# A New Global ZTD Forecast Model Based on Improved LSTM Neural Network

Lin He, Yibin Yao, Chaoqian Xu, Huan Zhang, Feifei Tang, Changquan Ji, Zhuoya Liu, and Wentan Wu

***Abstract*—Zenith tropospheric delay (ZTD), consisting of zenith hydrostatic delay (ZHD) and zenith wet delay (ZWD), is a significant contributor to errors in precise positioning using the global navigation satellite system (GNSS) precise point positioning (PPP) and real-time kinematic techniques. Accurate and timely predictions of ZTD on a global scale are crucial for enhancing GNSS positioning accuracy and expediting convergence. This study proposes an innovative global tropospheric prediction model that leverages long short-term memory (LSTM) neural networks, aiming to achieve both high precision and long-term prediction capability for ZTD. The experimental data utilized were sourced from the Vienna Mapping Functions 3-Optimized zenith total delay (ZTD) dataset. This study delves further into the analysis of ZTD residuals by extracting periodic signals. The ZTD residuals were then utilized to train a modified LSTM neural network model, enabling the prediction of global residuals. The final ZTD predictions were obtained by combining the modified LSTM ZTD residual forecast component with the ZTD periodic component. Our results demonstrate that the average root-mean-square error (RMSE) of the modified LSTM-ZTD model in 2020 was 1.44 cm. In addition, the average RMSE of the forecasted ZTD during spring, summer, autumn, and winter was found to be 1.43 cm, 1.47 cm, 1.56 cm, and 1.36 cm, respectively. Through the integration of the LSTM neural network and the ZTD periodic signal extracted using a physical algorithm, this work has successfully enhanced the accuracy and time span of ZTD forecasts on a global scale.**

***Index Terms*—Global navigation satellite system (GNSS), long short-term memory (LSTM), zenith tropospheric delay.**

## I. Introduction

THE troposphere, the atmospheric layer closest to the Earth's surface, constitutes approximately 75% of the total air mass. It plays a crucial role in maintaining the global atmospheric radiation balance, energy balance, and water cycle. In addition, it serves as a key meteorological parameter for studying global climate change and understanding the mechanisms behind extreme weather events. As a nondispersive medium, the troposphere introduces nondispersive refraction when global navigation satellite system (GNSS) signals pass through it. This type of refraction remains consistent across signal frequencies, as supported by studies conducted by Bevis et al. [3], Jin et al. [18], and Rocken et al. [33]. In the processing of GNSS positioning data solutions, it is necessary to account for the tropospheric delay, which occurs as the GNSS signal propagates through the atmosphere. This delay is accurately mapped in the zenith direction using a mapping function (MF). The resulting zenith tropospheric delay (ZTD) is approximately 2.5 m and exhibits variations in relation to latitude and altitude. Notably, when the local horizontal altitude angle reaches 10°, the oblique tropospheric delay can reach up to 20 m. As such, this delay must be corrected to ensure accurate positioning results. The ZTD represents a significant source of error in various radio geodetic techniques, such as GNSS, satellite laser ranging, and very long baseline interferometry. From a formation perspective, the ZTD can further be divided into two components: the zenith hydrostatic delay (ZHD), which depends on air pressure and temperature, and the zenith wet delay (ZWD), which relies on water vapor pressure and temperature [13]. Furthermore, the tropospheric delay is strongly influenced by weather conditions and exhibits spatial variability across different station locations. [7], [18]. Hence, a reliable priori zenith tropospheric delay (ZTD) is crucial for achieving precise GNSS positioning. The accuracy of the a priori ZTD has a substantial impact on the quality of GNSS positioning results, particularly in the vertical direction.

The estimation of tropospheric delay in numerical weather models (NWM) relies on the direct application of physical algorithms utilizing essential parameters such as air temperature, air pressure, water vapor pressure, and others. Prominent examples

Manuscript received 30 October 2023; revised 27 December 2023, 16 February 2024, and 14 March 2024; accepted 16 April 2024. Date of publication 22 April 2024; date of current version 9 May 2024. This work was supported in part by the Chongqing Natural Science Foundation Innovation and Development Joint Fund Project under Grant Z2302230017, in part by Chongqing Jiaotong University Talent Funds under Grant F1230102, in part by the Special Key Project for Technological Innovation and Application Development in Chongqing under Grant CSTB2022TIAD-KPX0098 and Grant CSTB2022TIAD-CUX0016, and in part by the National Natural Science Foundation of China under Grant Number 42388102. *(Corresponding author: Lin He.)*

Lin He is with the School of Geodesy and Geomatics, Wuhan University, Wuhan 430079, China, and also with the Guizhou Panjiang Coal and Electricity Group Technology Research Institute, Company Ltd., Guiyang 550081, China (e-mail: hunter5288@163.com).

Yibin Yao is with the School of Geodesy and Geomatics, Wuhan University, Wuhan 430072, China, and also with the Key Laboratory of Geospace Environment and Geodesy, Ministry of Education, Wuhan University, Wuhan 430079, China (e-mail: ybyao@whu.edu.cn).

Chaoqian Xu is with the Hubei Luojia Laboratory, Wuhan 430079, China, and also with the School of Geodesy and Geomatics, Wuhan University, Wuhan 430079, China (e-mail: cqxu@whu.edu.cn).

Huan Zhang is with the China Merchants Chongqing Communications Research and Design Institute, Company Ltd., Chongqing 400711, China, and also with the Chongqing Jiaotong University, Chongqing 400074, China (e-mail: johnhuan@whu.edu.cn).

Feifei Tang and Changquan Ji are with the Chongqing Jiaotong University, Chongqing 400074, China (e-mail: fftang80@126.com).

Zhuoya Liu is with the Electrical Science Institute of Guizhou Power Grid, Company Ltd., CSG, Guiyang 550000, China (e-mail: liuzhuoya_scu@163.com).

Wentan Wu is with the Natural Resources Archives of Hebei Province, Shijiazuang 050000, China (e-mail: 568511035@qq.com).

Digital Object Identifier 10.1109/JSTARS.2024.3391821

© 2024 The Authors. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License. For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

of such models include the UNB series models [9], [22], the EGNOS model [31], the TropGrid ZTD model [20], [38], and the GPT series model [6], [7], [12], [21]. These models employ sophisticated algorithms to accurately estimate tropospheric delay based on the physical characteristics of the atmosphere.

The IGGtrop series model (including GGtrop_ri, IGGtrop_Rh, The IGGtrop_SH) based on a 3-D grid adapted to the radial and zonal variations, which is more precise than the model based on latitude (e.g., UNB3, EGNOS, and UNB3m) and requires fewer parameters, reveals the annual and semiannual changes in ZTD [24], [25], [26]. The global zenith tropospheric delay (GZTD) series model is known for its high accuracy, typically achieving a precision level of approximately 3 cm. Compared to other models such as the EGNOS and UNB series, the GZTD model exhibits superior precision in estimating tropospheric delays at various locations worldwide [44], [45]. The improved tropospheric grid) model conducts a comprehensive analysis of multiple tropospheric parameters, including annual, semiannual, and daily variations. This model provides accurate estimates of temperature, pressure, weighted mean temperature (TM), and ZWD. Notably, the average root-mean-square error (RMSE) of this model is determined to be 3.73 cm, indicating its high precision in characterizing tropospheric conditions [46]. Empirical models for estimating tropospheric delay are typically developed based on meteorological parameters, average periodic changes in zenith tropospheric delay (ZTD), and weighted TM over multiple years. These models are designed to capture the repeating patterns of these parameters within each period, allowing for their application over extended periods of time. However, empirical models generally exhibit lower precision compared to other modeling approaches due to their reliance on historical data and simplifications of atmospheric processes.

