer can decrease the input dimension through convolutional operations. One-dimensional convolution is commonly regarded as a linear transformation. Therefore, using a convolutional kernel of size 3 in the first convolutional layer, each output extracts different feature information. For example, Output 1 extracts DOY and *Lat*; Outputs 2 and 3 extract spatial position information of the input data, and Output 4 extracts T_s and e_s while considering their variations with *Hs*. In the second convolutional layer, the feature dimension of the output was further compressed and reduced to 2, concurrently extracting important features from the input data. As depicted in Figure 9b, it is evident that the inputs of the second convolutional layer are most correlated with DOY, *Lat*, T_s , and e_s , consistent with the findings in Figure 7. This substantiates that the convolutional layer in the proposed algorithm effectively reduces the input data dimensionality and extracts spatial features through convolutional operations.

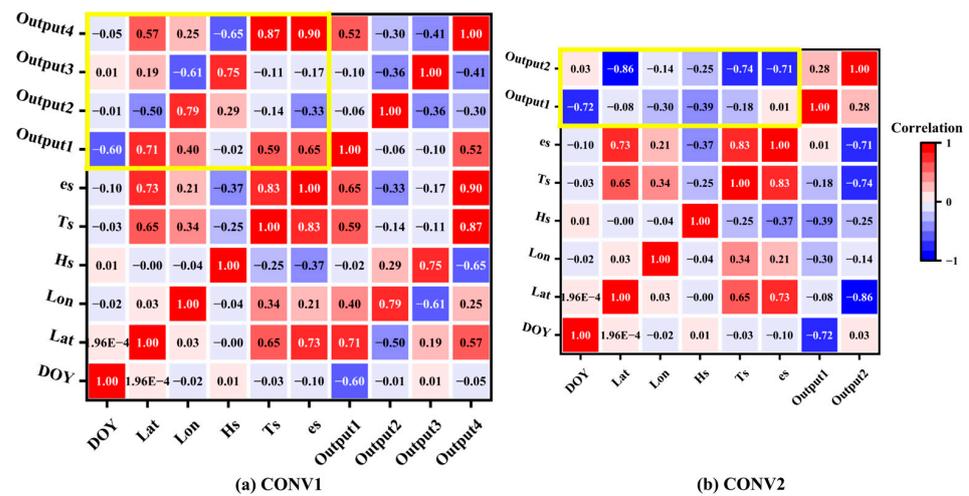


Figure 9. The Pearson correlation coefficients between input data and CNN intermediate layer outputs (The yellow box indicates the extraction ability of the CNN intermediate layer to capture important features of the input data). (a) for the first convolutional layer (CONV1); (b) for the second convolutional layer (CONV2).

4. Accuracy Analysis and Evaluation

4.1. Performance Analysis

Based on the above modeling strategies, three ZWD models are constructed by the hybrid algorithm, and the 46-radiosonde data from 2015–2021 are used to train the models and optimize the hyperparameters. The performance of the above models and other models is verified using the data from 2022. The joint probability density and evaluation indexes are shown in Figure 10.

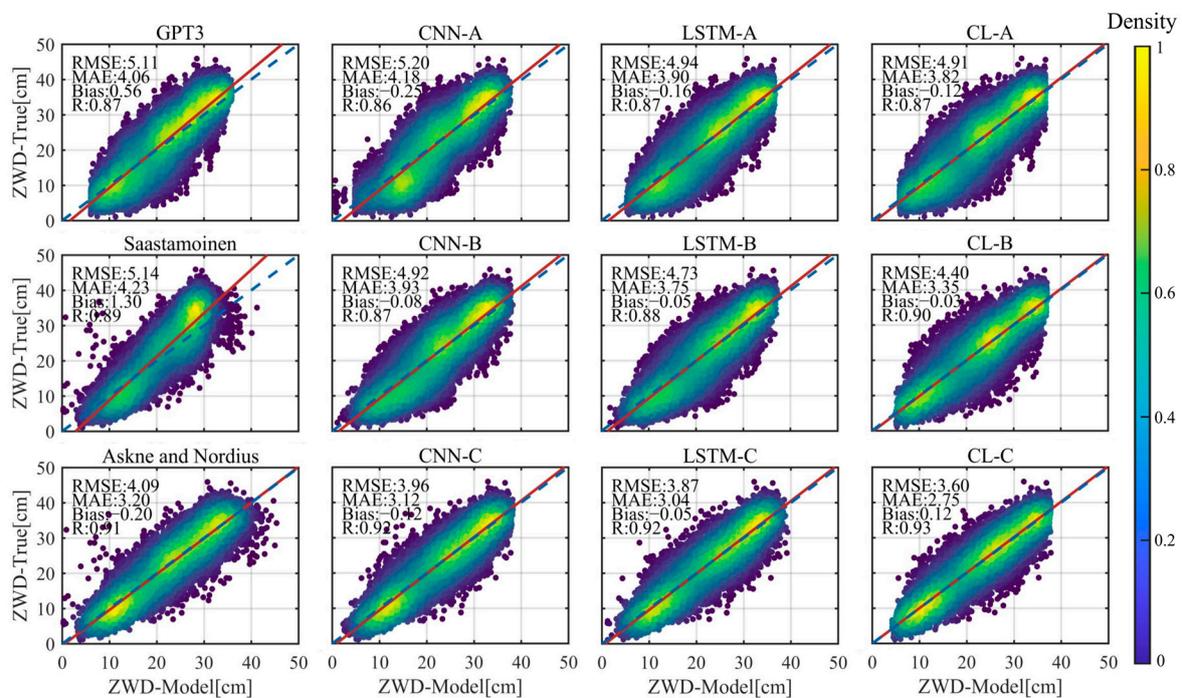


Figure 10. Comparison of the forecasting accuracy of the different models in 2022 (The dashed and solid lines represent the theoretical and practical correlation trends between ZWD-Model and ZWD-True, respectively).

