

Article

A Machine Learning-Based Tropospheric Prediction Approach for High-Precision Real-Time GNSS Positioning

Jianping Chen *  and Yang Gao 

Department of Geomatics Engineering, University of Calgary, Calgary, AB T2N 1N4, Canada; ygao@ucalgary.ca

* Correspondence: jian.chen1@ucalgary.ca

Abstract: For high-precision positioning applications, various GNSS errors need to be mitigated, including the tropospheric error, which remains a significant error source as it can reach up to a few meters. Although some commercial GNSS correction data providers, such as the Quasi-Zenith Satellite System (QZSS) Centimeter Level Augmentation Service (CLAS), have developed real-time precise regional troposphere products, the service is available only in limited regional areas. The International GNSS Service (IGS) has provided precise troposphere correction data in TRO format post-mission, but its long latency of 1 to 2 weeks makes it unable to support real-time applications. In this work, a real-time troposphere prediction method based on the IGS post-processing products was developed using machine learning techniques to eliminate the long latency problem. The test results from tropospheric predictions over a year using the proposed method indicate that the new method can achieve a prediction accuracy (RMSE) of 2 cm, making it suitable for real-time applications.

Keywords: LSTM; FFNN; troposphere prediction; deep learning; machine learning; neural network



Citation: Chen, J.; Gao, Y. A Machine Learning-Based Tropospheric Prediction Approach for High-Precision Real-Time GNSS Positioning. *Sensors* **2024**, *24*, 2957. <https://doi.org/10.3390/s24102957>

Academic Editor: Thomas Pany

Received: 31 March 2024

Revised: 24 April 2024

Accepted: 2 May 2024

Published: 7 May 2024



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

1. Introduction

For high-precision positioning applications, various GNSS errors need to be mitigated, including the tropospheric error, which can reach up to a few meters. The troposphere is a layer of the Earth's atmosphere close to the Earth's surface. Therefore, the tropospheric delay is a local effect. This layer contains atomic and molecular constituents capable of delaying and bending radio signals. As the bending effect is negligible compared to the delay effect, only the delay effect is considered in this research [1]. The tropospheric delays can be divided, based on physical parameters, into two components: the zenith hydrostatic delays (ZHDs) and the zenith wet delays (ZWDs). The zenith total delays (ZTDs) can be obtained by combining the two components [2]. With accurately known local surface atmospheric pressure, the ZHD can be calculated precisely and subsequently applied as a correction to remove its effect on positioning. Yao et al. proposed a local fusion method from the PPP model as well as a few empirical models with a real-time local troposphere fitting model to improve the accuracy and reliability of local Hong Kong tropospheric prediction [3]. Compared to the ZHD, the ZWD is more difficult to model and determine with high accuracy due to local water vapor variations [4]. In order to separate ZWD from ZTDs, a so-called UNB3m model was proposed by Leandro et al. to model the ZHD as well as ZWD based on a few predefined lookup tables. The model evolved from the original UNB3 model by accounting for possible negative humidity [5]. Better accuracy than the original Saastamoinen model was shown in Leandro et al. [5,6]. In this study, we employ the UNB3m model to characterize the ZHD and subtract it from the ZTD to estimate the wet delays. The International GNSS Service (IGS) has developed zenith troposphere products in the SINEX_TRO exchange format (TRO), with data storage dating back to 1997 [7]. The Center for Orbit Determination in Europe (CODE) has provided products since 2017, with an original sampling interval of 2 h, which was changed later to 1 h starting from 8 January 2023. The latency of the IGS products, however, is 1 to 2 weeks, which makes them unable to support real-time applications.

In the realm of real-time troposphere applications, several regional correction services utilize troposphere components generated using data from densely distributed regional Continuously Operating Reference Stations (CORSs). Examples of such regional services include the QZSS CLAS. However, these services cover only limited local areas due to the availability constraints of dense networks. Other studies have attempted to utilize IGS ultra-rapid products with the Precise Point Positioning (PPP) technique to derive ZTDs for near real-time applications [8,9]. Additionally, efforts have been made using CORSs in China to model tropospheric wet delays and utilize the resulting products to enhance PPP convergence [10]. While real-time estimation of the troposphere has achieved adequate accuracy in water vapor estimation, this approach relies heavily on a solid connection between processing centers and remote stations to provide real-time GNSS observation data, making its reliability contingent on infrastructure dependability. However, observation outages are not uncommon among IGS stations, and the possible long re-convergence caused by outages when employing the PPP technique presents further obstacles to applications [11].

Neural network models are well suited for highly complex non-linear problems and have been utilized across various applications for decades [12]. The widely used types of neural networks include the feedforward neural network (FFNN) and recurrent neural network (RNN). The FFNN possesses a simple structure, with information flowing unidirectionally only (forward) within the network, while the RNN has a bi-directional information flow [12–14]. In an effort to overcome the gradient vanishing problem in the traditional RNNs, the long-short term memory (LSTM) method was introduced [15]. GNSS atmospheric correction prediction is a suitable application for adopting the machine learning approach. By applying the machine learning method to conduct real-time troposphere prediction, the latency associated with post-processed troposphere products can be eliminated. Numerous studies have already been carried out to apply neural network approaches to prediction of the effect of the ionosphere, which is another important component in the atmospheric delay affecting GNSS positioning accuracy [14,16–23]. There are research results that consistently demonstrate high prediction accuracy on regional or global VTEC. In terms of applying machine learning to troposphere prediction, some researchers have been using machine learning methods in ZWD prediction. Among those studies, Lu et al. developed a tropospheric delay network (TropNet) to estimate the troposphere over the continental USA [2]. The wet delays used in the research are provided by the Geostationary Operational Environmental Satellite-R series and the global forecast system (GFS). The GFS products are updated every 6 h and continuously predicted forward. An average improvement of 11.9% was demonstrated. However, the prediction process requires the most recent 6 h of data for the next 6 h of prediction, making it complicated for field applications. Li et al. employed LSTM on GNSS-derived ZTDs to enhance the Global Pressure-Temperature (GPT3) model for Antarctica [24]. Zhang et al. investigated wide-area precise tropospheric corrections (WAPTCs) together with those derived from GNSS as well as numerical weather prediction (NWP) across mainland China, and validation shows sufficient accuracy [25]. Li et al. applied a back-propagation neural network (BPNN) with meteorological parameters in mainland China. The training data were based on radiosonde data at 182 sites over 7 years [26]. Bi et al. compared BPNNs and proposed a 1D convolutional neural network (1D-CNN) and concluded that the CNN slightly exceeds that of the BPNN with an RMSE of 2.69 cm in China [27]. Meanwhile, Shi et al. introduced an efficient deep learning approach for estimating troposphere delays for CORSs across China [28]. However, both studies encountered limitations due to availability issues of historical data from the CORSs.

Due to the limitations in the existing approaches, a machine-learning-based approach is proposed in this paper which simplifies the implementation of a machine learning algorithm in tropospheric prediction and can lead to existing post-processed IGS troposphere products to support real-time application. The proposed approach offers a practical and accessible solution to troposphere prediction without the need for continuous long-term data and hard-to-access parameters. Instead, it requires readily available variables including

time, local temperature, and relative humidity for the prediction, all of which are widely available from various sources. By utilizing such commonly available information, this approach simplifies the prediction process and enhances its feasibility for applications. This paper has examined two tropospheric delay prediction approaches, one based on the LSTM deep learning method and the other on the FFNN, aiming to enhance the accuracy of long-term local tropospheric delay prediction. By utilizing six years of data spanning from 2017 to 2022 for model training and the entire year of 2023 data for testing, the objective is to leverage extensive historical data to train models and subsequently forecast future ZWD. The proposed model has been applied to predict local tropospheric wet delays over a location in Alberta, Canada. The training data were downloaded from the IGS website, and the weather data were retrieved from the Environment Canada website. Evaluation of the prediction performance includes a comparison with the troposphere products obtained from the Center for Orbit Determination in Europe (CODE) over the last 7 years. This comparison also involves assessing the accuracy of the predictions and the estimation from the PPP method.