In meteorological parameter models, the ZHD and ZWD are typically calculated independently. Classical tropospheric delay models such as Hopfield, Saastamoinen, and Black require access to precise measurements of meteorological parameters at the station in order to effectively estimate tropospheric delay. These models utilize empirically derived relationships between meteorological parameters and tropospheric delay to provide accurate estimates of the ZHD and ZWD components [2], [4], [8], [16], [17], [37]. Nonetheless, numerous studies conducted by researchers have demonstrated that the ZTD calculated using meteorological parameters from models, such as Hopfield, Saastamoinen, and others, does not exhibit any significant advantage over empirical models. In fact, in some cases, the performance of these physically derived models is even inferior to that of empirical models. These findings highlight the limitations and potential shortcomings of relying solely on meteorological parameter-based models for accurately estimating ZTD in various atmospheric conditions.

The GNSS tropospheric delay model utilizes observation data collected by GNSS receivers for the estimation and modeling of zenith tropospheric delay (ZTD). In order to calculate ZTD using GNSS data, it is necessary to first convert the oblique path delay into the zenith direction via a tropospheric delay MF. The development of tropospheric delay MF has progressed significantly over time, resulting in more mature and stable conversion methods for accurately estimating ZTD from GNSS data. As a result, the use of GNSS data for ZTD calculation and modeling has become increasingly reliable and effective.

Herring [15] first established an MF based on the measured atmosphere by fitting the observational data of ten North American sounding ball stations. The Niell mapping function (NMF) is based on observation data of radio sounding stations in the Northern Hemisphere [30]. In NMF, continued fraction coefficients are only related to the latitude, altitude, and day of the year (DOY) of the station. Many researchers have used NWM to establish projection functions, such as the IMF [29] and VMF1 [5], [19]. The VMF1 is the most accurate tropospheric MF and has been used by GAMIT, Bernese, and other high-precision GNSS data processing software. The LSTM-MF model established VTEC product improved the STEC accuracy obtained from low-elevation-angle conversion [43].

When double-difference precision positioning technology is used to estimate ZTD in real-time to eliminate the correlation between the tropospheric parameters of the stations and obtain a stable and high-precision estimation, GNSS reference stations beyond 500 km are necessary [34]. Nondifferential precision positioning technology is also called precise point positioning (PPP) technology, which can estimate high-precision tropospheric delay for a single station and has more advantages than double-difference technology in estimating tropospheric delay [47]. In terms of real-time PPP technology, the earliest real-time PPP experiments used the real-time orbit and clock products provided by the Jet Propulsion Laboratory (Pasadena, America), which have comparable accuracies to the International GNSS Service (IGS) fast orbit and clock products released with a postevent delay of approximately 17 h [1], [14]. Although the ZTD calculation using the GNSS observation data shows a higher modeling accuracy than the other two types of key parameters, it is expensive for modeling and data acquisition.

Machine learning has increasingly been used in tropospheric delay time series regression prediction. Pikridas et al. [32] used artificial neural networks to predict ZTD; the results of the test data source conducted at the Euref Permanent Network Analysis Center demonstrated that the proposed model can achieve centimeter-level accuracy. Mohammed [28] studied the prediction of temperature, air pressure, and water vapor pressure by the artificial neural network, their root mean square (RMS) was approximately 3.0 cm. Suparta and Alhasa [39] established a zenith path delay estimation and prediction model for the Antarctic region via the adaptive neuro-fuzzy inference system method. Zheng et al. [49] improved the ZTD prediction accuracy by combining the BPNN and Hopfield models, and their accuracy is approximately 4 mm. Ding et al. [10] employed IGS ZTD data to study the residuals in the Saastamoinen model; they proposed a ZTD forecast model based on ground meteorological parameters and BPNN modeling, and the RMSE was approximately 20.4 mm in Russia. Yang et al. [42] employed the UNB3m model to calculate the temperature, air pressure, and relative humidity of a local area and established a "UNB3m+GA-BP" regional tropospheric delay model based on the BPNN, verified by GNSS data in Hong Kong; the accuracy is approximately 1.1 cm. Xiao et al. [40] established a regional ZTD prediction model in Japan using the BPNN and improved the accuracy in fitting and forecasting ZTD; the RMS in fitting is 7.83 mm and 8.52 mm in

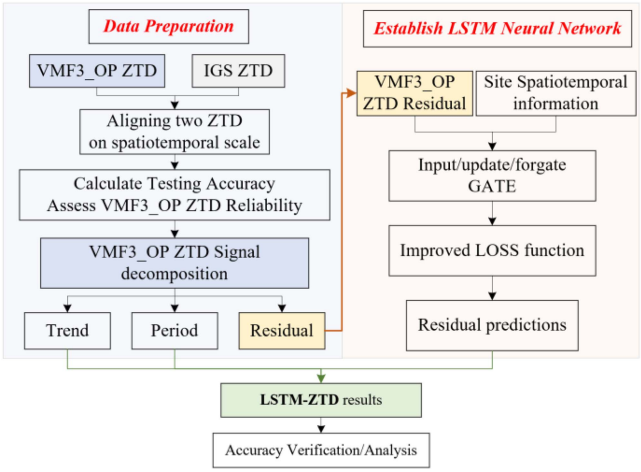


Fig. 1. Flowchart for developing the LSTM-ZTD model combined with ZTD decomposition and neural network.

forecasting. Zhang et al. [48] established a GNSS meteorological ZTD forecast model with GNSS stations in western Antarctica and two blind source separation algorithms, and the forecast accuracy is 7.2 mm for 6 h. Xu et al. [41] deduced the ZTD time-series variation trend with polynomial fitting, reconstructed and modeled the residual error by BPNN and least squares support vector machine (LSSVM) algorithm, and verified and compared spatial ZTD prediction based on BPNN and LSSVM with GNSS ZTD in Hong Kong. Their results demonstrated that the model accuracy depends on the season. Li et al. [23] used the National Center for Atmospheric Troposphere data in Japan and evaluated the accuracy of the troposphere based on the regional numerical forecast tropospheric delay inversion method of GRNN; the RMS is 12.7 mm.

The current tropospheric delay forecast models primarily emphasize short-term predictions within specific regions. Despite the rapid development of machine learning techniques, their full application in the field of geodesy is still in progress. However, the utilization of machine learning methods for modeling tropospheric delay products holds substantial scientific and practical value. Furthermore, extending the duration of tropospheric delay forecasts leads to a decrease in the frequency of ZTD updates, thereby enhancing convergence speed and positioning accuracy in PPP. It is crucial to improve the accuracy and timeliness of ZTD predictions in order to enhance PPP performance. The inherent spatial heterogeneity and temporal variability of atmospheric density, particularly related to water vapor, present significant challenges in accurately predicting and fitting ZTD.