Figure 10 shows that the bias of the models established by the deep learning algorithms under different strategies are millimeter level, and the model values are evenly distributed, significantly weakening the systematic bias compared to the traditional models. Figure 10 reveals that in South America, the CL-C model established with surface temperature and surface water vapor pressure performs better than other models, with an RMSE of 3.60 cm and a Bias of 0.12 mm. In comparison to the Saastamoinen model and the AN model, the CL-C model demonstrates an accuracy improvement of 30% and 12%, respectively. The statistical results show that the ZWD modeling performance of the hybrid algorithm exhibits a certain degree of improvement under three different strategies when compared to the CNN algorithm and LSTM algorithm, as well as the traditional models. RMSE exhibits varying degrees of improvement with different surface parameters can be accessed, as shown in Table 5.

Table 5. Statistical results of three combined models and the other model at 46 radiosonde sites in 2022.

Model	RMSE/cm	Improvement in GPT3	Improvement in Saastamoinen	Improvement in AN
GPT3	5.11	/	/	/
CNN-A	5.20	−2%	/	/
LSTM-A	4.94	3%	/	/
CL-A	4.91	4%	/	/
Saastamoinen	5.14	/	/	/
CNN-B	4.92	4%	4%	/
LSTM-B	4.73	7%	8%	/
CL-B	4.40	14%	14%	/
Askne and Nordius	4.09	/	/	/
CNN-C	3.96	23%	23%	3%
LSTM-C	3.87	24%	25%	5%
CL-C	3.60	30%	30%	12%

To explore the spatial distribution of the modeling accuracy, the calculated results of the accuracy of each model are shown in Figure 11, with the ZWD by integrating radiosonde as a reference.

Figure 11 illustrates that the CL-A model optimizes the ZWD estimation accuracy of individual offshore sites in northern South America and is slightly better than the CNN-A, LSTM-A, and GPT3 models. The empirical models mentioned earlier exhibit accuracy better than 5 cm in the northern and western regions of South America. However, their performance diminishes in the eastern region (20° S–30° S) with an RMSE of about 7 cm, likely influenced by the subtropical monsoon climate and sea winds in this area. The empirical models make it difficult to capture the complex changes in short-term non-trend of water vapor in monsoon climate regions, resulting in the accuracy of these models decreasing in specific regions.

After introducing the temperature parameter, the CL-B model shows a slight improvement in accuracy, reducing the RMSE in the eastern monsoon area to less than 6 cm. After reintroducing the water vapor pressure parameter, the CL-C model significantly reduces the RMSE to 3.5 cm in the subtropical monsoon climate area, where the RMSE of the empirical model and the AN model are 7 cm and 5 cm, respectively. It shows that the surface water vapor pressure has a significant effect on ZWD modeling. Additionally, even though the Saastamoinen model considers the water vapor pressure, it still exhibits a large error in the ZWD results for northern South America, with RMSE ranging from 5.5 cm to 7.5 cm. This discrepancy may be attributed to the presence of the Amazon Rain Forest in the northern region, where temperature and humidity remain high throughout the year and seasonal variations are less pronounced. These conditions challenge the

Saastamoinen model in accurately reflecting the spatial distribution of ZWD in this area, resulting in limited modeling capability for regions in the low latitudes, high temperatures, and frequent rainfalls [21,57].