The paper is structured as follows: Firstly, the methodology of the FFNN and LSTM machine learning methods, along with associated evaluation techniques, is discussed. Subsequently, test cases and results are presented, followed by an analysis of the performance of troposphere prediction using the proposed approach. Finally, conclusions are drawn, and future work is outlined.

2. Methodology

The slant troposphere delays can be retrieved as a byproduct of GNSS positioning. By reversing the mapping function and employing weighted averaging techniques, the station vertical troposphere delays can be generated. GPS-derived precipitable water vapor (PWV) was utilized for predicting precipitation events [29]. PPP technique was employed to derive PWV from zenith wet delay (ZWD). Another case study conducted in Turkey used the Artificial Neural Network (ANN) method to predict troposphere wet delay on selected dates representing summer, winter, and spring [30].

In December 2020, IGS finalized version 2 of the TRO file format in an effort to standardize the exchange format for tropospheric products, which can be traced back to 1997 [7]. The troposphere delay values in the TRO file are usually tropospheric zenith total delays (ZTDs) as well as total delay gradients. The zenith troposphere delays the relation between the ZTD and the hydrostatic component (ZHD) as well as the wet part (ZWD):

$$ZTD = ZWD + ZHD \quad (1)$$

The effort to model the ZHD has involved the realization of various models, such as the Saastamoinen model and the UNB3m model [5,6,31].

$$ZHD = \frac{[(0.0022779 \pm 0.0024)]P_0}{f_x(\phi, H)} \quad (2)$$

$$f_s(\phi, H) = 1 - 0.00266\cos(2\phi) - 0.00028H \quad (3)$$

Equations (2) and (3) describe the Saastamoinen model. P_0 is the surface pressure in mbar, and $f_s(\phi, H)$ is the function of geodetic latitude and height. In practice, local pressure at the receiver level is usually unavailable, and a default sea level pressure is applied instead. Therefore, the Saastamoinen ZHD remains constant at a particular location. To account for seasonal variations in zenith troposphere as well as the wet component, the UNB3m model was proposed, incorporating look-up tables for parameter variations across different latitudes. This model was developed by the University of New Brunswick, Canada [5,31].

In this study, we started with a focus on a single IGS station located near Priddis, Alberta, Canada, with data ranging from 29 January 2017, to 16 December 2023. After analyzing the single station, we subsequently extended the same approach to the additional 11

stations across Canada. Figure 1 depicts the Saastamoinen model and UNB3m model lines representing only the hydrostatic troposphere component along with the post-processing products over the past 7 years. The Saastamoinen model remains constant throughout the entire period, while the UNB3m model demonstrates a similar pattern to the post-processing products but with a lower magnitude. Notably, the dry part exhibits a larger magnitude during summer months and lower during winters, which is consistent with the observation from the total troposphere from IGS. Hence, the UNB3m hydrostatic part is employed to calculate the wet component of the troposphere.

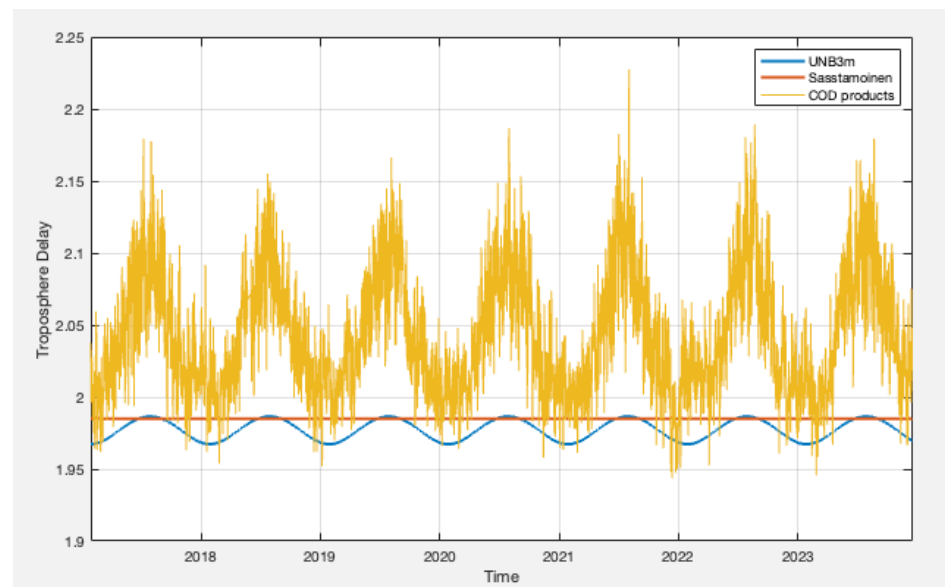


Figure 1. Troposphere comparison between hydrostatic UNB3m, Saastamoinen, and IGS-COD troposphere products for total zenith troposphere.

The converged weights and biases represent the best-fit model for the training inputs and outputs. Subsequently, this model can be utilized to predict unknown outputs based on a new set of inputs. In this study, both a fully connected FFNN and LSTM were employed to predict ZWD. The input list included seasonal and diurnal parameters, which were represented by the quadrature components of the day of the year (DoY) and the hour of the day (HoD). It is commonly known that LSTM is suitable for time sequence prediction, which requires continuous sequences in the training and following validation and prediction periods. However, in our specific application, such continuous data availability poses a challenge, particularly with meteorological data sources of Environment Canada. Therefore, the time inputs are given to both the FFNN and LSTM models. The time components can be described as follows [32]:

$$\text{HoDC} = \cos\left(\frac{2\pi \times t}{24}\right), \text{HoDS} = \sin\left(\frac{2\pi \times t}{24}\right) \quad (4)$$

$$\text{DoYC} = \cos\left(\frac{2\pi \times d}{365.25}\right), \text{DoYS} = \sin\left(\frac{2\pi \times d}{365.25}\right) \quad (5)$$

where t represents the hour of the day, and d denotes the day of the year.

The neural networks also make use of temperature and relative humidity as additional inputs. Consequently, the input data dimension is 6 by epochs, while the output data dimension is 1 by epochs.

The recurrent neural network (RNN) has the advantage of retaining temporal information through its internal loops, which pass the temporal information between different time frames. A chain-like neural network can be depicted by “unrolling” an RNN cell for easy understanding. The LSTM network is a special kind of RNN with characteristics of

learning long-term dependency [33]. Figure 2 illustrates the structure of a single LSTM cell. A LSTM cell primarily consists of three gates: the input gate i_t , the forget gate f_t , and the output gate o_t . With the help of the forget gate, an LSTM cell has the ability to retain or forget information to some extent.

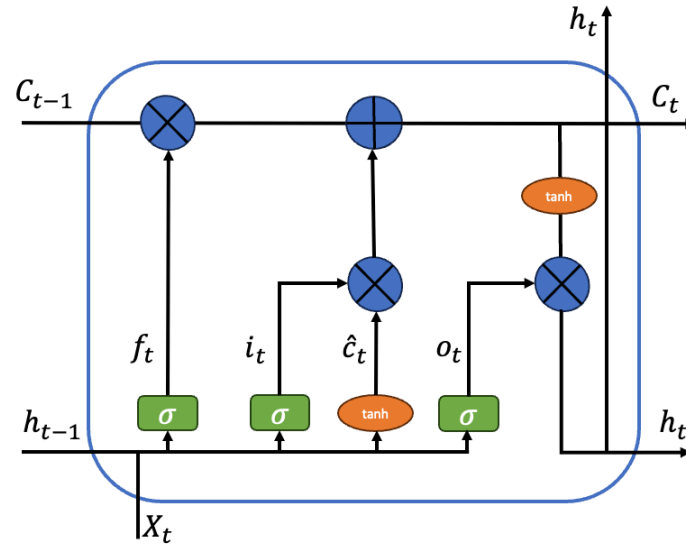


Figure 2. LSTM cell architecture.