To address these challenges, we sought to enhance the accuracy and duration of ZTD prediction utilizing machine learning techniques. Building upon previous research, we utilized a GPT2 empirical periodic model to extract periodic signals from the ZTD time series. We then trained an improved long short-term memory (LSTM) model on the residual series and extrapolated backward. Finally, the LSTM model's forecast residual was combined with the ZTD periodic signal to construct a global ZTD forecast model named LSTM-ZTD. The flowchart outlining the establishment of this model is depicted in Fig. 1.

The rest of the article is organized as follows. Section I provides an overview of the literature related to the tropospheric delay. Section II outlines the proposed model and residual data processing strategy. Section III evaluates the accuracy of the LSTM-ZTD model. Finally, Section IV concludes the article. Section V summarizes the findings and provides future perspectives.

## II. Dataset and Methods

The global geodetic observing system (GGOS) atmosphere provides meteorological products, contains ZHD and ZWD, the correlation coefficient of ray tracing delay, and atmospheric pressure. This work employed the global 5° × 5° microwave-range grid points Vienna Mapping Function 3 Operation products (VMF3_OP) version of ZTD time series data, which can be downloaded from the website (https://vmf.geo.tuwien.ac.at); the time resolution is 6 h. GGOS atmosphere provides daily observation data for four observation periods (00, 06, 12, and 18 UT) to estimate ZHD and ZWD at each grid point. These estimates are then combined to obtain zenith tropospheric delay (ZTD) predictions. To improve the accuracy of ZTD predictions, the periodic signal of ZTD was extracted using the GPT2 empirical period model expression and subsequently eliminated. The resulting residuals were treated as input parameters for training and extrapolating an LSTM neural network. By combining the extrapolated residuals with the ZTD periodic signals, the LSTM-ZTD model generated more accurate ZTD predictions.

The ZTD dataset was partitioned into two distinct sets: the first set encompassed the time period from 2008 to 2019 and was utilized for model development, while the second set consisted of data solely from the year 2020 and served as a means to assess and validate the accuracy of the model.

### A. Filtering the Periodic Signal

To generate empirical periodic ZTD signals that remained constant within every period, the ZTD datasets spanning from 2008 to 2019 were modeled using formula (2). Subsequently, the original ZTD sequences were subtracted from the ZTD periodic signals to obtain the residual ZTD sequences for the aforementioned time period

$$t = \frac{\mathrm{DOY}}{365.25} + \frac{HOD}{24} \tag{1}$$

$$\begin{aligned}\mathrm{ZTD}\ (\mathrm{t}) = A_0\ &+ A_1 \cos(2\pi \mathrm{t}) + B_1 \sin(2\pi t) \\ &+ A_2 \cos(4\pi t) + B_2 \sin(4\pi t)\end{aligned} \tag{2}$$

where DOY is the day of the year, HOD is the hour of the day, $A_0$ is the average ZTD, and ($A_1$, $B_1$) and ($A_2$, $B_2$) are the annual and semiannual amplitudes, respectively.

Fig. 2 depicts the spatial distribution of four representative stations randomly selected from the northern, southern, western, and eastern hemispheres in our experiment. The selection of IGS stations was random, whereas the GGOS grid had a spatial resolution of 5° × 5°, with the nearest grid points being chosen based on their proximity to IGS stations. All subsequent figures and tables in the following sections will be presented using these four representative stations.

Fig. 3 illustrates a comparison between the ZTD periodic signals, original ZTD, and residuals for four randomly selected grid points from both the Southern and Northern hemispheres.

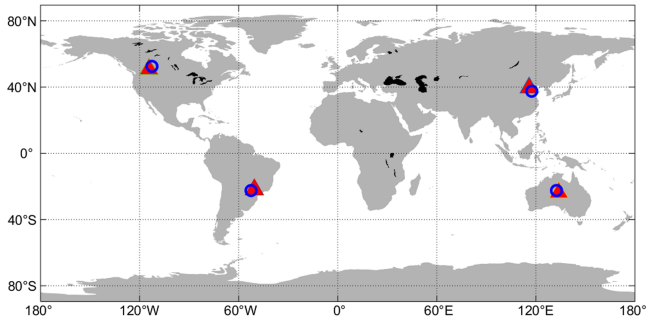


Fig. 2. Global spatial distribution map of representative GGOS (Hollow blue circles) and IGS (red triangles) stations.

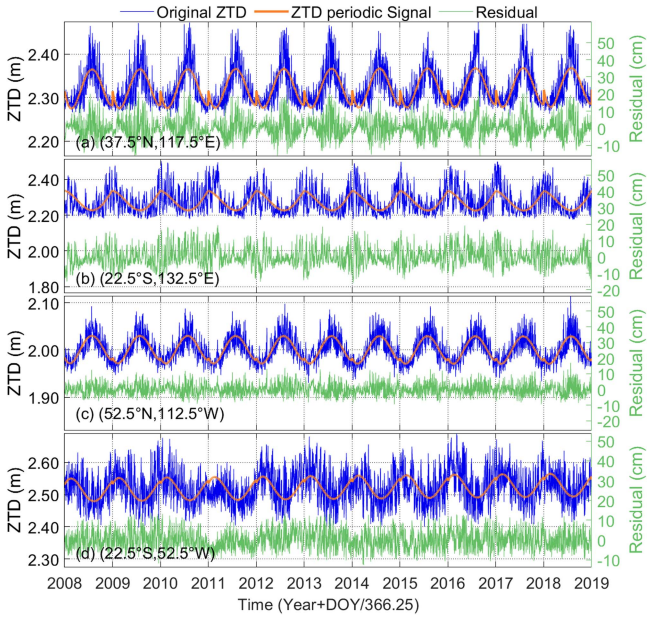


Fig. 3. Comparison of the ZTD periodic signals, original ZTD, and corresponding residuals.

Notably, the ZTD residual signals, obtained through the decomposition of the physical model described in (2), exhibit considerable magnitudes. These residuals display distinct patterns that are not readily identifiable using conventional methods. As a result, our subsequent objective involves training a neural network to capture these latent patterns by learning the residual terms associated with tropospheric delay. This endeavor aims to achieve a precise modeling of tropospheric delay with high accuracy.

### *B. Construction of the LSTM-ZTD Forecast Model*

In this section, the residuals were utilized as the training set, and modifications were made to the original LSTM neural network at various layers to forecast the residuals for each global grid in 2020. The predicted residuals were then integrated with the ZTD empirical periodic signals to generate forecast products using the LSTM-ZTD model. The original ZTD values from 2020 were employed as the validation set to assess the accuracy of the proposed model.