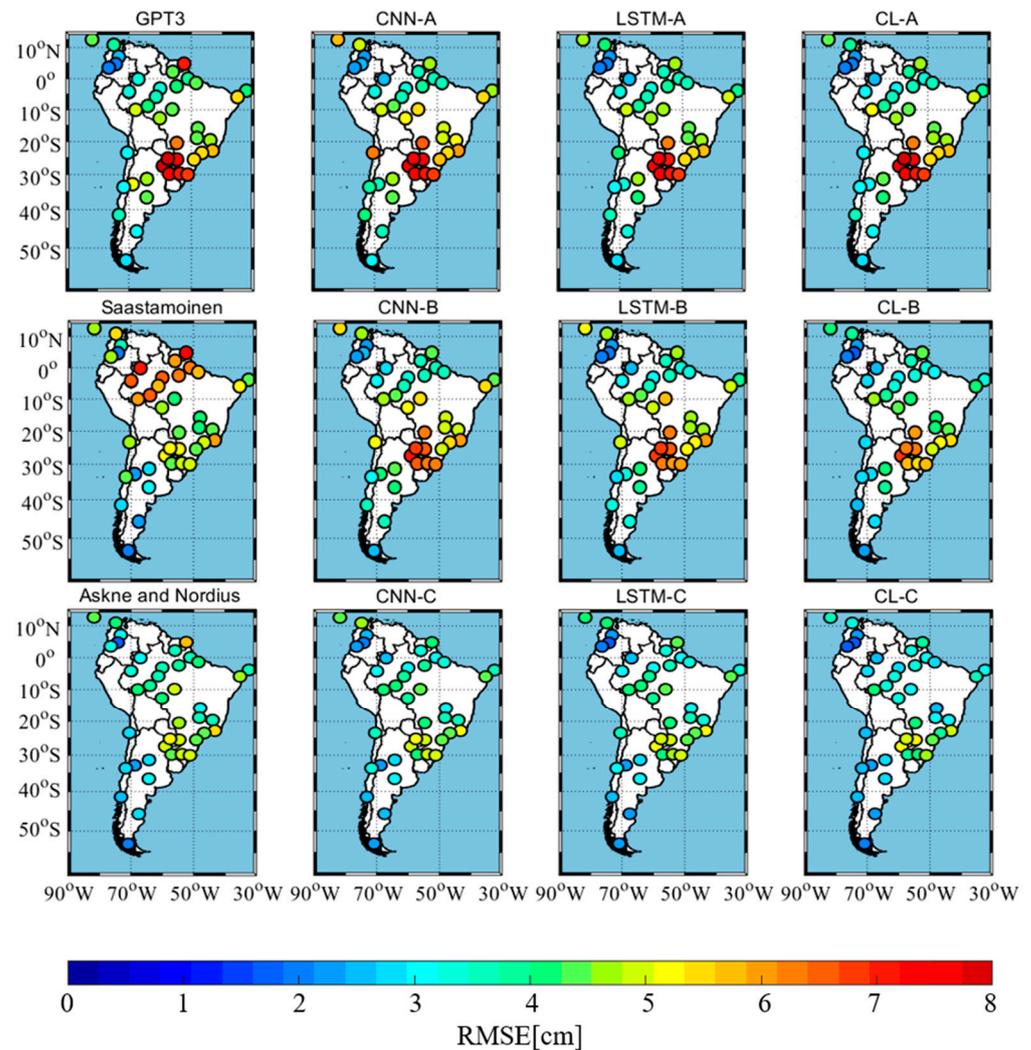


Figure 11. RMSE distribution of ZWD between models and reference for 46 sites in 2022.

Considering the distribution of sites and regional climate differences in South America, the modeling effects of different models in different latitude ranges are evaluated. The RMSE is used as the evaluation index, as shown in Table 6.

Table 6 illustrates that the combined ZWD models established based on the hybrid algorithm have different accuracy improvements compared to the models established based on CNN and LSTM and the traditional models at various latitudes. Notably, in the 15°N – 15°S region, the accuracy of the CL-A model is improved by 12% compared to the GPT3 model, and the accuracy of the CL-B model is improved by 38% compared to the Saastamoinen model, and the CL-C model demonstrates remarkable accuracy improvements of 44% and 19% compared to the Saastamoinen model and AN model, respectively. In the region of 30°S – 60°S , the water vapor changes are more stable, and the meteorological parameter models can express the nonlinear change in ZWD to a certain extent; the accuracy improvements of the combined models under the three strategies are lower than that in the low latitude area. However, the RMSE is still increased by 4–22%. In the region of 15°S – 30°S , it is difficult to express the change in ZWD accurately by the CL-B model due to the influence of the subtropical monsoon climate in southeastern South America. In this region, the accuracy of CL-B is lower than that of Saastamoinen, and the accuracy

of CL-A is comparable to GPT3. When the surface water vapor pressure is introduced in the CL-C model, it achieves 14% and 7% better than the Saastamoinen model and AN model, respectively.

Table 6. RMSE [cm] of different models in different latitude ranges.

Model	RMSE		
	15° N–15° S	15° S–30° S	30° S–60° S
GPT3	4.36	6.02	3.54
CNN-A	4.23	6.25	3.86
LSTM-A	3.94	6.02	3.43
CL-A	3.84	6.04	3.34
Saastamoinen	5.82	4.75	3.13
CNN-B	4.22	5.84	3.51
LSTM-B	3.96	5.67	3.21
CL-B	3.62	5.33	2.94
Askne and Nordius	4.03	4.38	2.56
CNN-C	3.72	4.35	2.79
LSTM-C	3.61	4.30	2.71
CL-C	3.26	4.09	2.45

In order to further explore the accuracy variations in the ZWD models based on the hybrid algorithm in the time domain, the seasonal influence is analyzed for the modeling performance. The monthly average RMSE is calculated for different models in 2022, and the results are shown in Figure 12.

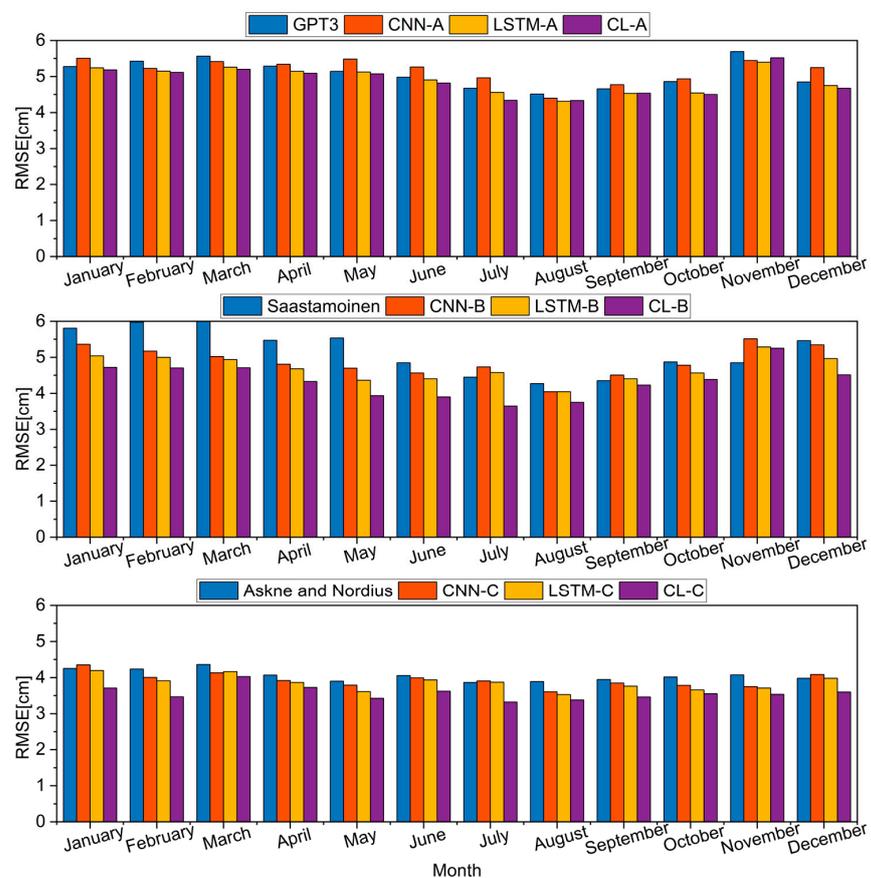


Figure 12. Monthly mean RMSE of different models in 2022.

Figure 12 shows the monthly average RMSE distribution of each ZWD model in 2022. The accuracy of the CL-A model is slightly better than that of the GPT3 model, and the accuracy in the time domain is more stable than that of the CNN-A model. When the meteorological parameters are introduced, the Saastamoinen model has a large error in the first half of the year compared to the GPT3 model and AN model, and the RMSE is greater than 5 cm. It shows obvious seasonal terms throughout the year, which may be caused by the inaccuracies in modeling ZWD in northern South America. The AN model uses the parameter of the exponential decay vertical trend of the water vapor over the site, and the RMSE is improved to 4 cm compared to the Saastamoinen model for the whole year. Compared to the above models, the CL-B and CL-C models can fully take into account the spatiotemporal characteristics and the annual RMSE of 3.6–5.2 cm and 3.3–4 cm, respectively. The CL-C model primarily enhances ZWD accuracy from January to May and from August to December, mitigating the impact of the rainy season changes in South America [58]. It can effectively identify the spatiotemporal characteristics of ZWD and strengthen the modeling ability of ZWD non-trend terms using the surface water vapor pressure. Compared to the ZWD models established by CNN and LSTM, the accuracy of CL-B and CL-C has improved by 5–23% and 5–20% within several months, respectively, which proves that the CL hybrid algorithm can effectively alleviate the influence of seasonal changes and enhance ZWD modeling abilities.

4.2. Generalization Assessment

The performance of the established models at modeling sites was validated using data from 46 sites in 2022, as detailed in Section 4.1. However, beyond these sites, it is essential to consider the generalization capability of the models to other locations in the South American region. Therefore, six non-modeled sites in South America, as illustrated in Figure 4, were selected, and the models were applied to predict their ZWD values using the data from the year 2022. Table 7 shows statistics of accuracy for these models in comparison with reference value over the year 2022.

Table 7. Statistical results of three combined models and the other models at 6 radiosonde sites in 2022.

Model	RMSE/cm	Improvement in GPT3	Improvement in Saastamoinen	Improvement in AN
GPT3	5.32	/	/	/
CNN-A	5.51	−4%	/	/
LSTM-A	5.26	1%	/	/
CL-A	5.24	2%	/	/
Saastamoinen	6.13	/	/	/
CNN-B	5.19	2%	15%	/
LSTM-B	5.09	4%	17%	/
CL-B	4.84	9%	19%	/
Askne and Nordius	4.59	/	/	/
CNN-C	4.30	19%	30%	6%
LSTM-C	4.33	19%	29%	6%
CL-C	4.13	22%	33%	10%

Table 7 shows that the ZWD models based on the hybrid algorithm have better generalization ability compared to the ZWD models based on CNN and LSTM under the three strategies and traditional models. The accuracy of the CL-A model is slightly better than that of the LSTM-A and GPT3 models, while the accuracy of the CNN-A model is lower than that of the above models. When the surface meteorological parameters can be obtained, the generalization ability of the CL-B model is stronger than that of the Saastamoinen model, and the accuracy is improved by 19%. The CL-C model achieves 33%

and 10% better accuracy than the Saastamoinen model and AN model, respectively. This demonstrates that the proposed hybrid algorithm is more effective for ZWD modeling than the CNN algorithm, the LSTM algorithm, and traditional models. Table 8 and Figure 13 show the locations of six non-modeled sites and the RMSE of different models, respectively.

Table 8. Location of six non-modeled sites.

