On the Importance of PWV in Detecting Precipitation

Shilpa Manandhar*, Soumyabrata Dev†, Yee Hui Lee* and Yu Song Meng‡

*School of Electrical and Electronic Engineering, Nanyang Technological University (NTU), Singapore
†The ADAPT Centre, School of Computer Science and Statistics, Trinity College Dublin, Ireland
‡National Metrology Centre, Agency for Science, Technology and Research (A*STAR), Singapore

Abstract—With a rapid increase in the number of geostationary satellites around the earth's orbit, there has been a renewed interest in using Global Positioning System (GPS) to understand several phenomenon in earth's atmosphere. Such study using GPS devices are popular amongst the remote sensing community, as they provide several advantages with respect to scalability and range of applications. In this paper, we discuss how GPS signals can be used to estimate the amount of water vapor in the atmosphere. Furthermore, we demonstrate the importance of such precipitable water vapor (PWV) in the atmosphere for the task of rainfall detection. We present a detailed analysis in our dataset of meteorological data of 3 years. Test dataset shows that use of PWV in rainfall detection helps to reduce the false alarm rate by almost 12%.

I. INTRODUCTION

Global positioning system satellites are now-a-days extensively used to study the total water content in the atmosphere (also known as precipitable water vapor) and surface air temperature [1]. This is important to study as PWV impacts the hydrological cycle and other weather phenomenon. In this paper [1], we study the importance of PWV for the task of rainfall detection. Using a set of weather- and temporal-parameters, we use a data-centric methodology to establish this fact.

II. GPS & PRECIPITATION

The amount of precipitable water vapor in the atmosphere can be estimated from the GPS signal delays [2]. The delay incurred by the GPS signals is often referred as the zenith wet delay (ZWD). We calculate the PWV values (measured in mm), from the ZWD delays via:

$$\text{PWV} = \frac{PI \cdot \delta L_w^o}{\rho_l}, \tag{1}$$

where δL_w^o is ZWD, ρ_l is the density of liquid water (1000 kg/m^3), and PI is a dimensionless constant. We compute this parameter PI using:

$$PI = [-1 \cdot sgn(L_a) \cdot 1.7$$
$$\cdot 10^{-5}|L_a|^{h_{fac}} - 0.0001] \cdot cos(\frac{(DoY - 28)2\pi}{365.25}) \tag{2}$$
$$+ [0.165 - (1.7 \cdot 10^{-5})|L_a|^{1.65}] + f.$$

[1]The source code of all simulations in this paper is available online at https://github.com/shilpa-manandhar/PWV-for-rainfall

In this equation, L_a is the latitude, DoY is day-of-year, h_{fac} is 1.48 for stations from northern- and 1.25 for stations from southern- hemisphere. The f is computed using $f = -2.38 \cdot 10^{-6}H$, where H is the station height. We compute the PWV values for the IGS GPS station (station ID: NTUS), situated at Nanyang Technological University (NTU). These PWV values have a resolution of 5 minutes and have shown to be a good indicator of rainfall [3]

In our experiments, the meteorological parameters recorded by Davis Instruments 7440 Weather Vantage Pro II at a particular rooftop ($1.3°$N, $103.68°$E) of NTU building are used. Different parameters like temperature, relative humidity, solar irradiance, wind speed, direction and rainfall rate are recorded. All these weather measurements are recorded at an interval of 1 minute. In addition to these weather sensor data, we also collect continuous stream of sky images using collocated whole sky imagers. These images captured by our sky camera [4], [5] provide a visual understanding of the various phenomenon in the earth's atmosphere. In our previous work [6], we used such sky camera images to detect the onset of precipitation.

III. RAINFALL DETECTION

The task of rainfall detection can be modeled as a supervised learning task, based on a set of features. We use a combination of weather- and temporal- variables as discriminatory features of this task. We consider 7 parameters as features – temperature, dewpoint temperature, relative humidity, solar radiation, PWV, time of day and day of year. These variables are used as features for training a Support Vector Machine (SVM) model in the detection of rainfall.

In our experiments, we consider all the weather observations for the year of 2010 and 2011 as the data observations for our SVM model. An independent year of 2012 is used for the validation set. However, this datasets are highly unbalanced as the number of *rain* observations is low, as compared to *no rain* observations. Therefore, we employ random down sampling technique to make the dataset balanced. We employ a 1:1 ratio for *rain* to *no rain* observations. In order to reduce random sample bias, we perform our experiments for randomly sampled 100 trials. The average result of these 100 trials are reported.

For an objective evaluation in understanding the importance of PWV, we compute two metrics – True Detection (TD)

TABLE I: Performance of rainfall detection for our SVM-based model. All values are indicated in percentage (%). For the case where PWV is considered, an increase in performance is indicated by ↑, and a decrease in performance is indicated by ↓.