To mitigate overfitting when applying LSTM, several strategies were designed. One approach involves the utilization of dropout layers within the LSTM model architecture, which

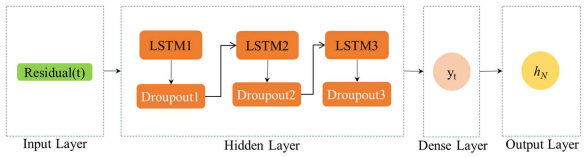


Fig. 4. Network structure of the LSTM neural network algorithm.

aids in preventing overfitting by randomly deactivating a proportion of neurons during each training iteration. In addition, we employed early stopping techniques based on validation performance to help prevent overfitting by halting the training process once the model's performance on a separate validation dataset ceases to improve. In addition, by adjusting the number of hidden units or layers, the complexity of the LSTM model can be effectively reduced and overfitting can be avoided.

The application of neural network regression algorithms in analyzing time series data shows great promise. Deep learning neural networks, with their multiple nonlinear mapping layers, effectively extract the features of input signals layer by layer, thereby uncovering the underlying patterns at deeper levels. Among these neural network architectures, the LSTM network has gained widespread adoption in the field of data mining. Its ability to control information flow through memory cells and gate control cells enables it to capture long-term dependencies within time series data by adjusting the weights of the forget gates [27], [36].

The network architecture of the LSTM-ZTD model is depicted in Fig. 4. It comprises several layers, including an input layer where the input data residual($t$) represent the sequence of residuals acquired by subtracting periodic signals from the original ZTD. A hidden layer is constructed with three LSTM layers (LSTM1, LSTM2, and LSTM3 as shown in Fig. 4), along with three dropout layers (Dropout1, Dropout2, and Dropout3 displayed in Fig. 4). The LSTM neural network is designed to extract informative features from the historical input data and connect to a fully connected layer (dense layer). In this layer, $y_t$ denotes the ZTD with filtered gross errors and noise, achieved through the recursive backward propagation of a sliding window mechanism facilitated by the dropout layer. The last one is the output layer; the output result $h_N$ is the ZTD residual predictions.

In the LSTM-ZTD model structure, the LSTM input space dimension was set to 360. The adaptive moment estimation algorithm was utilized as the optimizer, and the evaluation indicator for the loss function was the RMS. The activation function employed was tanh, while the recurrent activation function was sigmoid. A dropout rate of 0.01 was applied to each LSTM layer. The batch size used for training was 40, and a total of 10 epochs were conducted. The hyperparameters were determined through a tenfold cross-validation process [27]. The model was trained on the Centos Wuhan University supercomputing system using a CPU with a clock speed of 2580.175 MHz, 20 cores, and 128 GB of memory. The training duration on this system was recorded as 8 h and 42 min.

We performed an evaluation of the testing accuracy between VMF-ZTD and IGS-ZTD at GGOS stations. The detailed spatial information of the VMF grid and IGS site can be found in Table I.

TABLE I
IGS STATIONS AND NEAREST VMF GRID INFORMATION

| IGS site | Lat (°) | Lon (°) | H (m) | VMF Grid (5°×5°) | Country, Region |
|---|---|---|---|---|---|
| BJFS | 39.6N | 115.9E | 87.5 | (37.5°N, 117.5°E) | China, Fangshan |
| ALIC | 23.67S | 133.886E | 603.2 | (22.5°S, 132.5°E) | Australia, Alice Springs |
| PRDS | 50.871N | 114.296W | 1247.9 | (52.5°N, 112.5°W) | Canada, Calgary |
| SPTU | 21.929S | 50.492W | 508.78 | (22.5°S, 52.5°W) | Brazil, Tupa |

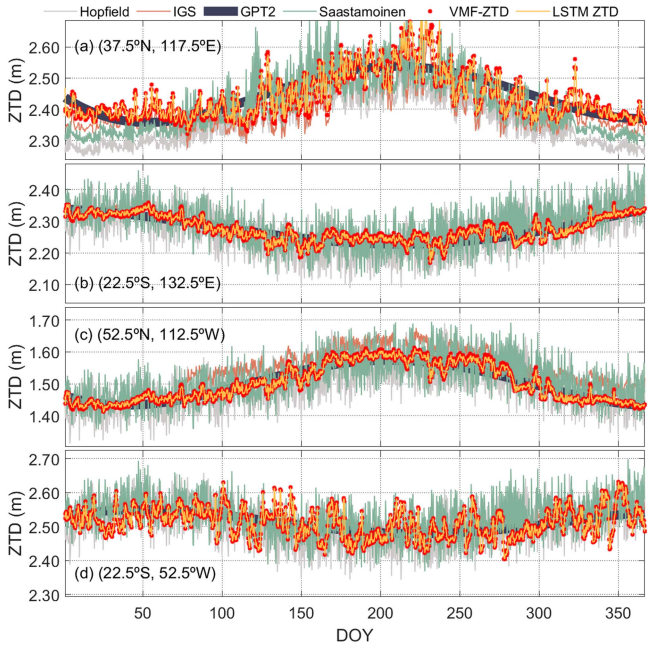


Fig. 5. ZTD time-series variations comparison of different models.

Furthermore, we conducted comparative experiments using the Hopfield and Saastamoinen methods. Meteorological parameters were obtained from the European Centre for Medium-Range Weather Forecasts Climate Change Service (https://cds.climate.copernicus.eu/).

The equations for accuracy indicators are listed as follows:

$$\text{Bias} = \frac{1}{N}\sum_{i=1}^{N}\left(\text{ZTD}_i^{\text{pre}} - \text{ZTD}_i^{o}\right) \tag{3}$$

$$\text{STD} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\text{ZTD}_i^{\text{pre}} - \text{ZTD}_i^{o} - \text{Bias}\right)^2} \tag{4}$$

$$\text{RMS} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\text{ZTD}_i^{\text{pre}} - \text{ZTD}_i^{o}\right)^2}. \tag{5}$$

Fig. 5 illustrates that the LSTM-ZTD model is consistent with the VMF-ZTD (original ZTD), further indicating that the LSTM-ZTD model significantly outperforms than other models. Table II demonstrates the detailed accuracy information among the different ZTD models at IGS stations.

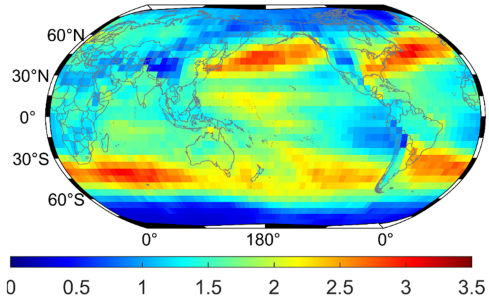


Fig. 6. RMS (cm) global distributions of the LSTM-ZTD model.

By associating Fig. 5 and Table II, it is illustrated that the LSTM-ZTD model is more precise than Hopfield and Saastamoinen models. The RMS of IGS-ZTD ranges from approximately 1 to 3 cm. The accuracy of the LSTM-ZTD model was significantly superior to the IGS-ZTD model at the BJFS, ALIC, PRDS, and SPTU stations, respectively.

## III. ACCURACY ANALYSIS AND VERIFICATION

We investigated the accuracy and reliability of the LSTM-ZTD model by analyzing the global distribution of accuracy and the variation in RMS with respect to seasons, latitude, and altitude. The average RMS indicator was employed to assess the accuracy of the ZTD predictions in comparison to the original values.

### A. Global Accuracy Analysis

The numerical accuracy results of the LSTM-ZTD model exhibited minimal bias, with RMS values of 0.256 cm (minimum), 3.076 cm (maximum), and 1.435 cm (average). The global distribution of RMS is depicted in Fig. 6, revealing superior accuracy over land compared to oceanic regions. Notably, the model demonstrated exceptional accuracy in the South and North Poles, as well as the Qinghai–Tibet Plateau. The heightened accuracy observed in Polar Regions can be attributed to the inactive troposphere, while the superior performance over the Qinghai–Tibet Plateau is primarily influenced by factors such as altitude and rainfall.

The statistical analysis reveals varying accuracies in different regions. Specifically, at the Antarctic, Arctic, and Qinghai–Tibet Plateau, the accuracies ranged from 0.25 to 0.32 cm, 0.4 to 1.1 cm, and 0.419 to 1.614 cm, respectively. However, at the North Atlantic, South Atlantic, and North Pacific, the accuracies were lower, ranging from 2 to 3 cm. Similarly, in the latitude zone of 37°—42° S, the accuracy was also comparatively poorer, with an RMS exceeding 2 cm. Notably, the western side of South America exhibited superior accuracy compared to the southeast side, with approximate accuracies of 0.9 cm and 2.3 cm, respectively.

The accuracy in the western region of the United States was superior to that in the eastern region, with approximate values of 1.5 cm and 2.3 cm, respectively. The North Pacific and North Atlantic Ocean exhibited the highest RMS values, exceeding 3 cm, while the lowest RMS value of 0.26 cm was observed in the Antarctic (82.5° S, 92.5° E). Overall, the numerical

TABLE II
ACCURACY COMPARISON AMONG DIFFERENT ZTD MODELS (CM)

| Site | Accuracy | Saastamoinen | Hopfield | GPT2 | IGS | LSTM-ZTD |
|---|---|---|---|---|---|---|
| BJFS | Bias | 2.95 | 3.05 | -0.21 | 2.05 | -0.01 |
| | STD | 3.42 | 3.19 | 2.19 | 2.16 | 1.92 |
| | RMS | 2.99 | 3.79 | 2.90 | 3.14 | 1.92 |
| ALIC | Bias | -4.25 | -0.06 | 0.09 | 2.75 | 0.00 |
| | STD | 2.03 | 2.69 | 3.32 | 2.83 | 1.43 |
| | RMS | 3.38 | 3.76 | 3.21 | 3.09 | 1.43 |
| PRDS | Bias | -1.12 | 1.16 | 0.11 | -2.82 | 0.00 |
| | STD | 3.6 | 3.59 | 2.10 | 2.73 | 0.82 |
| | RMS | 3.77 | 3.77 | 2.10 | 4.31 | 0.82 |
| SPTU | Bias | -2.58 | -2.67 | -0.28 | 3.17 | -0.01 |
| | STD | 3.26 | 3.90 | 1.6 | 2.14 | 0.90 |
| | RMS | 2.65 | 2.72 | 1.62 | 1.33 | 0.90 |

TABLE III
COMPARISON OF LSTM-ZTD MODEL ACCURACY ACROSS SEASONS

| Timespan | RMS (cm) | |
|---|---|---|
| | Average | [min, max] |
| Spring (N: Mar.–May; S: Sep.–Nov.) | 1.425 | [0.237, 3.266] |
| Summer (N: Jun.–Aug., S: Dec, Jan.–Feb.) | 1.467 | [0.277, 3.145] |
| Autumn (N: Sep.–Nov., S: Mar.–May) | 1.560 | [0.279, 3.704] |
| Winter (N: Dec., Jan.–Feb., S: Jun.–Aug.) | 1.364 | [0.147, 3.547] |

results demonstrate that the LSTM-ZTD model consistently outperformed in terms of prediction accuracy, thereby validating the rationale behind the modeling approach.

### B. Variation in Accuracy by Season

Numerous studies have been conducted by researchers to investigate the temporal and spatial distribution characteristics of GGOS atmosphere's zenith total delay (ZTD) data. These studies have revealed that ZTD exhibits annual and semiannual periodic changes and is correlated with latitude [11], [35]. In our investigation, we focused on forecasting ZTD values in different seasons, namely, spring (March–May in the Northern Hemisphere and September–November in the Southern Hemisphere), summer (June–August in the Northern Hemisphere and December–February in the Southern Hemisphere), autumn (September–November in the Southern Hemisphere and March–May in the Northern Hemisphere), and winter (December–February in the Northern Hemisphere and June–August in the Southern Hemisphere). Table III provides details on the global minimums, maximums, and average RMS values for each season. The average RMS values were found to be 1.425 cm, 1.467 cm, 1.560 cm, and 1.364 cm for spring, summer, autumn, and winter, respectively.

Table III illustrated that the accuracies in spring and winter were superior to summer and autumn. Fig. 7 shows the distribution of the RMS of the LSTM-ZTD model in spring, summer, autumn, and winter. In summary, the land area in winter had significantly outperformed RMS than the other three seasons, and the accuracies in autumn in the North Pacific and North Atlantic were worse than those in spring, summer, and winter.

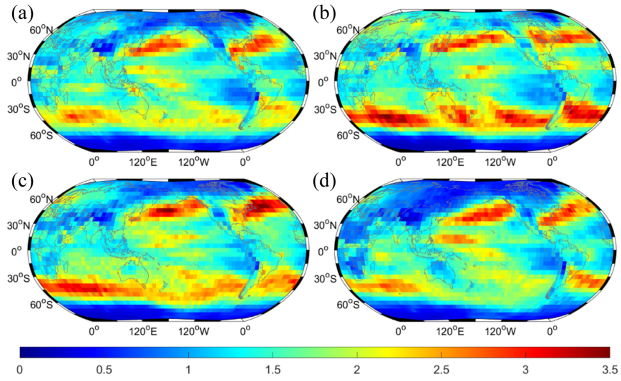


Fig. 7. Global distribution of LSTM-ZTD model accuracy (RMS in cm) in (a) spring, (b) summer, (c) autumn, and (d) winter.

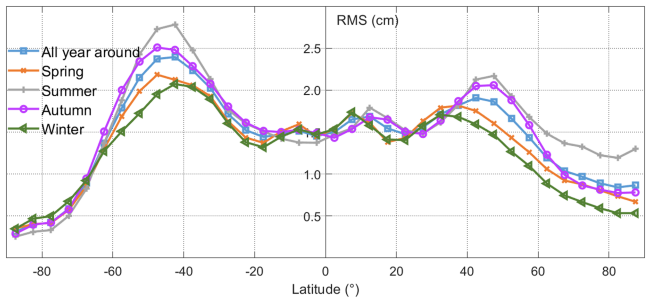


Fig. 8. ZTD time-series variations in different seasons along latitude.

The accuracy in the Arctic in summer was worse than that in spring, autumn, and winter. The accuracy in autumn in western South America is superior to spring, summer, and winter. At 37°–42° S, RMS distributions in spring and winter were smaller than in summer and autumn. In summer, the accuracy was the worst in this latitude zone, approximately 2.8–3.7 cm.