Station	Lat (°)	Lon (°)	H (m)
82022	2.83	−60.70	124.51
82705	−7.58	−72.77	193.09
83554	−19.00	−57.67	156.36
83899	−27.67	−48.55	5.73
87576	−34.82	−58.53	36.27
87715	−38.95	−68.13	288.64

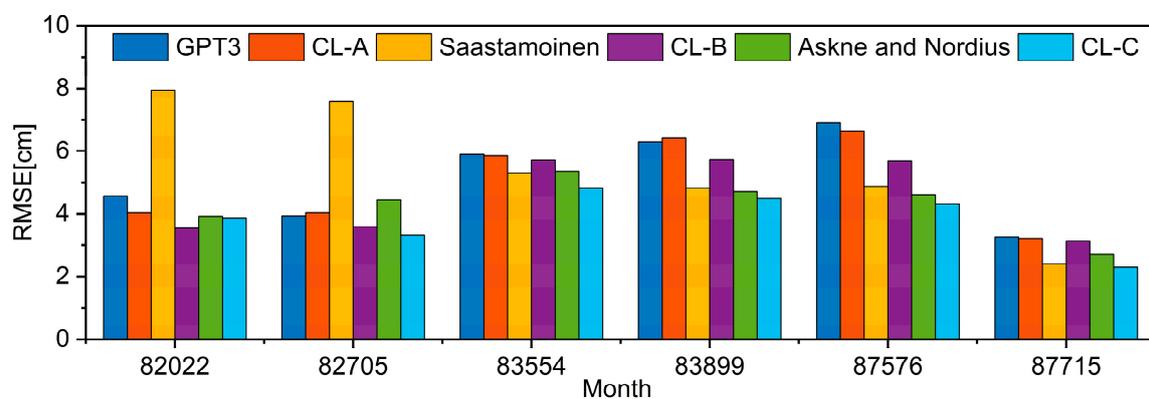


Figure 13. Monthly mean RMSE for estimated ZWD in 2022 for ZWD models based on the hybrid algorithm and traditional models at non-modeled sites.

Figure 13 shows that the accuracy of the CL-A model is comparable to the GPT3 model in most of the six sites, with a slight improvement for the low-latitude site (82022) and the monsoon climate region site (87576). The Saastamoinen model still exhibits RMSE of more than 7 cm in the northern sites (82022, 82705). The CL-C model generally shows higher generalization ability than the Saastamoinen model and AN model in the six sites, especially in the monsoon climate area. The proposed algorithm enhances the generalization ability better than the traditional models of the ZWD modeling in areas with low latitudes, specific climates, and complex terrains.

5. Conclusions

This paper introduces a novel spatiotemporal modeling algorithm, denoted as CNN-LSTM (CL), which integrates both CNN and LSTM algorithms. This algorithm is designed to capture the spatiotemporal characteristics of tropospheric Zenith Wet Delay (ZWD) sequences. The proposed algorithm utilizes an encoding–decoding framework, where CNN is responsible for encoding input parameters and extracting local spatial features, thereby constructing spatiotemporal features sequence. LSTM is employed to capture long-term spatiotemporal features and decode the output to estimate ZWD. This algorithm facilitates the recognition and modeling of both spatial and temporal characteristics of ZWD, which are crucial for accurate modeling. The research explores three strategies and investigates the influence of surface meteorological parameters on ZWD modeling, considering correlation analysis and diverse scenario requirements. This paper is conducted using a 7-year dataset from radiosonde in South America, covering the years from 2015 to 2021, with data from 2022 serving as a validation set. The accuracy of the proposed ZWD models established under the same modeling strategy is assessed in comparison

to the CNN algorithm and LSTM algorithm, as well as traditional models like the GPT3, Saastamoinen, and AN models. This aims to evaluate the accuracy and effectiveness of the proposed hybrid ZWD modeling algorithm.

The validation results of the 46 sites in South America for the year 2022 indicate that the overall accuracy of the CL-A model established based on the CL hybrid algorithm without surface meteorological parameters is slightly better than the GPT3 model, but the accuracy improvement in the 15° N–15° S region is 12%. The CL-B model, with the introduction of surface temperature, solves the problem of poor applicability of the Saastamoinen model in the Amazon Rain Forest area in northern South America, and the accuracy is improved by 38% in the area of 15° N–15° S. Reintroducing the surface water vapor pressure, the overall RMSE of ZWD estimated by the CL-C model is 3.60 cm, improved by 30% and 12%, respectively, compared to the Saastamoinen model and AN model. The CL-C model can significantly improve the performance by introducing the surface water vapor pressure, achieving approximately 44% and 19% improvements compared to the Saastamoinen and AN model in the 15° N–15° S region and 17% and 7% improvements in the 15° S–30° S region, respectively. The CL-C model effectively weakens the adverse effects of monsoon climate regions. Compared to the ZWD models established using the CNN algorithm and LSTM algorithm, the proposed hybrid algorithm performs better than the others.

To validate the performance of the proposed models at non-modeled sites in South America, data from an additional six sites for the year 2022 were selected for assessment. Results show that the ZWD models established using the CL hybrid algorithm exhibit strong generalization ability under the three strategies and reliability at different latitudes and climatic regions. They can effectively improve the overall accuracy of existing empirical models and meteorological parameter models in low latitudes and specific climate regions. Under the three strategies, the established models based on the proposed hybrid algorithm can improve the accuracy of the corresponding empirical and meteorological parameter models and describe the spatiotemporal variation in ZWD properly.