Based on the data flow in the architecture, the mathematical model of LSTM can be summarized as [33]:

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad (6)$$

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \quad (7)$$

$$\hat{c}_t = \tanh(W_c[h_{t-1}, x_t] + b_c) \quad (8)$$

$$c_t = f_t \circ c_{t-1} + i_t \circ \hat{c}_t \quad (9)$$

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \quad (10)$$

$$h_t = o_t \circ \tanh(c_t) \quad (11)$$

Determining the optimal number of hidden neurons in the FFNN hidden layer and the number of LSTM cells in LSTM layers is a trial-and-error process. Too few units may lead to underfitting, while too many units, on the contrary, can cause overfitting [14]. In order to find optimal numbers of neurons or cells for this troposphere research, various sets of numbers were experimented with to determine the most balanced configuration. The whole dataset covers the years from 2017 to 2023, with the year 2023 reserved for testing purposes. Therefore, data from 2017 to 2022 were utilized to achieve this setting verification. Just like regular training and validation, the 80% rule was applied, randomly dividing the dataset into two groups: one training dataset and one validation dataset. A total of 8 settings of 2, 4, 6, 8, 10, 20, 40, and 100 neurons or cells were tested. Figure 3 shows the statistical RMSE of all 8 settings and two options of having a single layer or double layers. Based on the plots, a 10-neuron single hidden layer FFNN and double LSTM layers with 10 cells each were selected for further analysis in this paper. It is important to note that due to the random shuffling of the training dataset, the numbers in the plots may vary slightly each time the test is re-run, but the overall trend remains consistent.

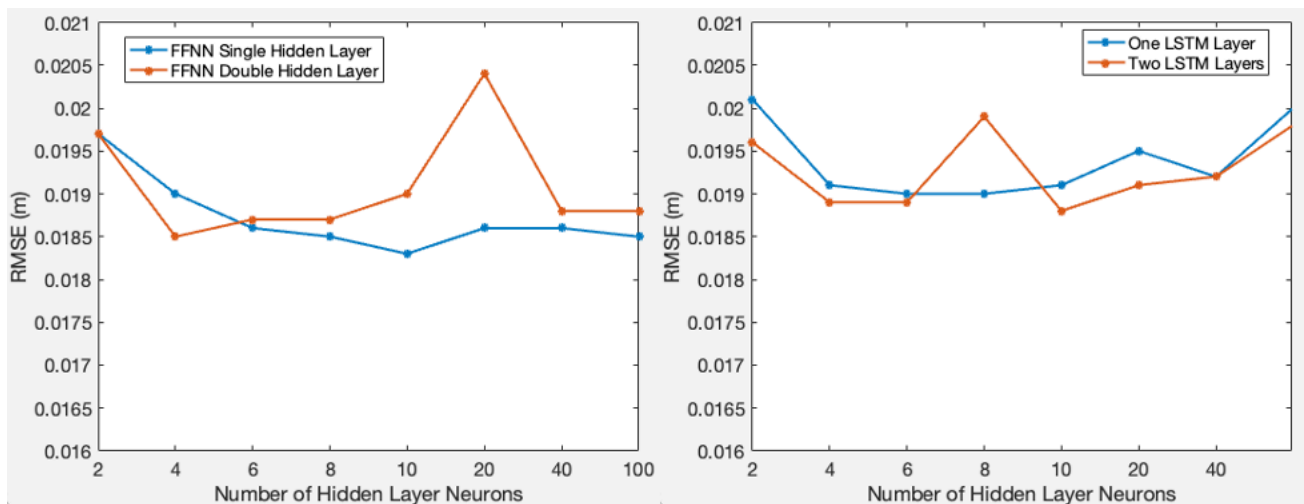


Figure 3. Number of nodes versus validation RMSE for FFNN and LSTM.

Figure 4 shows the architecture of the FFNN model being implemented in this study. As depicted, the model takes 6 inputs: Sin(DoY) (DoYS), Cos(DoY) (DoYC), Sin(HoD) (HoDS), Cos(HoD) (HoDC), temperature, and relative humidity. There is only one item as output: ZWD. One hidden layer resides in the middle with fully connected channels from inputs to output. During the training process, the model utilizes known inputs and outputs to adjust internal weights and biases. Finally, those trained weights and biases enable the model to make predictions. One of the main benefits of using neural networks is their ability to find the relationship between inputs and outputs without explicitly knowing the physical equations.

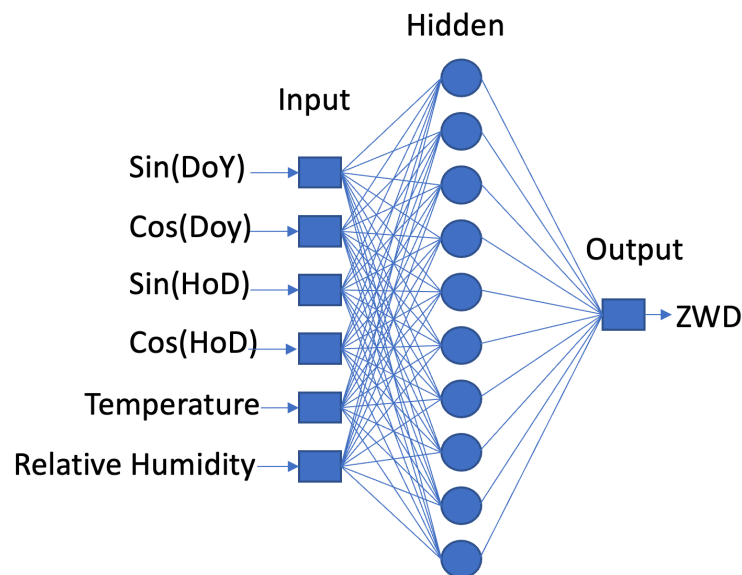


Figure 4. Architecture of FFNN.

Both FFNN and LSTM were chosen for troposphere prediction in this study. The FFNN model comprises one hidden layer with 10 neurons, while the LSTM model consists of two LSTM layers with 10 LSTM cells each. Figures 4 and 5 depict the proposed neural networks for FFNN and LSTM, respectively.

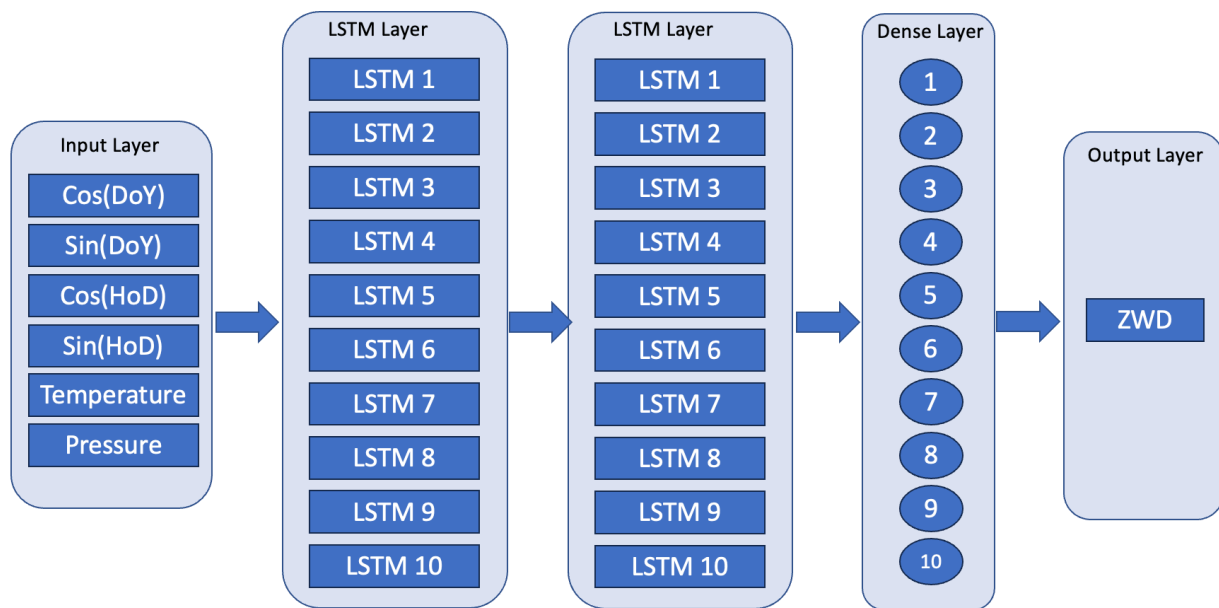


Figure 5. The architecture of LSTM.

Figure 4 displays the layers involved in the FFNN model utilized in this study. On the left side are all 6 inputs, while the single output is on the right. Temperature and relative humidity are included to account for local variations, as they affect the actual wet tropospheric delays. All information is transmitted to the output through a 10-neuron hidden layer.

The Levenberg–Marquardt backpropagation algorithm is employed for training parameters. Data division is based on random selection, and MSE is utilized for evaluation purposes.

The LSTM configuration shares the same set of input parameters and output parameters as the FFNN. The input layer transmits information to the first LSTM layer, followed by the second LSTM layer. After the two LSTM layers, a fully connected hidden layer is utilized to connect the dimensions of the output of the LSTM layer with the final output layer. All the LSTM training parameters are listed in Table 1.

Table 1. Parameters of the LSTM model.

Parameters	Value
Activation Function	Relu
Max Epochs	1000
Number of LSTM Layers	2 or 1
1st LSTM Layer Hidden Units	10
2nd LSTM Layer Hidden Units	10
Number of Dense Layers	1
Learning Rate	0.005
Optimizer	Adam
Dropout	0.2
Loss Function	Mean Squared Error (MSE)

3. Experiments and Results

3.1. Test Datasets

To test the troposphere prediction under various conditions, an IGS station (PRDS) located in Priddis, Alberta, Canada, approximately 25 km southwest of the University of Calgary campus, was selected. As CODE started providing troposphere data in 2017, the test data span from the year of 2017 to the end of 2023. The troposphere data from CODE are in TRO format, with a sampling interval of 2 h before 8 January 2023, which was changed to 1 h after that. For compatibility, the resolution of the 2023 data was diluted

to 2 h. The TRO files are accessible on the FTP CODE site (<ftp.aiub.unibe.ch>) (accessed on 16 December 2023). The first day with the PRDS station is 29 January 2017. There are missing data for 10 days in 2018, 3 days in 2019, and 2 days in 2022 for this particular station. Since the troposphere, especially the wet component of the troposphere, has a strong correlation of weather parameters, including temperature and relative humidity, these variables are included in the test datasets. The weather data for Priddis are from Environment Canada website (<https://climate.weather.gc.ca>) (accessed on 16 December 2023). There is also an outage in the weather data from 1 to 4 September 2020. The entire year of 2023 was used for testing the model, with data available until the data retrieval date of 16 December 2023. The downloaded weather data come with monthly files. The weather data contain temperature and relative humidity information but not pressure. Therefore, only temperature and relative humidity were included in the research. These parameters are also easy to obtain locally if needed for prediction purposes. The time tags in the weather data are in Local Standard Time (LST), which is 7 h behind UTC based on the Canadian Mountain time zone. When combining troposphere and weather data, the weather time tags need to be converted back to UTC. Additionally, the troposphere data in TRO files are total delays, and the wet part needs to be retrieved before feeding into neural networks based on the equation below.

$$ZWD = ZTD - ZHD \quad (12)$$

The ZHD is calculated using the UNB3m model. Figure 1 depicts the COD total delay as well as hydrostatic delays from UNB3m and Sasstamoinen. The figure shows that there are instances where negative wet troposphere delays occur. This could be from inaccuracies in modeled hydrostatic delays.

In terms of data splitting for training and validation, we allocated 80% of the data for training and the remaining 20% for validation. Additionally, the data from 2023 were kept aside for testing purposes. It is important to highlight that the training and validation datasets were randomly reshuffled before splitting to prevent any artificially imposed ordering. The time and data details are provided in Table 2.

Table 2. Time and data details.

	Training Data	Validation Data	Test Data
Time	1 January 2017 to 31 December 2022	1 January 2017 to 31 December 2022	1 January 2023 to 16 December 2023
Size	$25,765 \times 0.8 = 20,612$	$25,765 \times 0.2 = 5153$	4200

3.2. Accuracy of Tropospheric Prediction

In this study, we evaluated and compared two types of neural networks: the FFNN and LSTM. Due to outages in observing both weather data and troposphere products from CODE and unavailability of local surface pressure data, a dataset of approximately 6-year data was utilized for training and validation, while the data from the entire year of 2023 were for testing. Our analysis included both hourly errors and daily errors. Additionally, the maximum absolute error (MAXAE), the mean absolute error (MAE), and the root-mean-square error (RMSE) were analyzed on both an hourly and daily basis. The absolute error is given by:

$$|\epsilon| = \left| ZWD_{\text{pred}} - ZWD_{\text{tro}} \right| \quad (13)$$

where ϵ is the error between the estimate and known zenith wet tropospheric delay values, ZWD_{pred} is the estimated ZWD, and ZWD_{tro} is the reference ZWD from TRO files. The MAXAE, MAE, and RMSE values can be calculated using the following equations:

$$\text{MAXAE} = \max_{i=1:n}(\epsilon_i) \quad (14)$$

$$\text{MAE} = \frac{\sum_{i=1}^n |\epsilon_i|}{n} \quad (15)$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (\epsilon_i)^2}{n}} \quad (16)$$

where the subscript i means each sample point, and n is the total number of samples considered. Total statistics have 2-h intervals, and for daily, this number should be 12 items per day.

For the FFNN, as mentioned in the previous section, a hidden layer with 10 neurons was utilized for both training and prediction purposes. The FFNN training was configured with a predefined target iteration of 1000, and the training was terminated at the 115th iteration. The neural network settings are summarized in Table 3. In Figure 6, a comparison between the neural network prediction and actual troposphere from CODE, spanning the entire year of 2023. It is apparent that the predictions can closely follow the variations but with slightly less magnitude. Overall, the differences between the predicted and actual troposphere are generally small and within a few centimeters.

Table 3. Parameters of the FFNN Model.

Parameters	Value
Data Division	Random
Training Algorithm	Levenberg–Marquardt Backpropagation
Evaluation	Mean Squared Error (MSE)

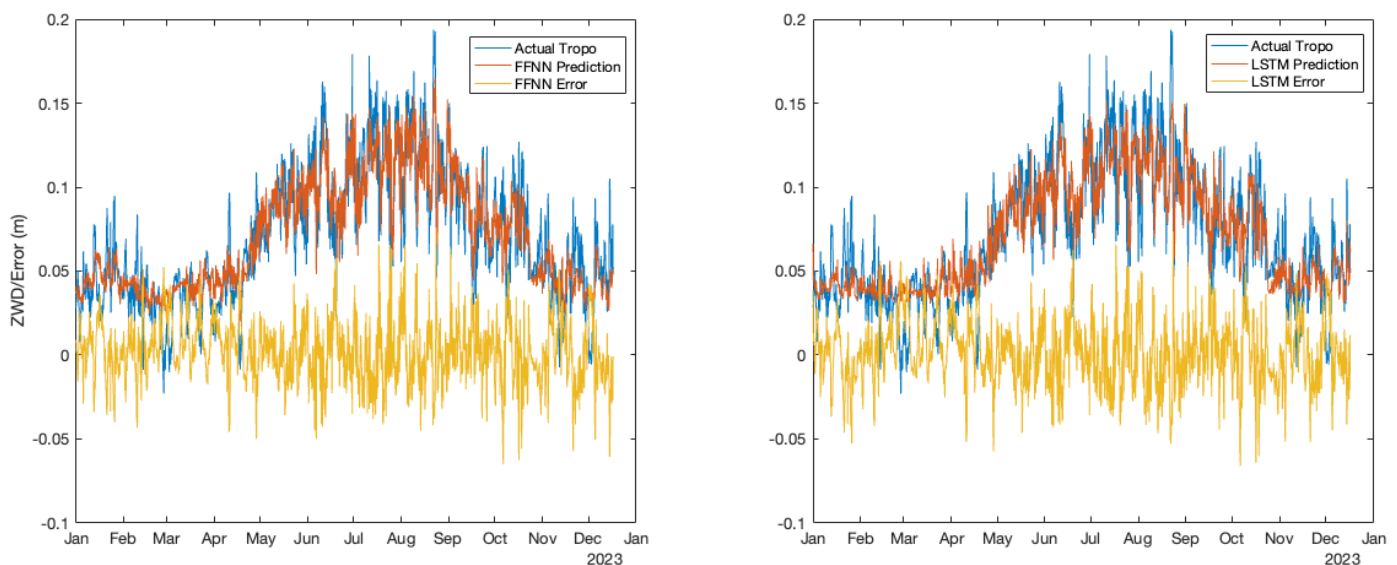


Figure 6. FFNN and LSTM prediction performance. The left side is for the FFNN, and the right side is for LSTM. Actual troposphere products and different neural network predictions as well as the prediction errors are shown in the plots.

In Figure 6, the actual difference between the troposphere from CODE and the machine learning prediction is also displayed. In the location of the station, which is very close to the campus of the University of Calgary, it is well known that the temperatures are relatively stable in summer and have more dramatic variations in winter. It is not unusual to have a 30 Celsius temperature change during a one-day timeframe. However, the figure demonstrates that the prediction is not seasonally dependent, as there is no obvious difference between summer and winter. This is relatively unexpected, considering that the temperature changes more significantly in winter than in summer in Southern Alberta. Figure 7 presents the statistical overview of the prediction throughout the entire year from

a different perspective. The figure contains RMSE, MAE, and MAXAE for each month of the year of 2023. It is evident that RMSE and MAE remain relatively stable throughout the year, with a magnitude mostly below 2 cm. However, for MAXAE, higher maximum absolute errors are observed during August.

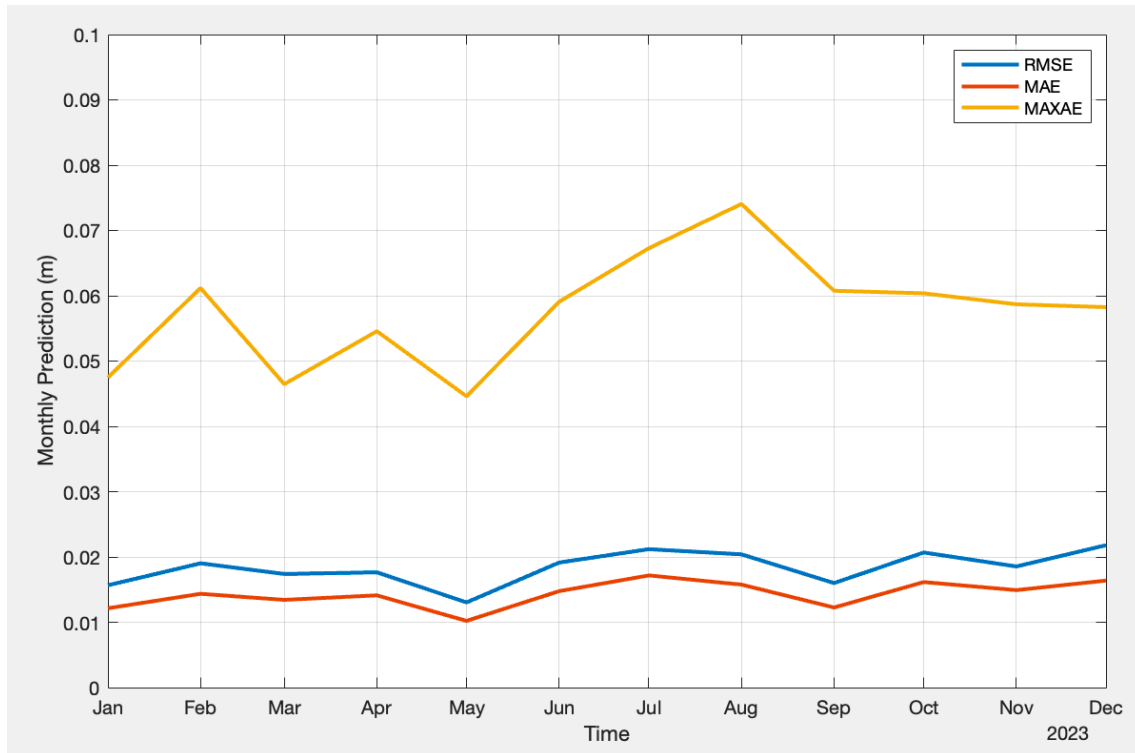


Figure 7. Monthly prediction Results (FFNN).

As for the LSTM model, there are no significant improvements compared to the FFNN model. Figure 6 depicts the comparison between the LSTM prediction and the actual troposphere from CODE.

The same figure displays the plot for the difference between the LSTM prediction and the actual troposphere delays from CODE. Again, it demonstrates a relatively small difference compared to the corresponding plot from the FFNN model. Figure 8 provides detailed statistical plots for RMSE, MAE, and MAXAE for individual months of the year 2023.

Table 4 presents the numerical results for the entire year of 2023 from both models. It is observed that the FFNN performs slightly better in terms of RMSE, while LSTM is better with absolute errors.

Table 4. Statistics of the two models.

Neural Network	RMSE (m)	MAXAE (m)	MAE (m)
FFNN	0.018	0.075	0.014
LSTM	0.019	0.068	0.015

Based on the single station performance, we extended the analysis to an extended selection across Canada. The stations cover the latitude from 45 degree to 64 degree north. And the longitude covers from coast to coast in Canada. The extra 11 stations are all available from IGS, which has recorded troposphere products since 2017.

The same method, which are 10 nodes for the FFNN and two LSTM layers with 10 neurons each, applies to all 11 extra stations, as shown in Figure 9. The results shown in Figure 10, reveal that the FFNN and LSTM are performing similarly within a few millimeters in RMSE. PRDS from Alberta and WILL from British Columbia demonstrate

the lowest RMSE, which indicates the best performance. FRDN from New Brunswick, HLFX of Nova Scotia, and SFJO of Newfoundland and Labrador exhibit poor performances in terms of prediction. Notably, these stations are situated along the Atlantic coast of Canada. One plausible explanation is the sparse distribution of CORSs in the region.

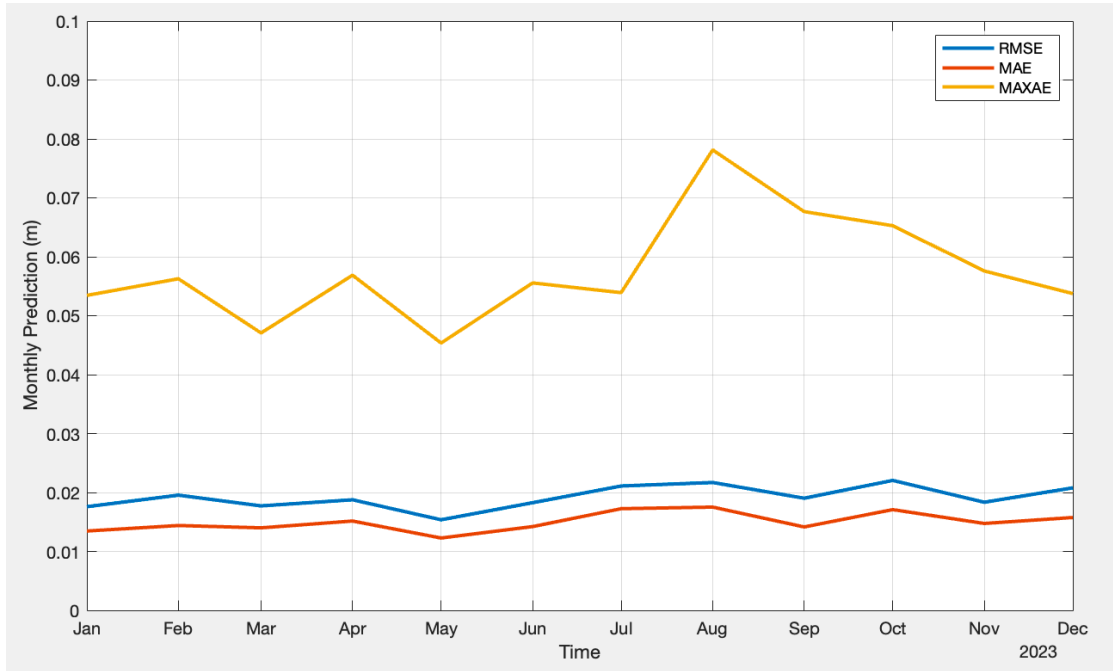


Figure 8. Monthly prediction results (LSTM).

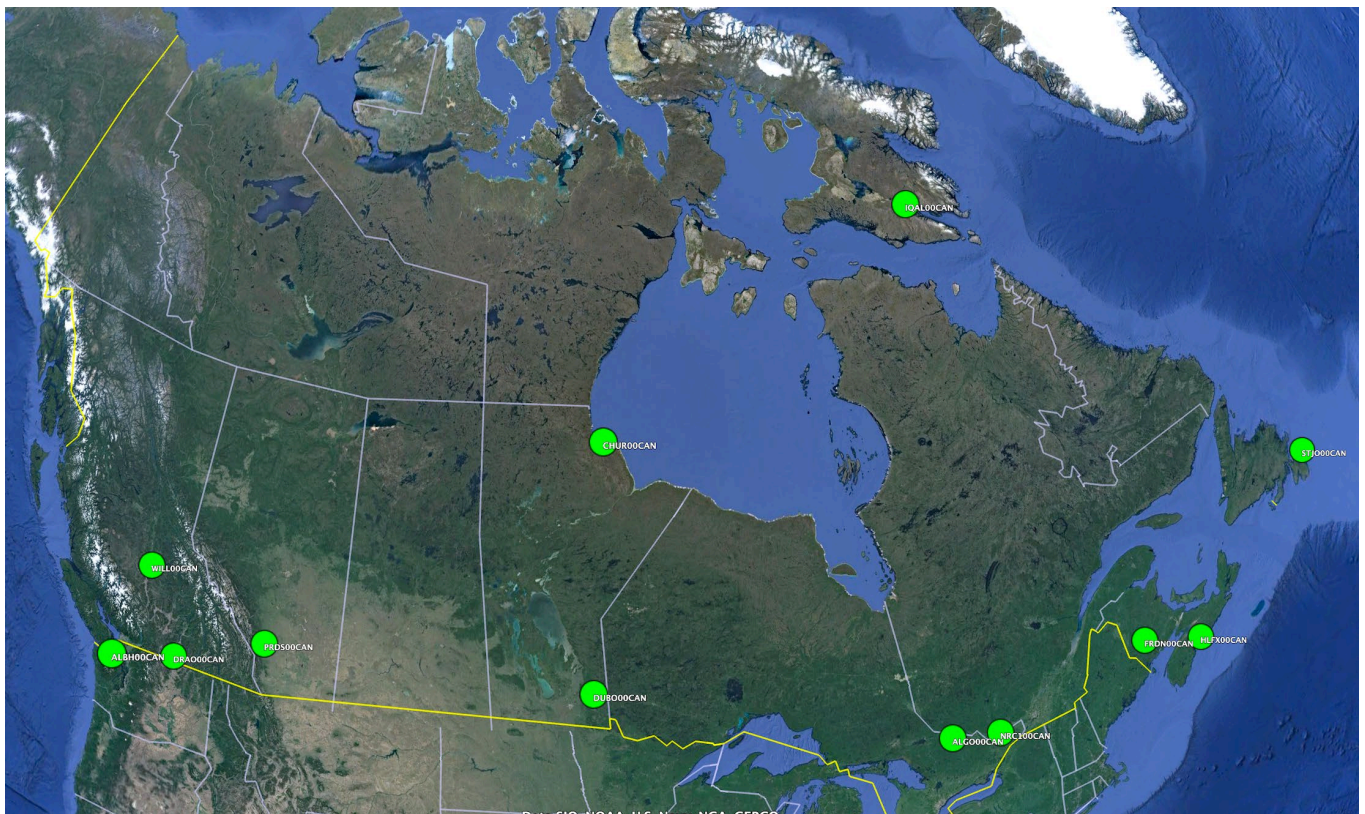


Figure 9. Extra stations included in the analysis.

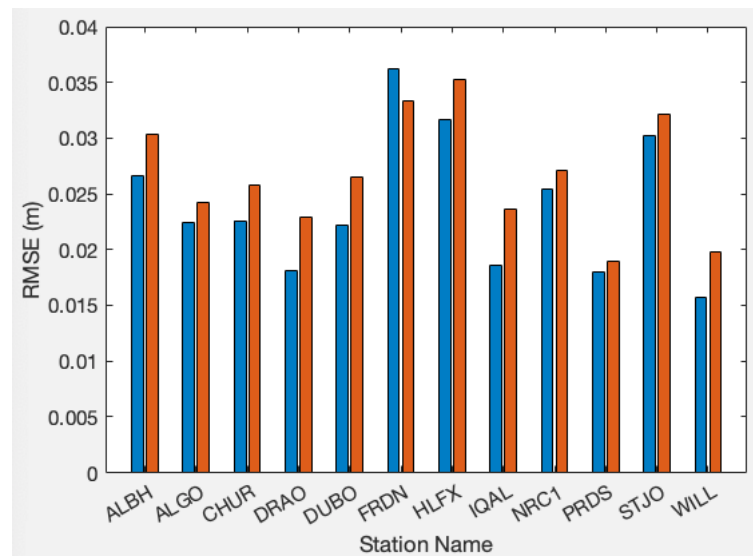


Figure 10. Neural network performance comparison between FFNN and LSTM. RMSEs are showing. Blue bars represent FFNN and red ones indicate LSTM.

4. Evaluation of Neural Network Models with Precise Point Positioning

In this section, the prediction performance of two neural network models will be assessed by comparing the results to the precise point positioning (PPP) estimation. The observation equations for pseudo-range and carrier phase are shown in the following equations.

$$P_k^j = \rho^j + dt_r - dT^j + m^j T + \left(\frac{f_1^2}{f_k^2} \right) I_1^j + b_{Pk} - B_{Pk}^j + \epsilon_P \quad (17)$$

$$\Phi_k^j = \rho^j + dt_r - dT^j + m^j T - \left(\frac{f_1^2}{f_k^2} \right) I_1^j + \lambda_k N_k + b_{\Phi k} - B_{\Phi k}^j + \epsilon_\Phi \quad (18)$$

where j represents the satellite index, k is frequency, P is the pseudo-range in meters, Φ is the carrier phase in meters, ρ is the geometric range, dt_r is the receiver clock in meters, dT^j is the satellite j clock, $m^j T$ is the slant troposphere delay, I_1^j is satellite j on frequency k slant ionosphere delay, b_{Pk} is the receiver clock bias for the pseudo-range on frequency k , B_{Pk}^j is the satellite j clock bias for the pseudo-range on frequency k , $b_{\Phi k}$ is the receiver clock bias for the carrier phase on frequency k , $B_{\Phi k}^j$ is the satellite j clock bias for the carrier phase on frequency k , and ϵ is the measurement error.