### C. Variation in Accuracy With Latitude

Fig. 8 illustrates the average RMS variation at the same latitude, indicating that the LSTM-ZTD model exhibited symmetrical distributions of RMS values across all seasons in both the Northern and Southern Hemispheres. In regions with

TABLE IV
VMF-ZTD AVERAGE RMS AND LSTM-ZTD AVERAGE RMSE WITH POLAR DAYS AND NIGHTS IN THE NORTH AND SOUTH POLAR REGIONS

| Seasons and Area | LSTM-ZTD average [min, max] (m) | LSTM-ZTD accuracy average [min, max] RMS (cm) |
|---|---|---|
| Summer and autumn, North Pole | 2.373 [2.295, 2.502] | 1.084 [1.051, 1.120] |
| Spring and winter, North Pole | 2.316 [2.235, 2.432] | 0.589 [0.546, 0.627] |
| Summer and autumn, South Pole | 1.595 [1.546, 1.634] | 0.356 [0.343, 0.361] |
| Spring and winter, South Pole | 1.588 [1.552, 1.632] | 0.298 [0.271, 0.328] |

low latitudes, the RMS value was approximately 1.5 cm. However, the accuracy gradually decreased beyond the latitude of 22° and then progressively increased toward latitudes above 42° N and below 42° S. Notably, the model demonstrated superior performance in regions characterized by a bipolar effect, with the Antarctic displaying better accuracy compared to the Arctic, with differences of within 0.5 cm and ranging from 0.5 to 1.3 cm, respectively.

In the midlatitudes, the accuracy was lowest at 42° S. The RMS values for spring, summer, autumn, and winter were measured at 2.2 cm, 2.7 cm, 2.5 cm, and 2.1 cm, respectively. In Antarctica, equatorial areas, and low-latitude regions, the RMSE exhibited slight fluctuations across the four seasons. Specifically, the accuracies in spring and winter surpassed those in summer and autumn within the latitudes of 22° to 62° S and 37° to 62° N. During summer, the RMS increased along the latitudes from 37° N to the Arctic, reaching a maximum value of 2.3 cm at 47° N. Conversely, in winter, the RMS values were relatively small, with a minimum of 0.53 cm observed at 87° N.

### *D. Accuracy Analysis at Polar Regions and Qinghai–Xizang Plateau*

The Qinghai–Tibet Plateau spans from 26°00'12" to 39°46'50" N in latitude and 73°18'52" to 104°46'59" E in longitude. Within this region, the test data encompassed 18 grid points, with elevations ranging approximately from 3000 to 5000 m, and an average altitude exceeding 4000 m across the plateau. The average annual ZTD was measured at 1.787 cm, with maximum and minimum values of 1.913 cm and 1.507 cm, respectively. Moreover, the average RMS was calculated at 0.824 cm, with minimum and maximum values of 0.419 cm and 1.614 cm, respectively. These findings suggest that the proposed model exhibits superior accuracy in the plateau area.

As the Earth rotates on its inclined axis, polar days and nights occur at the North and South poles. In Antarctica, polar days and nights take place in winter and summer, respectively. The seasonal variations in the Northern Hemisphere are opposite to those in the Southern Hemisphere. Table IV presents the average ZTD in the Arctic during summer and autumn, which was measured at 2.37 cm, with maximum and minimum values of 2.50 cm and 2.30 cm, respectively. Meanwhile, the average ZTD in the Arctic during spring and winter was recorded at 2.32 cm, with maximum and minimum values of 2.24 cm and 2.43 cm, respectively. Notably, the ZTD during a polar day was larger than that during a polar night. The ZTD in Antarctica was minimally affected by polar day and night, and the accuracy of the LSTM-ZTD model exhibited a similar trend.

In conclusion, the LSTM-ZTD model exhibited an average RMS of 1.435 cm in 2020. The average RMS values for spring, summer, autumn, and winter were recorded at 1.43 cm, 1.47 cm, 1.56 cm, and 1.36 cm, respectively. Notably, the LSTM-ZTD model demonstrated significantly superior prediction accuracies in the North and South Poles compared to low and middle latitudes, and even more so for the Tibetan Plateau. In addition, the distribution of RMS values was found to be symmetric between the Northern and Southern Hemispheres.

The relatively stable climatic conditions in the North and South poles contribute to slight periodic fluctuations in ZTD, thereby enhancing the prediction accuracy. However, the forecast accuracy for latitudes ranging from 37° to 42° S was poor, with a bias of approximately 4 cm. This indicates that the LSTM-ZTD model is unable to fully capture the complex changes occurring in this latitude range. The limitations may stem from the intricate topography and dynamic climatic variations in the area. Further research is necessary to explore the specific reasons behind these limitations.

## IV. CONCLUSION

This study presents the development of a global ZTD forecast model (LSTM-ZTD model) based on a modified LSTM neural network. The original dataset used in this study is the tropospheric product VMF3_OP time-series data provided by GGOS. We utilized the GPT2 empirical periodic model to extract the periodic signals in the ZTD sequence, which were subsequently used to filter the residual sequence. A modified LSTM neural network was then designed to train the residual sequence. Finally, the residual serial predictions and ZTD periodic signals were combined to output the ZTD forecast product.

The accuracy of the LSTM-ZTD model was carefully analyzed and verified, with results indicating an average RMS of 1.435 cm in 2020. The empirical periodic model is capable of extracting the general trend of true ZTD variations but a relatively large model deviation was observed. However, the proposed LSTM-ZTD model can almost completely compensate for this deviation. Importantly, the input parameters are only the residuals between the empirical periodic signals and the original ZTD, which have no relationship with the specific location (latitude and longitude).

This study trained the residual series of the empirical periodic model using the LSTM-ZTD model and examined global accuracy changes in the predicted products from both time and space

distributions. The model was also evaluated for seasonal changes in forecast accuracy. Results show that the ZTD accuracy of forecasting 12 months in advance performed well and significantly improved the forecast accuracy and timeliness of the ZTD. Overall, the LSTM-ZTD model presented in this study has the potential to enhance ZTD forecasting capabilities globally.

## V. Discussion

In this study, we have developed a ZTD forecasting model by integrating the decomposition of ZTD signals with LSTM neural networks. Through this approach, we have identified a distinct pattern within the residual terms obtained from decomposing the periodic signals of the atmospheric component, which eludes recognition by traditional physical models. Leveraging an enhanced LSTM neural network, we are able to capture and analyze this pattern effectively. Our proposed methodology exhibits significant improvements in global-scale accuracy compared to prior research, thus successfully overcoming the challenges associated with global-scale network integration.

Nevertheless, it should be noted that inadequate training dataset sizes impede the neural network's ability to learn these specific patterns, leading to compromised accuracy. In addition, there is room for enhancing the training time efficiency of our model. As part of our future work, we will prioritize optimizing the training time to further enhance the overall performance of the model.

## Acknowledgment

The authors would like to thank GGOS for providing VMF ZTD datasets; it can be accessed at https://vmf.geo.tuwien.ac.at/trop_products/GNSS/VMF3/VMF3_OP. They would also like to thank IGS for providing reference ZTDs at repository ftp://cddis.gsfc.nasa.gov/pub/gps/data.