The hybrid deep learning algorithm, utilizing long-term historical data with multiple parameters as input, effectively captures the complex nonlinear relationships between input and output to establish ZWD models. This algorithm demonstrates reliable applicability and accuracy in South America, showcasing excellent performance at both modeling and non-modeling sites within the region. In this paper, South America is considered the research area, and the modeling and generalization capabilities of the proposed models are validated within this region. Expanding the modeling region by incorporating data from stations worldwide could be considered as the next research work.

Author Contributions: Conceptualization, Y.W. and W.F.; methodology, Y.W. and W.F.; software, Y.W.; validation, Y.W.; visualization, Y.W.; writing—original draft, Y.W.; writing—review and editing, L.H., W.F. and S.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research is funded by the National Natural Science Foundation of China (Grant number 42171429).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The IGRA radiosonde data presented in this study are openly available in the National Centers for Environmental Information/National Oceanic and Atmospheric Administration (NCEI/NOAA) at <https://www.ncei.noaa.gov/pub/data/igra/>, accessed on 15 October 2023.

Acknowledgments: The author would like to thank the IGRA radiosonde data provided by the NCEI/NOAA and GPT3 codes provided by the Vienna University of Technology. This study is supported by the National Natural Science Foundation of China.

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

References

1. Rózsa, S.; Ambrus, B.; Juni, I.; Ober, P.B.; Mile, M. An advanced residual error model for tropospheric delay estimation. *GPS Solut.* **2020**, *24*, 103. [[CrossRef](#)]
2. Sun, L.; Chen, P.; Wei, E.; Li, Q. Global model of zenith tropospheric delay proposed based on EOF analysis. *Adv. Space Res.* **2017**, *60*, 187–198. [[CrossRef](#)]
3. Chen, P.; Ma, Y.; Liu, H.; Zheng, N. A new global tropospheric delay model considering the spatiotemporal variation characteristics of ZTD with altitude coefficient. *Earth Space Sci.* **2020**, *7*, e2019EA000888. [[CrossRef](#)]
4. Hofmann-Wellenhof, B.; Lichtenegger, H.; Collins, J. *Global Positioning System: Theory and Practice*; Springer Science & Business Media: Vienna, Austria, 2012.
5. Mendes, V. *Modeling the Neutral-Atmospheric Propagation Delay in Radiometric Space Techniques*; UNB Geodesy and Geomatics Engineering Technical Report No. 199; University of New Brunswick: Fredericton, NB, Canada, 1999.
6. Chen, B.; Liu, Z. A comprehensive evaluation and analysis of the performance of multiple tropospheric models in China region. *IEEE Trans. Geosci. Remote Sens.* **2015**, *54*, 663–678. [[CrossRef](#)]
7. Ifadis, I. Space to earth techniques: Some considerations on the zenith wet path delay parameters. *Surv. Rev.* **1993**, *32*, 130–144. [[CrossRef](#)]
8. Li, Z.; Ding, X.; Chen, W.; Liu, G.; Shea, Y.; Emerson, N. Comparative study of empirical tropospheric models for the Hong Kong region. *Surv. Rev.* **2008**, *40*, 328–341. [[CrossRef](#)]
9. Li, X.; Dick, G.; Lu, C.; Ge, M.; Nilsson, T.; Ning, T.; Wickert, J.; Schuh, H. Multi-GNSS meteorology: Real-time retrieving of atmospheric water vapor from BeiDou, Galileo, GLONASS, and GPS observations. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 6385–6393. [[CrossRef](#)]
10. Lu, C.; Zheng, Y.; Wu, Z.; Zhang, Y.; Wang, Q.; Wang, Z.; Liu, Y.; Zhong, Y. TropNet: A deep spatiotemporal neural network for tropospheric delay modeling and forecasting. *J. Geod.* **2023**, *97*, 34. [[CrossRef](#)]
