

Predicting GPS-based PWV Measurements Using Exponential Smoothing

Shilpa Manandhar^{†*}, Soumyabrata Dev^{†‡}, Yee Hui Lee^{*}, and Stefan Winkler[§]

^{*} School of Electrical and Electronic Engineering, Nanyang Technological University (NTU), Singapore

[‡] ADAPT SFI Research Centre, Trinity College Dublin, Ireland

[§] School of Computing, National University of Singapore (NUS)

Abstract—Global Positioning System (GPS) derived precipitable water vapor (PWV) is extensively being used in atmospheric remote sensing for applications like rainfall prediction. Many applications require PWV values with good resolution and without any missing values. In this paper, we implement an exponential smoothing method to accurately predict the missing PWV values. The method shows good performance in terms of capturing the seasonal variability of PWV values. We report a root mean square error of 0.1 mm for a lead time of 15 minutes, using past data of 30 hours measured at 5-minute intervals.

I. INTRODUCTION

Precipitable water vapor (PWV) values are an indicator of moisture content in the atmosphere and exhibit good correlation with rainfall events. There is a growing trend of using the PWV values derived from Global Positioning System (GPS) in detection and/or prediction of a rainfall event [1], [2]. There are algorithms reported in the literature which predict rainfall events with lead times starting from 5 minutes up to 6 hours. The prediction window is affected by the resolution and availability of the GPS-PWV data. Generally, GPS-PWV values can be derived with a resolution of 5 minutes [3]. However, there are missing PWV values at certain hours. In this paper, we address this issue of missing PWV values, by proposing a method to predict the PWV values based on the past PWV values.

II. GPS-BASED PWV MEASUREMENTS

A. Computing PWV

In this section, we briefly mention the methods to calculate PWV values from GPS measurements. The GPS signals are effected by two main delays in the troposphere layer of the atmosphere. They are Zenith Hydrostatic Delay (*ZHD*) and Zenith Wet Delay (*ZWD*). Out of these two delays, *ZWD* is due to the water vapor content of the atmosphere. Therefore, PWV is derived from the *ZWD* as shown by Eqs. (1,2) [4].

$$PWV = PI \cdot ZWD \quad (1)$$

$$PI = [-\text{sgn}(L_a) \cdot 1.7 \cdot 10^{-5} |L_a|^{h_{fac}} - 0.0001] \cdot \cos \frac{2\pi(DoY - 28)}{365.25} + 0.165 - 1.7 \cdot 10^{-5} |L_a|^{1.65} + f \quad (2)$$

[†] Authors contributed equally.

Send correspondence to Y. H. Lee, E-mail: EYHLee@ntu.edu.sg.

where L_a is the latitude, DoY is day-of-year, $h_{fac} = 1.48$ for stations from the northern hemisphere and 1.25 for the southern hemisphere. $f = -2.38 \cdot 10^{-6} H$, where H is the station height, which can be ignored for stations below 1000m in altitude. In this paper, the *ZWD* values are processed using the GIPSY OASIS software for a tropical IGS GPS station, ID: NTUS (1.30°N, 103.68°E), with a temporal resolution of 5 minutes.

B. Predicting PWV Values

Suppose p_1, p_2, \dots, p_t indicate the PWV values measured up to time t . We use triple exponential smoothing (TES) [5] to model the seasonal variations of the PWV values. The principal idea behind triple exponential smoothing is to apply exponential weights on the observations, with more weightage on recent observations. The TES weights are assigned on the level, trend and seasonal components of the time series. We use the 5 minute interval PWV values in the form of time series data, and use only the historical PWV values to predict the future PWV values after time t . The future PWV values p_{t+m} are modelled as:

$$p_{t+m} = s_t + mb_t + c_{t-L+1+(m-1) \bmod L}, \quad (3)$$

where L is the length of a season, s_t is the smoothed version of the constant part of observation, b_t is the best estimate of the linear trend, and c_t is the series of seasonal corrections. We benchmark our proposed method with two popular forecasting techniques – persistence model and average model. The persistence model assumes that the forecasted PWV value remains constant as the latest PWV value, and is modelled as $p_{t+m} = p_t$. The average model works under the assumption that the future PWV values is the same as the average of the historical PWV values. It is modelled as $p_{t+m} = \frac{1}{t} \sum_t p_t$.

III. RESULTS & DISCUSSIONS

In this section, we provide a detailed analysis ¹ of the forecasting of PWV values using exponential smoothing. The PWV values for the year 2010 are computed for our chosen station ID NTUS.

¹The code of all simulations in this paper is available online at <https://github.com/Soumyabrata/predicting-pwv>.

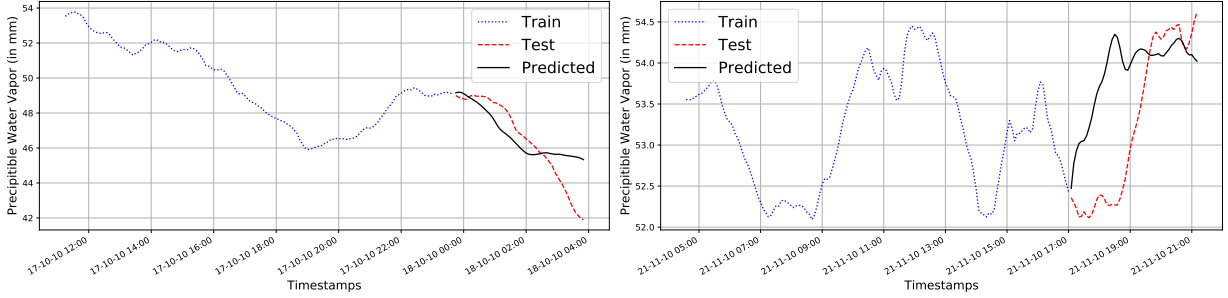


Fig. 1: Sample illustrations of the prediction of PWV values, along with the ground-truth values. We observe that the trend and seasonality of the future PWV values are captured with a good degree of accuracy.

A. Qualitative Evaluation

Our proposed method can efficiently capture the seasonal variation of the PWV values, and provide a foundation for short- and long- term forecasting. Figure 1 shows sample illustrations of the accuracy of PWV prediction. We use historical data of 30 hours to predict the future PWV values. We observe that our proposed technique can capture the *trend* of the future PWV values accurately.

B. Quantitative Evaluation

We use the Root Mean Square Error (RMSE) between the measured and predicted PWV values, in order to provide an objective evaluation of our proposed method. The performance of the prediction is dependent on two primary factors – the amount of historical data that is considered for training the time series model, and the length of lead times to the forecast data. We use a varying range of historical data and lead times, in order to understand the impact of these variables on the forecasting performance. Figure 2 shows the impact of these two independent variables.

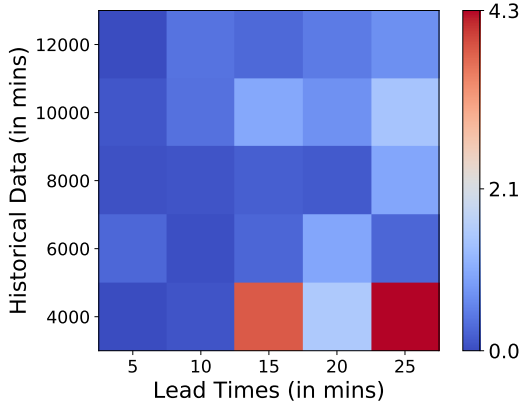


Fig. 2: Distribution of RMSE (mm) for a range of lead times and historical data.

The corresponding error is color coded for a particular value of historical data and lead time. We repeat this experiment 10 times for a chosen value of the two variables, in order to reduce any sampling bias. We observe that the error gradually increases with a lower value of historical data, and larger value of lead time. This makes sense as high amount of training data

is required to model the seasonality properly, and the error accumulates as we predict higher lead times.

As a final comparison, we benchmark our proposed method with two popular baseline models, namely persistence and average. Table I reports the average RMSE values (in mm) of the different methods in our dataset. We observe that the average model performs very poorly. Our proposed method shows a consistent improvement over the persistence model. This is due to the fact that PWV values remain fairly constant for shorter lead times. The forecasting performance can be further improved by incorporating other sensor data in addition to historical PWV values.

TABLE I: RMSE (mm) for varying lead times.

Lead Time	Proposed	Persistence	Average
5 min	0.061	0.086	10.433
10 min	0.078	0.144	9.525
15 min	0.101	0.259	7.028

IV. CONCLUSION & FUTURE WORK

This paper applies an exponential smoothing method for predicting future PWV values using past PWV data. The exponential method shows better performance compared to the two other techniques. The RMSE values increase with longer lead time and less historical data. In future work, other meteorological data [6] will be included for a better prediction of PWV values.

REFERENCES

- [1] S. Manandhar, Y. H. Lee, Y. S. Meng, F. Yuan, and J. T. Ong, "GPS derived PWV for rainfall nowcasting in tropical region," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 8, pp. 4835–4844, Aug 2018.
- [2] S. Manandhar, S. Dev, Y. H. Lee, Y. S. Meng, and S. Winkler, "A data-driven approach to detecting precipitation from meteorological sensor data," in *Proc. International Geoscience and Remote Sensing Symposium (IGARSS)*, 2018.
- [3] S. Manandhar, S. Dev, Y. H. Lee, and Y. S. Meng, "On the importance of PWV in detecting precipitation," in *2018 USNC-URSI Radio Science Meeting (Joint with AP-S Symposium)*. IEEE, 2018, pp. 89–90.
- [4] S. Manandhar, Y. H. Lee, Y. S. Meng, and J. T. Ong, "A simplified model for the retrieval of precipitable water vapor from GPS signal," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 11, pp. 6245–6253, Nov 2017.
- [5] E. S. Gardner, "Exponential smoothing: The state of the art," *Journal of forecasting*, vol. 4, no. 1, pp. 1–28, 1985.
- [6] S. Dev, S. Manandhar, Y. H. Lee, and S. Winkler, "Multi-label cloud segmentation using a deep network," in *Proc. IEEE AP-S Symposium on Antennas and Propagation and USNC-URSI Radio Science Meeting*, 2019.