## References

[1] Y. Altiner, L. Mervart, W. Soehne, and G. Weber, "Real-time PPP results from global orbit and clock corrections," in *Proc. EGU Gen. Assem. Conf. Abstr.*, 2010.

[2] J. Askne and H. Nordius, "Estimation of tropospheric delay for microwaves from surface weather data," *Radio Sci.*, vol. 22, no. 3, pp. 379–386, 1987.

[3] M. Bevis, S. Businger, T. A. Herring, C. Rocken, R. A. Anthes, and R. H. Ware, "GPS meteorology: Remote sensing of atmospheric water vapor using the Global Positioning System," *J. Geophysical Res., Atmospheres*, vol. 97, no. D14, pp. 15787–15801, 1992.

[4] H. D. Black, "An easily implemented algorithm for the tropospheric range correction," *J. Geophysical Res., Solid Earth*, vol. 83, no. B4, pp. 1825–1828, 1978.

[5] J. Boehm, B. Werl, and H. Schuh, "Troposphere mapping functions for GPS and very long baseline interferometry from European centre for medium-range weather forecasts operational analysis data," *J. Geophysical Res., Solid Earth*, vol. 111, no. B2, pp. 1–22, 2006.

[6] J. Böhm, R. Heinkelmann, and H. Schuh, "Short note: A global model of pressure and temperature for geodetic applications," *J. Geodesy*, vol. 81, no. 10, pp. 679–683, 2007.

[7] J. Böhm, G. Möller, M. Schindelegger, G. Pain, and R. Weber, "Development of an improved empirical model for slant delays in the troposphere (GPT2w)," *GPS Sol.*, vol. 19, no. 3, pp. 433–441, 2015.

[8] P. Callahan, "Prediction of tropospheric wet-component range error from surface measurements," JPL, Tech. Rep. 32-1526, 1973.

[9] J. P. Collins and R. B. Langley, "A tropospheric delay model for the user of the wide area augmentation system," Dept. Geodesy Geomatics Eng., Univ. New Brunswick, Fredericton, NB, Canada, 1997.

[10] M. Ding, W. Hu, X. Jin, and L. Yu, "A new ZTD model based on permanent ground-based GNSS-ZTD data," *Surv. Rev.*, vol. 48, no. 351, pp. 385–391, 2016.

[11] J. Dousa and P. Vaclavovic, "The evaluation of ground-based GNSS tropospheric products at geodetic observatory Pecný," in *Proc. IAG 150 Years*, 2015, pp. 759–765.

[12] Z. Du, Q. Zhao, W. Yao, and Y. Yao, "Improved GPT2w (IGPT2w) model for site specific zenith tropospheric delay estimation in China," *J. Atmos. Sol.-Terr. Phys.*, vol. 198, 2020, Art. no. 105202.

[13] J. Duan et al., "GPS meteorology: Direct estimation of the absolute value of precipitable water," *J. Appl. Meteorol. Climatol.*, vol. 35, no. 6, pp. 830–838, 1996.

[14] Y. Gao and K. Chen, "Performance analysis of precise point positioning using real-time orbit and clock products," *J. Glob. Positioning Syst.*, vol. 3, no. 1/2, pp. 95–100, 2004.

[15] T. Herring, "Modeling atmospheric delays in the analysis of space geodetic data," in *Proceedirws of Refraction of Transatmospheric Simals in Geodesy*, J. C. De Munck and T. A. Spoelstra, Eds., Delft, The Netherlands: Netherlands Geodetic Commission Publications on Geodesy, 1992.

[16] H. Hopfield, "Two-quartic tropospheric refractivity profile for correcting satellite data," *J. Geophysical Res.*, vol. 74, no. 18, pp. 4487–4499, 1969.

[17] I. M. Ifadis, "The atmospheric delay of radio waves: Modelling the elevation dependence on a global scale," Ph.D. dissertation, Aristotle Univ. Thessaloniki, Division Geotechnical Eng., Washington, DC, USA, 1986.

[18] S. Jin, O. Luo, and S. Gleason, "Characterization of diurnal cycles in ZTD from a decade of global GPS observations," *J. Geodesy*, vol. 83, no. 6, pp. 537–545, 2009.

[19] J. Kouba, "Implementation and testing of the gridded Vienna mapping function 1 (VMF1)," *J. Geodesy*, vol. 82, no. 4, pp. 193–205, 2008.

[20] E. Krueger, T. Schueler, and B. Arbesser-Rastburg, "The standard tropospheric correction model for the European satellite navigation system Galileo," in *Proc. Gen. Assem. URSI*, 2005, pp. 23–29.

[21] K. Lagler, M. Schindelegger, J. Böhm, H. Krásná, and T. Nilsson, "GPT2: Empirical slant delay model for radio space geodetic techniques," *Geophysical Res. Lett.*, vol. 40, no. 6, pp. 1069–1073, 2013.

[22] R. Leandro, M. Santos, and R. Langley, "UNB neutral atmosphere models: Development and performance," in *Proc. Nat. Tech. Meeting Inst. Navigation*, 2006, pp. 1–5.

[23] L. Li, Y. Xu, L. Yan, S. Wang, G. Liu, and F. Liu, "A regional NWP tropospheric delay inversion method based on a general regression neural network model," *Sensors*, vol. 20, no. 11, 2020, Art. no. 3167.

[24] W. Li et al., "New versions of the BDS/GNSS zenith tropospheric delay model IGGtrop," *J. Geodesy*, vol. 89, no. 1, pp. 73–80, 2015.

[25] W. Li, Y. Yuan, J. Ou, and Y. He, "IGGtrop_SH and IGGtrop_rH: Two improved empirical tropospheric delay models based on vertical reduction functions," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 9, pp. 5276–5288, Sep. 2018.

[26] W. Li, Y. Yuan, J. Ou, H. Li, and Z. Li, "A new global zenith tropospheric delay model IGGtrop for GNSS applications," *Chin. Sci. Bull.*, vol. 57, no. 17, pp. 2132–2139, 2012.

[27] L. Liu, S. Zou, Y. Yao, and Z. Wang, "Forecasting global ionospheric TEC using deep learning approach," *Space Weather*, vol. 18, no. 11, 2020, Art. no. e2020SW002501.

[28] J. Mohammed, "Artificial neural network for predicting global sub-daily tropospheric wet delay," *J. Atmos. Sol.-Terr. Phys.*, vol. 217, 2021, Art. no. 105612.

[29] A. Niell, "Preliminary evaluation of atmospheric mapping functions based on numerical weather models," *Phys. Chem. Earth, A, Solid Earth Geodesy*, vol. 26, no. 6-8, pp. 475–480, 2001.

[30] A. E. Niell, "Global mapping functions for the atmosphere delay at radio wavelengths," *J. Geophysical Res., Solid Earth*, vol. 101, no. B2, pp. 3227–3246, 1996.

[31] N. Penna, A. Dodson, and W. Chen, "Assessment of EGNOS tropospheric correction model," *J. Navigation*, vol. 54, no. 1, pp. 37–55, 2001.

[32] C. Pikridas, S. Katsougiannopoulos, and I. Ifadis, "Predicting zenith tropospheric delay using the artificial neural network technique. application to selected EPN stations," *J. Nat. Cancer Inst.*, vol. 88, no. 24, pp. 1803–1805, 2010.

[33] C. Rocken, T. Van Hove, and R. Ware, "Near real-time GPS sensing of atmospheric water vapor," *Geophysical Res. Lett.*, vol. 24, no. 24, pp. 3221–3224, 1997.

[34] C. Rocken et al., "Sensing atmospheric water vapor with the global positioning system," *Geophysical Res. Lett.*, vol. 20, no. 23, pp. 2631–2634, 1993.

[35] W. Rohm and J. Bosy, "The verification of GNSS tropospheric tomography model in a mountainous area," *Adv. Space Res.*, vol. 47, no. 10, pp. 1721–1730, 2011.
[36] A. Ruwali, A. S. Kumar, K. B. Prakash, G. Sivavaraprasad, and D. V. Ratnam, "Implementation of hybrid deep learning model (LSTM-CNN) for ionospheric TEC forecasting using GPS data," *IEEE Geosci. Remote Sens. Lett.*, vol. 18, no. 6, pp. 1004–1008, Jun. 2020.
[37] J. Saastamoinen, "Contributions to the theory of atmospheric refraction," *Bull. Géodésique*, vol. 105, no. 1, pp. 279–298, 1972.
[38] T. Schüler, "The TropGrid2 standard tropospheric correction model," *GPS Sol.*, vol. 18, no. 1, pp. 123–131, 2014.
[39] W. Suparta and K. M. Alhasa, "Modeling of zenith path delay over Antarctica using an adaptive neuro fuzzy inference system technique," *Expert Syst. With Appl.*, vol. 42, no. 3, pp. 1050–1064, 2015.
[40] G. Xiao, J. Ou, G. Liu, and H. Zhang, "Construction of a regional precise tropospheric delay model based on improved BP neural network," *Chin. J. Geophys.*, vol. 61, no. 8, pp. 3139–3148, 2018.
[41] T. Xu, S. Li, and N. Jiang, "Zenith troposphere delay prediction based on BP neural network and least squares support vector machine," in *Proc. EGU Gen. Assem. Conf. Abstr.*, 2020, pp. 1–6.
[42] Y. Yang, T. Xu, and L. Ren, "A new regional tropospheric delay correction model based on BP neural network," in *Proc. Forum Cooperative Positioning Service*, 2017, pp. 96–100.
[43] Y. Wang, Y. Qin, J. Kong, Y. Yao, and X. Gao, "Ionospheric refined mapping function construction based on LSTM," *IEEE Trans. Geosci. Remote Sens.*, vol. 62, 2024, Art. no. 4100115, doi: 10.1109/TGRS.2023.3338777.
[44] Y.-B. Yao, C.-Y. He, B. Zhang, and C.-Q. Xu, "A new global zenith tropospheric delay model GZTD," *Chin. J. Geophys.*, vol. 56, no. 7, pp. 2218–2227, 2013.
[45] Y. Yao, Y. Hu, C. Yu, B. Zhang, and J. Guo, "An improved global zenith tropospheric delay model GZTD2 considering diurnal variations," *Nonlinear Processes Geophys.*, vol. 23, no. 3, pp. 127–136, 2016.
[46] Y. Yao, C. Xu, J. Shi, N. Cao, B. Zhang, and J. Yang, "ITG: A new global GNSS tropospheric correction model," *Sci. Rep.*, vol. 5, no. 1, pp. 1–9, 2015.
[47] H. Zhang et al., "A tropospheric zenith delay forecasting model based on a long short-term memory neural network and its impact on precise point positioning," *Remote Sens.*, vol. 14, 2022, Art. no. 5921.
[48] Q. Zhang, F. Li, S. Zhang, and W. Li, "Modeling and forecasting the GPS zenith troposphere delay in West Antarctica based on different blind source separation methods and deep learning," *Sensors*, vol. 20, no. 8, 2020, Art. no. 2343.
[49] D. Zheng, W. Hu, J. Wang, and M. Zhu, "Research on regional zenith tropospheric delay based on neural network technology," *Surv. Rev.*, vol. 47, no. 343, pp. 286–295, 2015.

**Lin He** received the M.Sc. degree in geodesy and surveying engineering from the University of Chinese Academy of Sciences, Wuhan, China in 2014. He is currently working toward the Ph.D. degree with the School of Geodesy and Geomatics, Wuhan University, Wuhan, China.

His current research focuses on establishing a tropospheric delay correction model by machine learning method and its application in positioning and the method of surveying data processing.

**Yibin Yao** received the Ph.D. degree in geodesy and surveying engineering from Wuhan University, Wuhan, China, in 2004.

He is currently a Professor with the School of Geodesy and Geomatics, Wuhan University. He is also with the Collaborative Innovation Center for Geospatial Technology, Wuhan. His main research interests include global navigation satellite systems and ionospheric atmospheric meteorological studies, theory, and methods of surveying data processing.

**Chaoqian Xu** received the Ph.D. degree in geodesy and surveying engineering from the School of Geodesy and Geomatics, Wuhan University, Wuhan, China, in 2017.

He is currently a Postdoctoral Scholar with Wuhan University. His main research interests include real-time high-precision GNSS positioning and methods of surveying data processing.

**Huan Zhang** received the Ph.D. degree in geodesy and surveying engineering from the School of Geodesy and Geomatics, Wuhan University, Wuhan, China, in 2022.

He is currently a Lecturer with the School of Smarty City, Chonqing Jiaotong University, Chongqing, China. He is also a Postdoctoral Scholar with China Merchants Chongqing Communications Research and Design Institute, Co., Ltd., Chongqing. His research focuses on tropospheric delay correction model by machine learning method and its application in positioning.

**Feifei Tang** received the Ph.D. degree in geodesy and surveying engineering from Wuhan University, Wuhan, China, in 2009.

She is currently a Professor and Vice Dean with the School of Smart City, Chongqing Jiaotong University, Chongqing, China. Her research interests include feature extraction, change detection, and object recognition of laser scanning data and UAV data.

**Changquan Ji** received the master's degree in computer science from Northeast Normal University, Changchun, China, in 2011.

He is a Teacher with Chongqing Jiaotong University, Chongqing, China. His current research interests include network education and distant education.

**Zhuoya Liu** received the master's degree in water conservancy and hydropower engineering from Sichuan University, Chengdu, China, in 2014.

She is a Senior Engineer. She studies key technologies of monitoring, early warning, risk prevention, and control of power grid geological hazards within mountain areas. Her main research interest includes real-time high-precision GNSS positioning and methods of surveying data processing.

**Wentan Wu** received the master's degree in geodesy and surveying engineering from Wuhan University, Wuhan, China, in 2006.

He is a Professorate Senior Engineer with the Natural Resources Archives of Hebei Province, Hebei, China. His current research interests include high-precision GNSS navigation and positioning technology, GNSS time-series analysis, and GNSS engineering application.