11. Shi, J.; Xu, C.; Guo, J.; Gao, Y. Local troposphere augmentation for real-time precise point positioning. *Earth Planets Space* **2014**, *66*, 30. [[CrossRef](#)]
12. Lu, C.; Li, X.; Zus, F.; Heinkelmann, R.; Dick, G.; Ge, M.; Wickert, J.; Schuh, H. Improving BeiDou real-time precise point positioning with numerical weather models. *J. Geod.* **2017**, *91*, 1019–1029. [[CrossRef](#)]
13. de Oliveira, P., Jr.; Morel, L.; Fund, F.; Legros, R.; Monico, J.; Durand, S.; Durand, F. Modeling tropospheric wet delays with dense and sparse network configurations for PPP-RTK. *GPS Solut.* **2017**, *21*, 237–250. [[CrossRef](#)]
14. Guangwei, J.; Panlong, W.; Chunxi, G.; Bin, W.; Yuanxi, Y. Short-term GNSS network solution and performance in large height difference region with tropospheric delay constraint. *Acta Geod. Cartogr. Sin.* **2022**, *51*, 2255. [[CrossRef](#)]
15. Collins, J. *Assessment and Development of a Tropospheric Delay Model for Aircraft Users of the Global Positioning System*; National Library of Canada/Bibliothèque Nationale du Canada: Ottawa, ON, Canada, 2001.
16. Böhm, J.; Heinkelmann, R.; Schuh, H. Short note: A global model of pressure and temperature for geodetic applications. *J. Geod.* **2007**, *81*, 679–683. [[CrossRef](#)]
17. Hopfield, H.S. Tropospheric effect on electromagnetically measured range: Prediction from surface weather data. *Radio Sci.* **1971**, *6*, 357–367. [[CrossRef](#)]
18. Saastamoinen, J. Atmospheric correction for the troposphere and stratosphere in radio ranging satellites. *Use Artif. Satell. Geod.* **1972**, *15*, 247–251. [[CrossRef](#)]
19. Black, H.D. An easily implemented algorithm for the tropospheric range correction. *J. Geophys. Res. Solid Earth* **1978**, *83*, 1825–1828. [[CrossRef](#)]
20. Askne, J.; Nordius, H. Estimation of tropospheric delay for microwaves from surface weather data. *Radio Sci.* **1987**, *22*, 379–386. [[CrossRef](#)]
21. Li, Q.; Yuan, L.; Jiang, Z. Modeling tropospheric zenith wet delays in the Chinese mainland based on machine learning. *GPS Solut.* **2023**, *27*, 171. [[CrossRef](#)]
22. Sun, J.; Wu, Z.; Yin, Z.; Ma, B. A simplified GNSS tropospheric delay model based on the nonlinear hypothesis. *GPS Solut.* **2017**, *21*, 1735–1745. [[CrossRef](#)]
23. Ding, W.; Qie, X. Prediction of Air Pollutant Concentrations via RANDOM Forest Regressor Coupled with Uncertainty Analysis—A Case Study in Ningxia. *Atmosphere* **2022**, *13*, 960. [[CrossRef](#)]
24. Tran, T.T.K.; Lee, T.; Kim, J.-S. Increasing neurons or deepening layers in forecasting maximum temperature time series? *Atmosphere* **2020**, *11*, 1072. [[CrossRef](#)]
25. Yang, F.; Meng, X.; Guo, J.; Yuan, D.; Chen, M. Development and evaluation of the refined zenith tropospheric delay (ZTD) models. *Satell. Navig.* **2021**, *2*, 21. [[CrossRef](#)]
26. Li, S.; Xu, T.; Jiang, N. Tropospheric Delay Modeling Based on Multi-source Data Fusion and Machine Learning Algorithms. In Proceedings of the China Satellite Navigation Conference (CSNC 2021) Proceedings, Nanchang, China, 26–28 May 2021; Springer Nature: Singapore, 2021; Volume 1, pp. 145–158.
27. Gao, W.; Gao, J.; Yang, L.; Wang, M.; Yao, W. A novel modeling strategy of weighted mean temperature in China using RNN and LSTM. *Remote Sens.* **2021**, *13*, 3004. [[CrossRef](#)]
28. Lu, C.; Zhang, Y.; Zheng, Y.; Wu, Z.; Wang, Q. Precipitable water vapor fusion of MODIS and ERA5 based on convolutional neural network. *GPS Solut.* **2023**, *27*, 15. [[CrossRef](#)]

29. Osah, S.; Acheampong, A.A.; Fosu, C.; Dadzie, I. Deep learning model for predicting daily IGS zenith tropospheric delays in West Africa using TensorFlow and Keras. *Adv. Space Res.* **2021**, *68*, 1243–1262. [[CrossRef](#)]
30. Ding, M. Developing a new combined model of zenith wet delay by using neural network. *Adv. Space Res.* **2022**, *70*, 350–359. [[CrossRef](#)]
31. Liu, G.; Huang, G.; Xu, Y.; Ta, L.; Jing, C.; Cao, Y.; Wang, Z. Accuracy evaluation and analysis of GNSS tropospheric delay inversion from meteorological reanalysis data. *Remote Sens.* **2022**, *14*, 3434. [[CrossRef](#)]
32. Xiao, X.; Lv, W.; Han, Y.; Lu, F.; Liu, J. Prediction of CORS Water Vapor Values Based on the CEEMDAN and ARIMA-LSTM Combination Model. *Atmosphere* **2022**, *13*, 1453. [[CrossRef](#)]
33. Long, F.; Hu, W.; Dong, Y.; Wang, J. Neural network-based models for estimating weighted mean temperature in China and adjacent areas. *Atmosphere* **2021**, *12*, 169. [[CrossRef](#)]
34. LeCun, Y.; Bengio, Y. Convolutional networks for images, speech, and time series. In *The Handbook of Brain Theory and Neural Networks*; MIT Press: Cambridge, MA, USA, 1995; Volume 3361.
35. Ding, Y.; Tian, X.; Yin, L.; Chen, X.; Liu, S.; Yang, B.; Zheng, W. Multi-scale relation network for few-shot learning based on meta-learning. In Proceedings of the International Conference on Computer Vision Systems, Thessaloniki, Greece, 23–25 September 2019; pp. 343–352. [[CrossRef](#)]
36. Zhang, Z.; Tian, J.; Huang, W.; Yin, L.; Zheng, W.; Liu, S. A haze prediction method based on one-dimensional convolutional neural network. *Atmosphere* **2021**, *12*, 1327. [[CrossRef](#)]
37. Albawi, S.; Mohammed, T.A.; Al-Zawi, S. Understanding of a convolutional neural network. In Proceedings of the 2017 International Conference on Engineering and Technology (ICET), Antalya, Turkey, 21–23 August 2017; pp. 1–6. [[CrossRef](#)]
38. Wei, Z.; Zhou, Z.; Wang, P.; Ren, J.; Yin, Y.; Pedersen, G.F.; Shen, M. Fast and Automatic 3D Modeling of Antenna Structure Using CNN-LSTM Network for Efficient Data Generation. *arXiv* **2023**, arXiv:2306.15530. [[CrossRef](#)]
39. Gu, J.; Wang, Z.; Kuen, J.; Ma, L.; Shahroudy, A.; Shuai, B.; Liu, T.; Wang, X.; Wang, G.; Cai, J. Recent advances in convolutional neural networks. *Pattern Recognit.* **2018**, *77*, 354–377. [[CrossRef](#)]
40. Tang, J.; Li, Y.; Ding, M.; Liu, H.; Yang, D.; Wu, X. An ionospheric TEC forecasting model based on a CNN-LSTM-attention mechanism neural network. *Remote Sens.* **2022**, *14*, 2433. [[CrossRef](#)]
41. Yilun, W.; Qingbin, L.; Yu, H.; Yajun, W.; Xuezhou, Z.; Yaosheng, T.; Chunfeng, L.; Lei, P. Deformation prediction model based on an improved CNN+ LSTM model for the first impoundment of super-high arch dams. *J. Civ. Struct. Health Monit.* **2023**, *13*, 431–442. [[CrossRef](#)]
42. Bengio, Y.; Courville, A.; Vincent, P. Representation learning: A review and new perspectives. *IEEE Trans. Pattern Anal. Mach. Intell.* **2013**, *35*, 1798–1828. [[CrossRef](#)]
43. Pascanu, R.; Mikolov, T.; Bengio, Y. On the difficulty of training recurrent neural networks. In Proceedings of the International Conference on Machine Learning, Atlanta, GA, USA, 16–21 June 2013; pp. 1310–1318. [[CrossRef](#)]
44. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735–1780. [[CrossRef](#)]
45. Rawat, W.; Wang, Z. Deep convolutional neural networks for image classification: A comprehensive review. *Neural Comput.* **2017**, *29*, 2352–2449. [[CrossRef](#)]
46. Ishida, K.; Ercan, A.; Nagasato, T.; Kiyama, M.; Amagasaki, M. Use of 1D-CNN for input data size reduction of LSTM in Hourly Rainfall-Runoff modeling. *arXiv* **2021**, arXiv:2111.04732. [[CrossRef](#)]
47. Zeiler, M.D.; Fergus, R. Visualizing and understanding convolutional networks. In Proceedings of the Computer Vision—ECCV 2014: 13th European Conference, Zurich, Switzerland, 6–12 September 2014; Proceedings, Part I 13. pp. 818–833. [[CrossRef](#)]
48. Yibin, Y.; Shun, Z.; Jian, K. Research progress and prospect of GNSS space environment science. *Acta Geod. Et Cartogr. Sin.* **2017**, *46*, 1408. [[CrossRef](#)]
49. Durre, I.; Vose, R.S.; Wuertz, D.B. Overview of the integrated global radiosonde archive. *J. Clim.* **2006**, *19*, 53–68. [[CrossRef](#)]
50. Alves, D.B.M.; Sapucci, L.F.; Marques, H.A.; de Souza, E.M.; Gouveia, T.A.F.; Magário, J.A. Using a regional numerical weather prediction model for GNSS positioning over Brazil. *GPS Solut.* **2016**, *20*, 677–685. [[CrossRef](#)]
51. 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. Meteorol. (1988–2005)* **1994**, *33*, 379–386. [[CrossRef](#)]
52. Rüeger, J.M. *Refractive Indices of Light, Infrared and Radio Waves in the Atmosphere*; School of Surveying and Spatial Information Systems, University of New South: Sydney, NSW, Australia, 2002.
53. Wang, X.; Zhang, K.; Wu, S.; Fan, S.; Cheng, Y. Water vapor-weighted mean temperature and its impact on the determination of precipitable water vapor and its linear trend. *J. Geophys. Res. Atmos.* **2016**, *121*, 833–852. [[CrossRef](#)]
54. Pavlis, N.; Kenyon, S.; Factor, J.; Holmes, S. Earth gravitational model 2008. In *SEG Technical Program Expanded Abstracts 2008*; Society of Exploration Geophysicists: Las Vegas, NV, USA, 2008; pp. 761–763. [[CrossRef](#)]
55. Zheng, D.; Hu, W.; Wang, J.; Zhu, M. Research on regional zenith tropospheric delay based on neural network technology. *Surv. Rev.* **2015**, *47*, 286–295. [[CrossRef](#)]
56. Liu, Y.H.; Maldonado, P. *R Deep Learning Projects: Master the Techniques to Design and Develop Neural Network Models in R*; Packt Publishing Ltd.: Birmingham, UK, 2018.

-
57. Yang, F.; Guo, J.; Meng, X.; Li, J.; Zou, J.; Xu, Y. Establishment and assessment of a zenith wet delay (ZWD) augmentation model. *GPS Solut.* **2021**, *25*, 148. [[CrossRef](#)]
 58. Chou, S.; Bustamante, J.; Gomes, J. Evaluation of Eta Model seasonal precipitation forecasts over South America. *Nonlinear Process. Geophys.* **2005**, *12*, 537–555. [[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.