The receiver clock biases for the pseudo-range and carrier phase can be absorbed by the receiver clocks, satellite pseudo-range biases are applied, and satellite carrier-phase biases are ignored. Consequently, the orbit error, clock error, receiver clock error, troposphere, and ionosphere need to be estimated. Since the satellite carrier-phase biases are ignored, the ambiguity integers cannot be restored, and only float solutions are processed.

$$m^j T = m_h ZHD + m_w ZWD \quad (19)$$

where $m^j T$ is the total slant troposphere delay, m_h is the hydrostatic Niell mapping function, m_w is the wet Niell mapping function, ZHD is the hydrostatic zenith troposphere delay, and ZWD is the wet zenith troposphere delay. Rewriting the observation equations, we can obtain the following:

$$P_k^j = \rho^j + dt_{rk} - dT^j + m_h ZHD + m_w ZWD + \left(\frac{f_1^2}{f_k^2} \right) I_1^j + \epsilon_P \quad (20)$$

$$\Phi_k^j = \rho^j + dt_{rk} - dT^j + m_h ZHD + m_w ZWD - \left(\frac{f_1^2}{f_k^2} \right) I_1^j + \lambda_k N_k + \epsilon_\Phi \quad (21)$$

The slant troposphere delay can be separated into a hydrostatic part and a wet part based on Equation (13). As the hydrostatic part of the zenith troposphere is known from the UNB3m model, the estimated troposphere from the PPP engine is the zenith wet delay.

To carry out the validation test using PPP, data were logged using a Hemisphere GNSS Phantom 40 board at Hemisphere GNSS' Calgary office on 14 October 2023. The Hemisphere GNSS P40 board has the capability to track the Hemisphere Atlas LBand signal [34]. The Atlas correction was also logged with the receiver along with observations. Figures 11 and 12 illustrate the performance of the FFNN and LSTM models against the PPP estimation. The plots show that the PPP estimation exhibits more variations compared to the neural network ZWD.

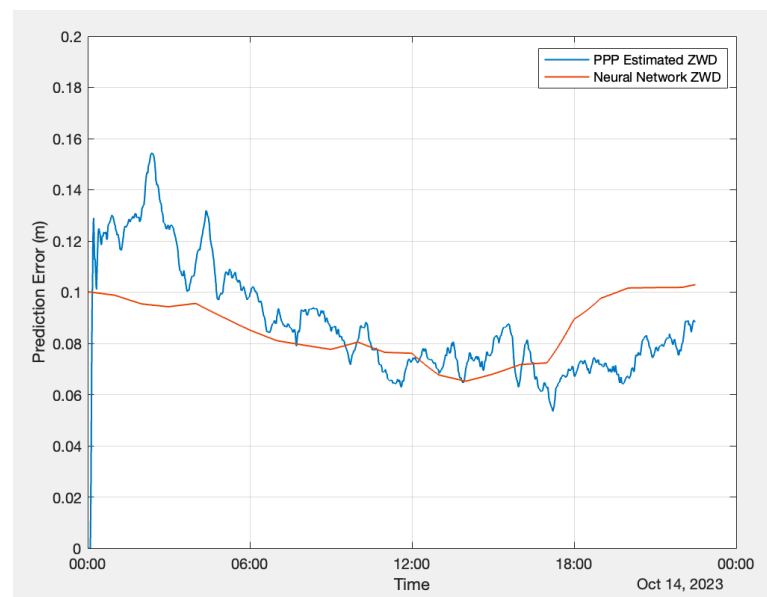


Figure 11. FFNN prediction versus PPP estimation.

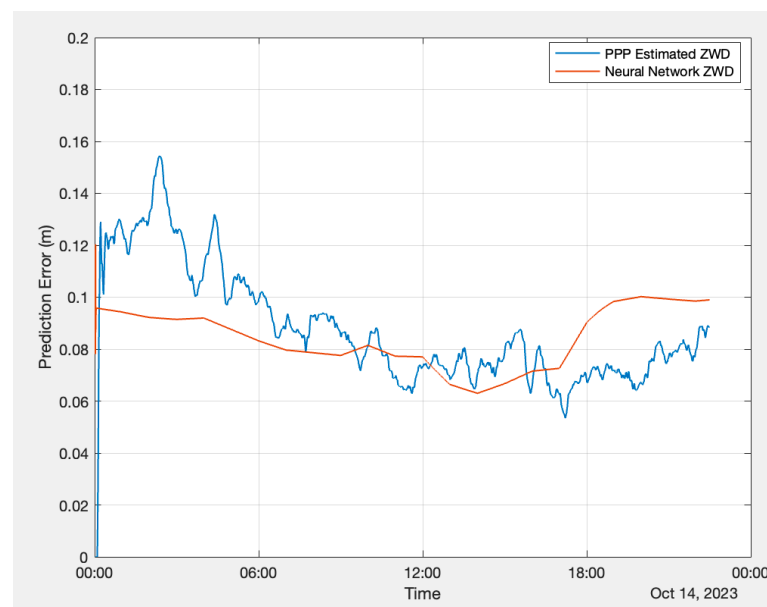


Figure 12. LSTM prediction versus PPP estimation.

As the PPP Atlas corrections do not include carrier-phase biases, the estimated carrier-phase ambiguities do not converge to integers. Because float PPP requires a longer convergence time, to ensure accurate statistics evaluation, the first 6 h of data were excluded from the analysis, and they are not accounted for in the statistical data in the results. Table 5 below presents the RMSE, MAXAE, and MAE of the FFNN and LSTM performance. Both the FFNN and LSTM models exhibit an RMSE of 1.6 cm and their MAEs are at 1.3 cm level, while the maximum absolute error from LSTM is slightly better than that from the FFNN.

Table 5. Prediction compared to PPP results.

Neural Network	RMSE (m)	MAXAE (m)	MAE (m)
FFNN	0.016	0.037	0.013
LSTM	0.016	0.036	0.013

5. Conclusions

In this work, we conducted a comparative analysis of a FFNN-based machine learning model and employed an LSTM-based deep learning neural network for long-term troposphere prediction over a duration of one year. Various machine learning settings were tested, and the results indicated that a single hidden layer FFNN model with 10 neurons, as well as a two-LSTM-layer model with 10 hidden cells each, demonstrated the best performance. Both the FFNN and LSTM models achieved RMSE of less than 2 cm for the entire year of 2023. Using real-receiver observation and Hemisphere GNSS Atlas correction, the generated ZWD on 14 October 2023, exhibited an RMSE of approximately 1.6 cm when compared to both FFNN and LSTM predictions. The research suggests that both machine learning methods can serve as simpler alternatives to IGS troposphere products. Unlike other studies that employ machine learning models with more complex inputs and outputs, our proposed models utilize local time and temperature and relative humidity, which are parameters readily available from various sources. Therefore, this method can lead to accuracy of postprocessing troposphere products with slight degradation in real-time applications. Future work will include training with additional aiding information, such as the local air pressure, to further enhance the prediction accuracy. Furthermore, it is essential to examine the predictive accuracy over a dense network and explore potential applications of PPP-RTK.

Author Contributions: Conceptualization, J.C. and Y.G.; methodology, J.C. and Y.G.; software, J.C.; validation, J.C. and Y.G.; formal analysis, J.C.; investigation, J.C.; writing—original draft preparation, J.C.; writing—review and editing, J.C. and Y.G.; supervision, Y.G.; project administration, Y.G.; funding acquisition, Y.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Natural Sciences and Engineering Research Council of Canada (NSERC) (RT691884).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: As described in the paper, IGS data are from the IGS FTP site, and meteorological data are from the Natural Resources Canada website.

Acknowledgments: We gratefully acknowledge the financial support from the Natural Sciences and Engineering Research Council of Canada (NSERC). We extend our appreciation to IGS and CODE for providing troposphere products and to Hemisphere GNSS for their assistance and collaboration.