	Without PWV feature		With PWV feature	
	True Detection (TD)	False Alarm (FA)	True Detection (TD)	False Alarm (FA)
Testing data	92.6	44.0	87.4 (↓)	32.2 (↑)
Validation data	88.2	33.6	85.8 (↓)	28.5 (↑)

percentage and False Alarm (FA) percentage. The metrics TD and FA are defined as:

$$TD = \frac{TP}{TP + FN} \qquad (3)$$

$$FA = \frac{FP}{TN + FP}, \qquad (4)$$

where TP, TN, FP and FN are true positive, true negative, false positive and false negative respectively, with respect to rainfall detection task. We consider 20% of the data in the balanced dataset of 2010 and 2011, as the training set. The remaining observations are considered as the testing set. We perform two experiments to prove the importance of PWV in detecting rainfall. We consider all the 7 considered features, and compute the TD- and FA- percentage. The results are evaluated for testing- and validation- dataset. Furthermore, we consider a 6 dimensional feature vector without considering the parameter PWV.

The results are reported in Table I. We observe that, if we consider PWV in our model, the false alarm performance greatly improves, as compared to the case where PWV is not considered. The FA decreases by 11.8% for testing data, and the decreases by 5.1% for the validation set. However, there is a slight degradation in the performance of true detection percentage. The TD slightly decreases by 5.2% and 2.4% for the testing- and validation- data respectively. These experiments are conducted via down-sampling techniques at 20% training dataset.

We also perform experiments with varying percentage of training dataset. Figure 1 shows the TD and FA percentages for varying percentage of training dataset. All these observations are the average of 100 trials, that are randomly chosen from the dataset. It is clear that the performance of false alarm is significantly better, when PWV is considered in the model, for all the varying training size. However, the performance of the true detection slightly reduces by considering PWV. for smaller training dataset size. This difference in performance of TD diminishes for large training dataset size. Furthermore, we observe an over-saturation effect after 60% of training size – the TD and FA performances remains the same. This is because no further discriminatory variability of the dataset gets incorporated in our model, with the increasing training size.

IV. CONCLUSION & FUTURE WORK

The detection of water vapor content in the atmosphere is important amongst the remote sensing community to understand several phenomenon in the earth's atmosphere. It
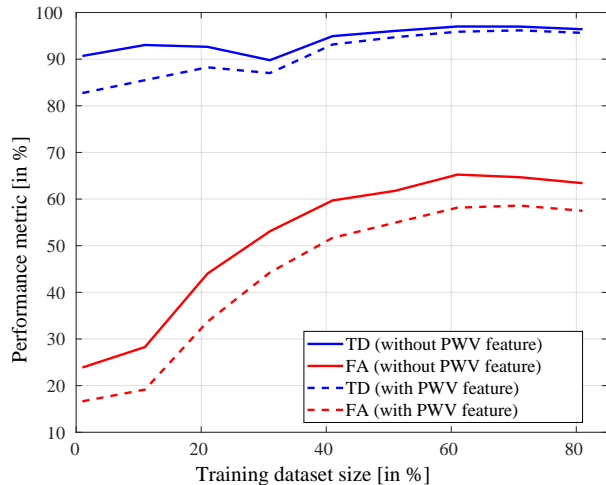

Fig. 1: True detection- and false alarm- percentage of rainfall recognition for varying percentage of training dataset.

is especially, important for the detection and prediction of rainfall. In this paper, we have demonstrated the importance of PWV for the task of precipitation detection. Using a machine-learning framework, we observed that the false alarm reduces, when PWV is considered as one of the discriminatory features. In the future, we also plan to integrate image-based features extracted from ground-based sky cameras [7] in a multi-modal fashion, for further improving the current benchmark.

REFERENCES

[1] M. Fujita and T. Sato, "Observed behaviours of precipitable water vapour and precipitation intensity in response to upper air profiles estimated from surface air temperature," *Nature Scientific Reports*, vol. 7, no. 1, pp. 4233, 2017.

[2] 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.

[3] S. Manandhar, Y. H. Lee, and S. Dev, "GPS derived PWV for rainfall monitoring," in *Proc. International Geoscience and Remote Sensing Symposium (IGARSS)*, 2016.

[4] S. Dev, F. M. Savoy, Y. H. Lee, and S. Winkler, "WAHRSIS: A low-cost, high-resolution whole sky imager with near-infrared capabilities," in *Proc. IS&T/SPIE Infrared Imaging Systems*, 2014.

[5] S. Dev, F. M. Savoy, Y. H. Lee, and S. Winkler, "Design of low-cost, compact and weather-proof whole sky imagers for High-Dynamic-Range captures," in *Proc. International Geoscience and Remote Sensing Symposium (IGARSS)*, 2015, pp. 5359–5362.

[6] S. