Machine learning-based modelling of zenith wet delay using terrestrial meteorological data in the Brazilian territory

Afonso Marques Albuquerque[1], Raphael Silva Nespolo[1], Antonio Maria Garcia Tommaselli[1], Rorai Perreira Martins-Neto[2], Nilton Nobuhiro Imai[1], Daniele Barroca Marra Alves[1], Tayna Aparecida Ferreira Gouveia[1], Gabriel Oliveira Jerez[1]

[1] Faculty of Science and Technology, São Paulo State University, Presidente Prudente, Brazil
[2] Faculty of Forestry and Wood Sciences, Czech University of Life Sciences Prague, Prague, Czech Republic

Keywords: Zenith Wet Delay (ZWD), Random Forest, Machine learning, Meteorological stations.

Abstract

The Zenith Total Delay (ZTD) is one of the primary error sources derived from the neutral atmosphere associated with the GNSS (Global Navigation Satellite Systems) technique. Zenith Wet Delay (ZWD) is the smallest part of the ZTD, but the high variability is caused by spatial-temporal variation, making the modelling of this component a challenging task. Although ZWD is considered an error in GNSS positioning, it is also a variable composed mainly of water vapour and can, therefore, be used for atmospheric investigations, and assists in climate studies for precipitation events. In this work, a model was trained to estimate the delay wet component from surface atmospheric parameters. The training data comes from 29 radiosonde stations around Brazil, for a six-year period (2017 to 2022), with data collected at 12 h UTC (Universal Time Coordinated). The model was validated using the holdout technique, with 70% of the data used in training and 30% for validation (cross-validation analysis). The generated model achieved a RMSE (Root Mean Squared Error) of approximately 38 mm, with an 81% of determination coefficient.

1. Introduction

The neutral atmosphere is a gaseous layer extended from the Earth's surface to an altitude of approximately 50 km. This definition, assumed by the Geodesy, considers the presence or absence of ions (Seeber, 2003; Elgered and Wickert, 2017; Teunissen and Montenbruck, 2017).

The neutral atmosphere is a non-ionized layer, i.e., it does not have the presence of ions, but it contains dry gases (hydrostatic component), mainly hydrogen and oxygen, and water vapour (wet component) (Davis et al., 1985; Vianello and Alves, 2000; Sapucci, 2001; Elgered and Wickert, 2017). In this layer, according to the level of existing water vapour, combined with the presence of energy provided by solar radiation, a meteorological phenomena usually occur (Jacob, 1999; Vianello and Alves, 2000). In addition, electromagnetic signals travelling through the neutral atmosphere can be refracted (Davis et al., 1985; Elgered and Wickert, 2017).

The presence of the wet component in this layer does not occur uniformly. Its highest concentration is found mainly in the initial kilometres, extending from the surface to approximately 4 km in height, gradually decreasing up to 10 km (Vianello and Alves, 2000; Wallace and Hobbs, 2005). In higher altitudes, their presence is very low or non-existent. The geographical location is also relevant because in tropical regions, characterized by hot and humid climates, the concentration tends to be higher due to the evaporation of water. The opposite can be stated in desert and pole regions. Related to the hydrostatic component, it behaves in a more homogeneous way (Vianello and Alves, 2000; Wallace and Hobbs, 2005; Elgered and Wickert, 2017).

The effects caused by the neutral atmosphere can impact the propagation of the GNSS (Global Navigation Satellite Systems) signal, with errors ranging from a few meters to approximately 30 m, with the wet component accounting for 10 % of the total errors. In addition to the variation in different atmospheric conditions, this magnitude is also associated with the elevation angle of the observed satellites. Near the zenith, this delay can cause errors of more than 2.5 m. This effect, known as ZTD (Zenith Total Delay), is influenced by atmospheric elements such as pressure, temperature, and water vapour. It occurs due to atmospheric refraction, which causes a delay in the signal reception, when passing through the neutral atmosphere. The variation of the refractivity index along this path is due to the presence of hydrostatic and wet components. When this effect is primarily influenced by gases, it is called ZHD (Zenith Hydrostatic Delay), and when influenced by water vapour, it is called ZWD (Zenith Wet Delay) (Thayer, 1974; Mendes, 1998; Vianello and Alves, 2000; Sapucci, 2001; Seeber, 2003; Hofmann-Wellenhof et al., 2007; Monico, 2008; Nievinski, 2009; Gouveia, 2013; Gouveia et al., 2020).

The ZWD values exhibit high spatio-temporal variability, making it difficult to determine this component using surface measurements. The wet component can be obtained by radiometers or radiosonde data through the numerical integration of atmospheric profiles (Sapucci, 2001; Nievinski, 2009). Although instruments such as radiometers and radiosondes can provide accurate measurements of the wet component, their high costs can limit their use, preventing high sampling rates (Bevis et al., 1992; Sapucci et al., 2006).

Although the neutral atmosphere represents a source of error for GNSS positioning (ZTD, ZHD and ZWD), for other sciences, such as Meteorology, it can represent a source of information, since delay values are associated with atmospheric behaviour. In this context, ZWD is an important variable in climatic studies of precipitation, due to its relationship with precipitable water vapour (PWV) (Bevis et al., 1992; Sapucci, 2001, 2014, 2019).

Several studies have been conducted in order to model neutral atmospheric delay, focusing especially on ZWD. Some authors like De Oliveira et al. (2017) and Lu et al. (2017) have developed models for ZWD estimates in PPP (Precise Point Positioning) GNSS processing. However, the accuracy of existing models is still limited due to their unsuitability to account for spatial variations in atmospheric water vapour (Davis et al., 1985).

Recently, machine learning techniques have been employed to enhance ZWD modelling. Chen et al. (2022) and Xiong et al. (2021), for example, developed machine-learning techniques to model atmospheric parameters, while Gao et al. (2021), Li et al. (2024), and Li, Yuan, and Jiang (2023) proposed modelling ZWD through artificial intelligence.

However, even with the advances made with ZWD modelling, current models have a major flaw in their ability to predict non-stationary variations or short-term fluctuations in the neutral atmosphere. To solve this problem, it was proposed in the present study the development of a more accurate and flexible ZWD model, capable of predicting the temporal and spatial variations of atmospheric water vapour.

There are about 100 radiosonde stations in Brazil, with an average sample of two daily releases (University of Wyoming, 2024). There are also more than 500 INMET (*Instituto Nacional de Meteorologia*) automatic surface stations with hourly sampling (INMET, 2023). Obtaining a model capable of determining precise ZWD from surface measurements may represent a significant advancement not only for Geodesy and positioning studies but also for climatic and atmospheric research in Brazil. This advancement could enable the estimation of values in more isolated locations across the country within a much larger and cost-effective sample.

In this context, this study investigates the hypothesis that it is feasible to develop a model for obtaining ZWD from surface measurements. The model development leverages advanced machine learning techniques, such as Random Forest (RF), to forecast the wet component of delay based on atmospheric parameters at the station level.

The section 2 details the methodology employed in this study. Initially, the data used are presented in Section 2.1, which shows the sources of information, the temporal resolution of the data, and the study locations. The section 2.2 presents the methodology for obtaining the ZWD from radiosonde in-situ measurements. Subsequently, Section 2.3 addresses the method of applying random forest to the radiosonde data for estimating the wet component delay. Section 2.4 presents the methodology for validating the results using the holdout and cross-validation techniques. Section 2.5 shows a correlation analysis of the variables used to train the model. The section 3 presents the results obtained from the model trained using Random Forest. Section 3.1 discusses the results of the model validation phase, evaluating the variation of the estimates in comparison to the reference values. In Section 3.2, tests are conducted with surface stations (INMET) to estimate the ZWD using the trained model. These estimates are then evaluated using radiosonde stations as a reference, calculating the RMSE obtained at the test stations and representing the values on a map of the Brazilian territory.

2. Methodology

2.1 Dataset

The main dataset used in this work were the atmospheric profiles obtained by radiosonde stations. The equipment consists of a weather balloon with an attached sensor (radiosonde) that allows the measurement of atmospheric information in situ as it gains altitude. Based on the radiosonde profiles, data on atmospheric pressure, dew point temperature, relative humidity, geopotential height, wind speed, directions, and other parameters can be obtained in a series of layers that reach high levels of the neutral atmosphere (Sapucci, 2001).

The radiosonde data used in this work was provided by the University of Wyoming from 29 radiosonde stations in the Brazilian territory, over 6 years (2017-2022) (University of Wyoming, 2024). The collections were performed at 12h UTC (Universal Time Coordinated). The stations and the data period were selected based on availability, seeking the maximum number of stations with the least data absence and the largest continuity during the years.

In addition, surface weather data from INMET stations were used as input to the model. The data used was from 2023, collected at 12 UTC. Ten stations were used for the tests, and the stations were chosen based on their geographic location in the Brazilian regions: North, Northeast, Midwest, Southeast, and South.

Figure 1 shows the map of stations used in the study. The radiosonde stations were used for training the model (shown in red), while in blue are shown all the automatic surface stations available in Brazil. In green are the automatic stations tested in the training model obtained.

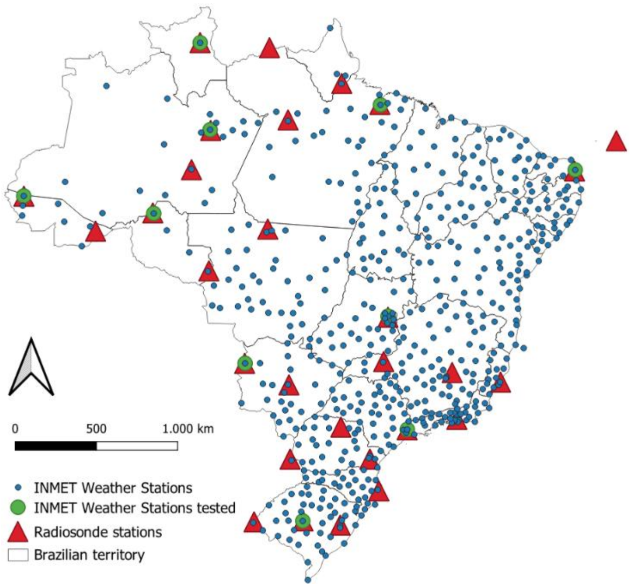

Figure 1. Map of INMET and radiosonde stations and surface meteorological stations in Brazil.

In order to mitigate large atmospheric variations, radiosonde and INMET station pairs were selected with distances smaller than 50 km (Monico, 2008). Due to the high density of INMET surface stations, it was feasible to select stations within distances smaller than 10 km (Table 1). The ZWD obtained from the model was compared with the on-site measurements obtained by the nearest radiosonde on the same date and time.

Due to the data sources employed, the final model must ensure a temporal resolution of 24 hours and a spatial resolution dependent on the number of stations.

https://doi.org/10.5194/isprs-annals-X-3-2024-13-2024

Local	Stations		Distance (km)
	Radiosonde	INMET	
Manaus (AM)	SBMN	A101	6.57
Natal (RN)	SBNT	A304	9.28
Brasília (DF)	SBBR	A001	7.86
São Paulo (SP)	SBMT	A701	2.81
Santa Maria (RS)	SBSM	A803	2.06
Belém (PA)	SBBE	A201	5.68
Corumbá (MS)	83554	A724	3.44
Porto Velho (RO)	SBPV	A925	7.98
Boa Vista (RR)	SBBV	A135	1.77
Cruzeiro do Sul (AC)	82705	A705	1.61

Table 1. Location of the radiosondes from INMET and surface stations, along with the distance between them.

2.2 ZWD-Radiosonde

Using the tool improved by Lima (2020) and developed by Sapucci (2001) in MATLAB software, it was possible to perform the processing of radiosonde data, obtaining the components of the zenith delay, including the ZWD. The tool obtains the values from the integration of atmospheric radiosonde profiles, from the surface (h_s) to the top of the neutral atmosphere (h_{top}), based on Equation 1 (Sapucci, 2001; Nievinski, 2009; Elgered; Wickert, 2017b; Gouveia et al., 2020).

$$ZWD = 10^{-6} \int_{h_s}^{h_{top}} \left(k_2' \frac{e}{T} Z_w^{-1} + k_3 \frac{e}{T^2} Z_w^{-1} \right) dh \qquad (1)$$

In which, T and e represent the absolute temperature in [K] and the partial pressure of water vapour in [hPa] in the atmosphere, respectively; k_2', k_3 are the atmospheric refractivity coefficients determined by Rüeger (2002). Finally, Z_w^{-1} is the inverse of compressibility factor of atmospheric water vapour that indicates the deviation of the atmosphere from an ideal gas.

2.3 Random Forest

For this work, the random forest (RF) algorithm was chosen for modelling. The selection of RF was based on the results obtained by Li, Yuan and Jiang (2023) and Li et al. (2024), who also proposed the generation of machine-learning models for the estimation of ZWD, using data from radiosondes.

The RF technique consists of a combination of learning methods for classification, regression, and other tasks that operate by building a multitude of decision trees at the time of training. For sorting tasks, the random forest output is the class selected by the majority of trees. For regression tasks, the mean predictions of the individual trees are returned. This is a very interesting method because random decision forests correct the tendency of decision trees to overfitting their training set (Breiman, 2001).

Before estimating the ZWD values, the data was filtered based on the quality of the radiosonde data. The raw data was filtered to remove noisy profiles, which made it impossible to calculate the ZWD due to the low amount of atmospheric information collected. In this way, only stations with the least amount of missing data in the selected period were used, in order to obtain greater continuity. Then, with the removal of outliers caused by failures in data collection with the radiosonde, modelling was carried out using the RF algorithm in the R environment. For this task, the packages Caret and Random Forest were used (Liaw & Wiener, 2002; Kuhn, 2015; R Core Team, 2017). The number of trees constructed was 500, and the number of variables selected in each node of the tree was one-third of the number of input data. The input data, that is, predictor variables, to estimate the ZWD (output variable in m) were day of the year, ZHD (m), average temperature (°C), Dew point temperature (°C), Atmospheric pressure (hPa), mean relative humidity (%), Latitude (°dec), Longitude (°dec), Altitude (m), Surface water vapour, Surface wet delay.

2.4 Model Validation

The observations were divided in a 70:30 ratio, with 70% being used for model training and 30% for validation. The use of cross-validation was also defined in order to evaluate the performance of the generated model in an unseen dataset.

Finally, the 10 stations at INMET, presented in Table 1, were also used in order to perform a validation with data that the model did not have contact during the training. Thus, the aim was to verify if the input variables used for the validation of the model are consistent with real-use situations. It is worth noting that these stations were selected so that they had geographic proximity to radiosondes and data availability during the same period from 2017 to 2022.

2.5 Variable correlation analysis

To avoid using highly correlated variables, which would not add much information to the trained model, a correlation analysis was performed to identify the predictors that are highly correlated with each other. Predictors with correlations higher than 0.70 or lower than -0.70 were eliminated (Figure 2).

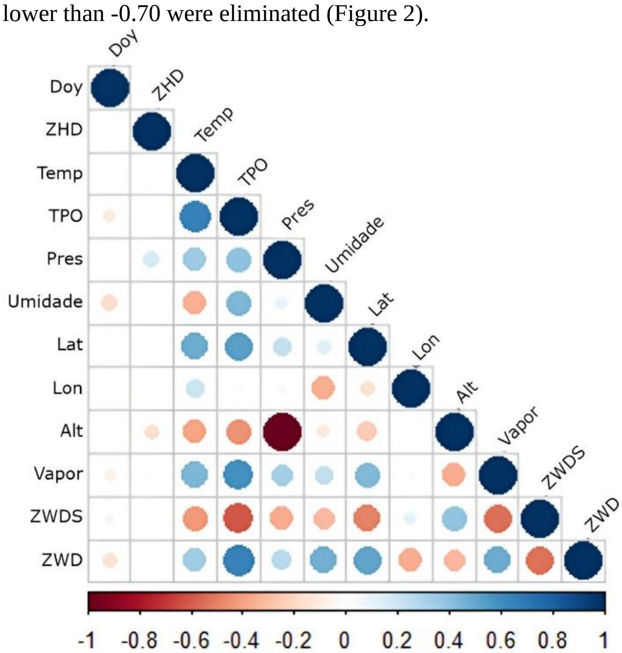

Figure 2. Correlation between predictor variables and response variables.

3. Results

3.1 Validation of the trained model

Table 2 presents the statistics obtained in the validation of the model, including the Root Mean Square Error (RMSE) in meters, the coefficient of determination (R^2) of the model and the Mean Absolute Error (MAE).

Model	RMSE (cm)	R^2	MAE (cm)
Random Forest	3.82	0.81	2.94

Table 2. Results of the validation of the model for the determination of ZWD from Random Forest.

The model presents 3.82 cm of RMSE, showing that the predictions of the RF-trained model present this deviation in relation to the values of the sample selected for validation. This value is within the expected range since the difference between models used in practice is centimetric (Wu et al., 2024). Related to the R^2, a value of 0.81 suggests that the model explains about 81% of the variability in the data, which is close to what was obtained by Wu et al. (2024), which was approximately 90%. The model presented by Wu et al. (2024) demonstrated improved performance, likely credited to the density of stations used in training, which provided extensive geographical variability of the data and positively contributed to model training. Additionally, the Brazilian territory exhibits high atmospheric variability due to the Amazon rainforest, characterized by complex variations in temperature and humidity that may pose challenges to model training.

Figure 3 shows the dispersion analysis between the observed and predicted values in the validation. In general, it can be seen that the dots are fairly evenly distributed along the trend line, indicating the predictions have good accuracy and without expressive bias. In addition, the presence of outliers is not observed.

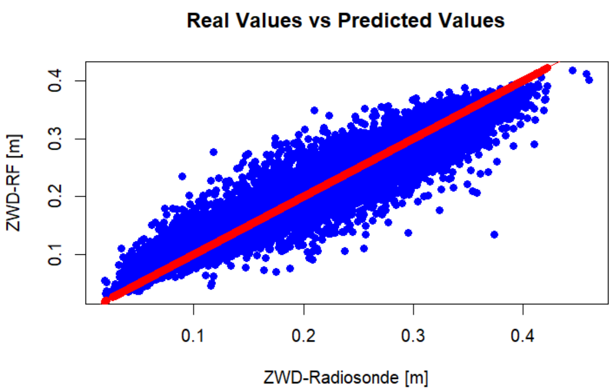

Figure 3. Dispersion of the predicted ZWD values from the model and the real values.

Based on the time series of data predicted and used for the training (Figure 4), it is possible to observe that the model behaves similarly to the real data. This fact becomes clear when noticing the great overlap between the values presented.

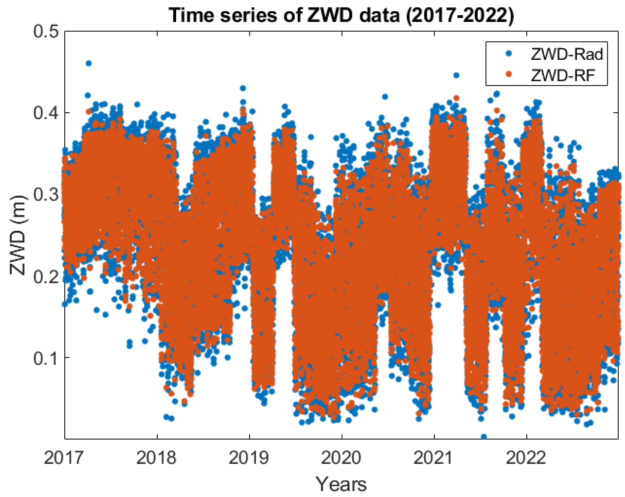

Figure 4. Time series of ZWD data estimated from the model and calculated in radiosonde.

In order to complement the previous analysis, a quantile evaluation was performed on the predicted data provided for validation (Figure 5). When looking at the graph, it is noted that the model faced significant challenges in determining values at the extremes (less than 10 cm and above 40 cm.), explaining the coefficient of determination obtained. In general, it is observed that a part of the predicted values follows the diagonal line, while many values deviate significantly from this line, indicating that the set does not follow the theoretical normal distribution.

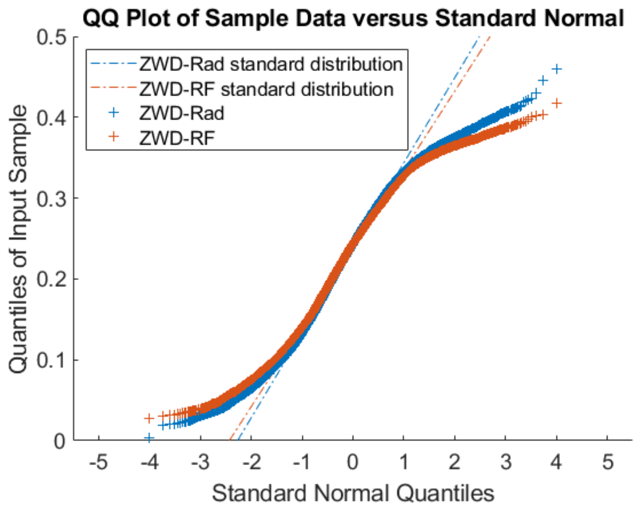

Figure 5. Quantile analysis for the ZWD data estimated and calculated in radiosonde.

3.2 Surface Station Testing (INMET)

INMET's surface meteorological stations were used as *inputs* to the ZWD determination model. The model evaluated in 10 stations obtained a maximum RMSE of 4.8 cm in the Santa Maria region (RS), while the lowest value was approximately 2.7 cm in Brasília (DF), as shown in Table 3. From these values, it can be concluded that the RMSE value obtained for the RF model matches the results obtained when using a different source of information, such as surface weather stations. On the other hand, the bias shows that, except in the region of Boa Vista – RR, it tends to be mostly negative, which means that the predictions of the trained model tend to underestimate the values used in the validation. In other words, the model is predicting lower values, which was observed in Figure 3, in which the data respect a certain range.