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

References

- Shi, J. *Precise Point Positioning Integer Ambiguity Resolution with Decoupled Clocks*; University of Calgary: Calgary, AB, Canada, 2012.
- 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. Geodesy* **2023**, *97*, 34. [[CrossRef](#)]

3. Yao, Y.; Xu, X.; Xu, C.; Peng, W.; Wan, Y. Establishment of a Real-Time Local Tropospheric Fusion Model. *Remote Sens.* **2019**, *11*, 1321. [[CrossRef](#)]
4. Bevis, M.; Businger, S.; Herring, T.A.; Rocken, C.; Anthes, R.A.; Ware, R.H. GPS meteorology: Remote sensing of Atmospheric Water Vapor Using the Global Positioning System. *J. Geophys. Res. Atmos.* **1992**, *97*, 15787–15801. [[CrossRef](#)]
5. Leandro, R.F.; Santos, M.C.; Langley, R.B. UNB Neutral Atmosphere Models: Development and Performance. In Proceedings of the 2006 National Technical Meeting of the Institute of Navigation, Monterey, CA, USA, 18–20 January 2006.
6. Saastamoinen, J. Atmospheric Correction for the Troposphere and Stratosphere in Radio Ranging of Satellites. *Use Artif. Satell. Geod.* **1972**, *15*, 247–251.
7. Pacione, R.; Douša, J. *SINEX_TRO—Solution (Software/Technique) INdependent EXchange Format for TROospheric and Meteorological Parameters*; International GNSS Service: Pasadena, CA, USA, 2020.
8. Ding, W.; Teferle, F.N.; Kazmierski, K.; Laurichesse, D.; Yuan, Y. An Evaluation of Real-Time Troposphere Estimation Based on GNSS Precise Point Positioning. *J. Geophys. Res. Atmos.* **2017**, *122*, 2779–2790. [[CrossRef](#)]
9. Hadas, T.; Hobiger, T.; Hordyniec, P. Considering Different Recent Advancements in GNSS on Real-Time Zenith Troposphere Estimates. *GPS Solut.* **2020**, *24*, 99. [[CrossRef](#)]
10. Zheng, F.; Lou, Y.; Gu, S.; Gong, X.; Shi, C. Modeling Tropospheric Wet Delays with National GNSS Reference Network in China for BeiDou Precise Point Positioning. *J. Geodesy* **2018**, *92*, 545–560. [[CrossRef](#)]
11. Li, H.; Li, X.; Kang, Q. Handling Method for Outages of IGS Real-Time Service (RTS) in GNSS Real-Time Sensing of Atmospheric Water Vapor. *IEEE J. Sel. Top. Appl. Earth Obs. Remote. Sens.* **2023**, *16*, 8310–8318. [[CrossRef](#)]
12. Haykin, S. *Neural Networks and Learning Machines*, 3rd ed.; Pearson: London, UK, 2009.
13. Chen, J.; Gao, Y. Real-Time Ionosphere Map Prediction Based on IGS Rapid Products Using Long Short-Term Memory Deep Learning. In Proceedings of the 2022 International Technical Meeting of The Institute of Navigation, Long Beach, CA, USA, 25–27 January 2022. [[CrossRef](#)]
14. Chen, J.; Gao, Y. Real-Time Ionosphere Prediction Based on IGS Rapid Products Using Long Short-Term Memory Deep Learning. *NAVIGATION J. Inst. Navig.* **2023**, *70*. [[CrossRef](#)]
15. Hochreiter, S.; Schmidhuber, J. Long Short-Term Memory. *Neural Comput.* **1997**, *9*, 1735–1780. [[CrossRef](#)]
16. Sun, W.; Xu, L.; Huang, X.; Zhang, W.; Yuan, T.; Chen, Z.; Yan, Y. Forecasting of Ionospheric Vertical Total Electron Content (TEC) Using LSTM Networks. In Proceedings of the 2017 International Conference on Machine Learning and Cybernetics, Ningbo, China, 9–12 July 2017; pp. 340–344.
17. Kaselimi, M.; Doulamis, N.; Delikaraoglou, D. *A Sequence-To-Sequence Temporal Convolutional Neural Network for Ionosphere Prediction Using GNSS Observations*; ISPRS: Nice, France, 2020; pp. 813–820.
18. Liu, L.; Zou, S.; Yao, Y.; Wang, Z. Forecasting Global Ionospheric TEC Using Deep Learning Approach. *Space Weather* **2020**, *18*, e2020SW002501. [[CrossRef](#)]
19. Tang, R.; Zeng, F.; Chen, Z.; Wang, J.-S.; Huang, C.-M.; Wu, Z. The Comparison of Predicting Storm-Time Ionospheric TEC by Three Methods: ARIMA, LSTM, and Seq2Seq. *Atmosphere* **2020**, *11*, 316. [[CrossRef](#)]
20. Xiong, P.; Zhai, D.; Long, C.; Zhou, H.; Zhang, X.; Shen, X. Long Short-Term Memory Neural Network for Ionospheric Total Electron Content Forecasting Over China. *Space Weather* **2021**, *19*, e2020SW002706. [[CrossRef](#)]
21. Ulukavak, M. Deep Learning for Ionospheric TEC Forecasting at Mid-Latitude Stations in Turkey. *Acta Geophysica* **2021**, *69*, 589–606. [[CrossRef](#)]
22. Xia, G.; Zhang, F.; Wang, C.; Zhou, C. ED-ConvLSTM: A Novel Global Ionospheric Total Electron Content Medium-Term Forecast Model. *Space Weather* **2021**, *20*, e2021SW002959. [[CrossRef](#)]
23. Han, Y.; Wang, L.; Fu, W.; Zhou, H.; Li, T.; Chen, R. Machine Learning-Based Short-Term GPS TEC Forecasting during High Solar Activity and Magnetic Storm Periods. *IEEE J. Sel. Top. Appl. Earth Obs. Remote. Sens.* **2021**, *15*, 115–126. [[CrossRef](#)]
24. Li, S.; Xu, T.; Xu, Y.; Jiang, N.; Bastos, L. Forecasting GNSS Zenith Troposphere Delay by Improving GPT3 Model with Machine Learning in Antarctica. *Atmosphere* **2022**, *13*, 78. [[CrossRef](#)]
25. Zhang, H.; Yuan, Y.; Li, W. Real-Time Wide-Area Precise Tropospheric Corrections (WAPTCs) Jointly Using GNSS and NWP Forecasts for China. *J. Geodesy* **2022**, *96*, 44. [[CrossRef](#)]
26. 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](#)]
27. Bi, H.; Huang, L.; Zhang, H.; Xie, S.; Zhou, L.; Liu, L. A Deep Learning-Based Model for Tropospheric Wet Delay Prediction Based on Multi-Layer 1D Convolution Neural Network. *Adv. Space Res.* **2024**, *73*, 5031–5042. [[CrossRef](#)]
28. Shi, J.; Li, X.; Ouyang, C.; Xu, C. An Efficient Deep Learning-Based Troposphere ZTD Dataset Generation Method for Massive GNSS CORS Stations. *IEEE Trans. Geosci. Remote Sens.* **2023**, *61*, 5801311. [[CrossRef](#)]
29. Zhao, Q.; Yao, Y.; Yao, W. GPS-Based PWV for Precipitation Forecasting and Its Application to a Typhoon Event. *J. Atmospheric Solar-Terrestrial Phys.* **2018**, *167*, 124–133. [[CrossRef](#)]
30. Selbesoglu, M.O. Spatial Interpolation of GNSS Troposphere Wet Delay by a Newly Designed Artificial Neural Network Model. *Appl. Sci.* **2019**, *9*, 4688. [[CrossRef](#)]
31. Leandro, R.F.; Santos, M.C. A Neural Network Approach for Regional Vertical Total Electron Content Modelling. *Stud. Geophys. Et Geod.* **2007**, *51*, 279–292. [[CrossRef](#)]

32. Mckinnell, L.A.; Poole, W.V. Neural Network-Based Ionospheric Modelling Ove the South African Region. *S. Afr. J. Sci.* **2004**, *100*, 519–523.
33. Fang, W.; Jiang, J.; Lu, S.; Gong, Y.; Tao, Y.; Tang, Y.; Yan, P.; Luo, H.; Liu, J. A LSTM Algorithm Estimating Pseudo Measurements for Aiding INS during GNSS Signal Outages. *Remote Sens.* **2020**, *12*, 256. [[CrossRef](#)]
34. Chen, J.; Duong, V.; Kanji, A. Enhancing Global PPP Service Reliability with Hemisphere Atlas®and Galileo HAS: A Dual Redundant Approach. In Proceedings of the 36th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GNSS+ 2023), Denver, CO, USA, 11–15 September 2023. [[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.