Local	Stations		RMSE (cm)	BIAS (cm)
	Radiosonde	INMET		
Manaus (AM)	SBMN	A101	3.13	-1.63
Natal (RN)	SBNT	A304	3.27	-0.16
Brasília (DF)	SBBR	A001	2.69	-0.10
São Paulo (SP)	SBMT	A701	4.81	-2.37
Santa Maria (RS)	SBSM	A803	4.87	-1.08
Belém (PA)	SBBE	A201	2.87	-0.54
Corumbá (MS)	83554	A724	5.60	-3.75
Porto Velho (RO)	SBPV	A925	5.03	-1.61
Boa Vista (RR)	SBBV	A135	4.86	1.78
Cruzeiro do Sul (AC)	82705	A705	3.84	-1.89

Table 3. INMET surface stations tested on the model with RMSE and BIAS calculated from the nearest radiosonde.

Figure 6 shows the locations of the 10 INMET stations used in this study, and their respective RMSE, in order to verify any spatial pattern and the behaviour of the ZWD estimation. This fact is not easy to verify due to the low number of stations. However, it was observed that the best RMSE values are concentrated in the north region, while the highest are in the south. This may be due to insufficient information for modelling the region, since it is a region of high atmospheric variability.

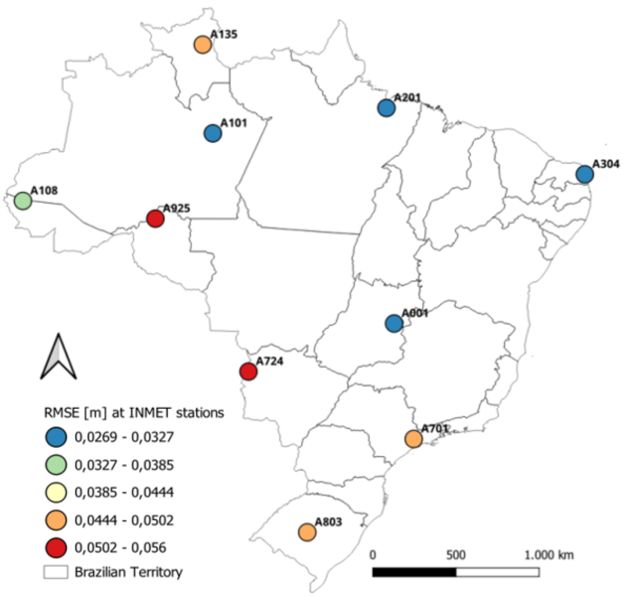

Figure 6. Cartographic representation of the behaviour of the RMSE for the INMET stations analyzed.

4. Conclusion

The wet component of atmospheric delay is an extremely important parameter for geodetic and meteorological studies. ZWD represents not only an error assessment in GNSS delay studies but also a crucial variable in climate studies due to its association with the water vapour. In this work, a decision tree-based model was trained to estimate the wet component using only surface input values obtained by radio sounding. The results showed an accuracy of approximately 3.8 cm based on the RMSE of the model, obtained in the data validation, using the *cross-validation* technique, with 30% of the observations used for validation. In addition, the model presented a coefficient of determination (R^2) of 81% and an MAE close to 2.9 cm, characterizing the model's fit with the data as within the expected.

The quantile analysis of the time series revealed that the model faced significant challenges in identifying extreme values, particularly those exceeding 40 cm of delay in the wet component. This result shows that the model fails to predict extreme events. Subsequently, an analysis was performed using 2023 data from INMET surface weather stations as atmospheric input parameters to test the model with data from another technique. The tests were carried out in 10 locations around Brazil, where the stations were selected based on the availability of data and proximity of the radiosonde stations, looking for stations with a distance of less than 10 km.

The results obtained using INMET data to obtain the ZWD showed values consistent with the model’s statistical evaluation. The region with the highest accuracy observed was Brasília (DF), with an MSR close to 2.7 cm. On the other hand, the municipality of Corumbá (MS) had the highest RMSE value of approximately 5.6 cm.

Based on the results obtained in this work, the model trained with the Random Forest technique indicates to be a useful tool for determining delay values of the wet component. The model consists of using surface atmospheric parameters as input values. This possibility represents an important advance since the density of surface stations is much higher compared to radiosonde stations. Although the radiosonde has a high quality due to its in-situ measurements, there is a low density of stations and a high launch cost. In addition, the surface stations have hourly sampling, making it possible to estimate the ZWD up to 24 times a day.

The random forest method offers several advantages over traditional models. Firstly, it is a robust ensemble learning technique that combines multiple decision trees to improve prediction accuracy and generalization. Unlike the deterministic nature of traditional models, RF can handle non-linear relationships and interactions between variables more effectively. This capability allows it to capture the complex dependencies between atmospheric conditions and ZWD, potentially leading to more accurate and reliable estimates.

Given the significance of ZWD for studies in Geodesy and Meteorology, future research should focus on expanding and enhancing the model by increasing the number of stations and extending data coverage across South America. Enhancing the model is considered crucial through the integration of additional information sources in training, such as ZWD-GNSS, Radio Occultation, and other remote sensing techniques for obtaining water vapour values. The enhanced model can be tested in geodetic and meteorological applications, including nowcasting rainfall events. Furthermore, future work can explore alternative machine learning techniques, which may yield different results due to the improved spatialization of training data. The development of models capable of accurately determining the wet component can significantly contribute to understanding atmospheric behaviour, thereby facilitating data acquisition in remote and challenging areas.

Acknowledgments

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) (88887.310313/2018-00, 88887.898553/2023-00 and 88887.961778/2024-00), by the National Council for Scientific and Technological Development, CNPq (grant n. 303670/2018-5, 151351/2019-8, 308747/2021-6, 306112/2023-0 and 116545/2023-2) and by the São Paulo Research Foundation, FAPESP (Grant: 2021/06029-7, 2021/05285-0 and 2023/14739-0) and the Czech University of Life Sciences Prague (CULS).

References

Bevis, M., Businger, S., Herring, T.A., Rocken, C., Anthes, R.A., Ware, R.H., 1992: GPS Meteorology: Remote Sensing of Atmospheric Water Vapor Using the Global Positioning System. *Journal of Geophysical Research*, 97, 15787–15801.

Breiman, L., 2001: Random Forests. *Machine Learning*, 45(1), 5–32. https://doi.org/10.1023/A:1010933404324.

Chen, J., Zhi, N., Liao, H., Lu, M., Feng, S., 2022: Global forecasting of ionospheric vertical total electron contents via ConvLSTM with spectrum analysis. *GPS Solutions*, 26(3), 1–10.

Davis, J.L., Herring, T.A., Shapiro, I.I., Rogers, A.E.E., Elgered, G., 1985: Geodesy by radio interferometry: effects of atmospheric modelling errors on estimates of baseline length. *Radio Science*, 20(6), 1593–1607.

De Oliveira, P.S., Morel, L., Fund, F., Legros, R., Monico, J.F.G., Durand, S., Durand, F., 2017: Modelling tropospheric wet delays with dense and sparse network configurations for PPP-RTK. *GPS Solutions*, 21(1), 237–250.

Elgered, G., Wickert, J., 2017: *Monitoring of the Neutral Atmosphere*. In: Teunissen, P.J.G., Montenbruck, O. (Eds.). *Springer Handbook of Global Navigation Satellite Systems*. Springer Handbooks. Cham: Springer International Publishing, 1109–1138.

Fund, F., 2009: Observations GPS et retards troposphériques: modélisations et application aux effets de surcharge océanique dans l'Ouest de la France. Ph.D. thesis, Université de Nantes.

Gao, W., Gao, J., Yang, L., Wang, M., Yao, W., 2021: A novel modelling strategy of weighted mean temperature in China using RNN and LSTM. *Remote Sensing*, 13, 3004.

Gouveia, T.A.F., 2013. Modelagem Neutrosférica sobre a América do Sul baseada em PNT e assimilação de dados locais utilizando observações GNSS. Master's dissertation, Faculdade de Ciências e Tecnologia, Universidade Estadual "Júlio de Mesquita Filho", Presidente Prudente, Brazil.

Gouveia, T.A.F., Monico, J.F.G., Alves, D.B.M., Sapucci, L.F., Nievinski, F.G., 2020: 50 years of synergy between Space Geodesy and Meteorology: from GNSS positioning error to very short-term precipitation forecasting applications. *Revista Brasileira de Cartografia*, 72, 1509-1535.

INMET, 2024: Maps. National Institute of Meteorology (INMET). URL https://mapas.inmet.gov.br/ (accessed 3.10.24).

Jacob, D.J., 1999: *Introduction to atmospheric chemistry*. Harvard University.

Kuhn, M., 2015: Caret: Classification and regression training. Astrophysics Source Code Library, ascl-1505.

Liaw, A., Wiener, M., 2002: Classification and regression by randomForest. R News, 2(3).

Lima, T.M.A., 2019. Modelagem Troposférica: Estudo e Análise de Diferentes Modelos. Undergraduate Research Report, FAPESP – IC, Faculdade de Ciências e Tecnologia, Universidade Estadual "Júlio de Mesquita Filho", Presidente Prudente, Brazil.

Li, Q., Böhm, J., Yuan, L., Weber, R., 2024: Global zenith wet delay modeling with surface meteorological data and machine learning. *GPS Solutions*, 28, 57.

Li, Q., Yuan, L., Jiang, Z., 2023: Modeling tropospheric zenith wet delays in the Chinese mainland based on machine learning. *GPS Solutions*, 27, 171.

Lu, C., Li, X., Zus, F., Heinkelmann, R., Dick, G., Ge, M., Schuh, H., 2017: Improving BeiDou real-time precise point positioning with numerical weather models. *Journal of Geodesy*, 91(9), 1019–1029.

MathWorks, 2024. MATLAB: The Language of Technical Computing. The MathWorks, Inc. URL: https://www.mathworks.com/products/matlab.html (accessed 1.01.24).

Mendes, V.B., 1999: Modeling the Neutral-Atmosphere Propagation Delay in Radiometric Space Techniques. Ph.D. thesis, University of New Brunswick, Dept. of Geodesy and Geomatics Engineering, Fredericton, N.B., Canada, Technical Report 199.

Monico, J.F.G., 2008: *Posicionamento pelo GNSS: Descrição, fundamentos e aplicações*. 2nd ed. São Paulo: Editora UNESP.

Nievinski, F.G., 2009: Ray-tracing options to mitigate the neutral atmosphere delay in GPS. Master's dissertation, University of New Brunswick, Canada.

QGIS Development Team, 2024. QGIS Geographic Information System. Open-Source Geospatial Foundation. URL: https://www.qgis.org (accessed 1.01.24).

R Core Team, 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing. URL: https://www.R-project.org (accessed 2.01.24).

Sapucci, L.F., Bastarz, C.F., Cerqueira, F., Avanço, L.A., Herdies, D.L., 2014: Impacto de perfis de Rádio Ocultação GNSS na qualidade das previsões de tempo do CPTEC/INPE. *Revista Brasileira de Meteorologia*, 29(4), 551–567.

Sapucci, L.F., 2001. Estimativa do vapor d'água atmosférico e avaliação da modelagem do atraso zenital troposférico utilizando

GPS. Master's dissertation, Faculdade de Ciências e Tecnologia, Universidade Estadual Paulista, Presidente Prudente, Brazil.

Sapucci, L.F., Machado, L.A.T., Monico, J.F.G., 2006: Previsões do atraso zenital troposférico para a América do Sul: variabilidade sazonal e avaliação da qualidade. *Revista Brasileira de Cartografia*.

Sapucci, L.F., Souza, E.M., Machado, L.A.T., Campos, T., 2019: Global Positioning System Precipitable Water (GPS-PWV) Jumps Before Intense Rain Events: A Potential Application to Nowcasting. *Meteorological Applications*.

Seeber, G., 2003: *Satellite Geodesy: Foundations, Methods, and Applications*. 1st ed. Berlin, New York: Walter de Gruyter.

Spilker, J.J. Jr., 1996: *Tropospheric Effects on GPS*. In: Parkinson, B.W., Spilker, J.J. Jr., *Global Positioning System: Theory and Applications*. Cambridge: American Institute of Aeronautics and Astronautics, 1, 517–546.

Thayer, D., 1974: An improved equation for the radio refractive index of air. *Radio Science*, 9, 803–807.

Teunissen, P., Montenbruck, O., 2017: *Springer Handbook of Global Navigation Satellite Systems*. 1st ed. Trento: Printer Trento s.r.l.

University of Wyoming, 2024: Soundings. University of Wyoming. URL http://www.weather.uwyo.edu/upperair/sounding.html (accessed 3.10.24).

Vianello, R.L., Alves, A.R., 2000: *Meteorologia básica e aplicações*. Universidade Federal de Viçosa (UFV), Viçosa: Editora UFV.

Wallace, J.M., Hobbs, P.V., 2005: *Atmospheric Science: An introductory survey*. University of Washington, 2nd ed.

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

Xiong, P., Zhai, D., Long, C., Zhou, H., Zhang, X., Shen, X., 2021: Long short-term memory neural network for ionospheric total electron content forecasting over China. *Space Weather*, 19(4), e2020SW002706.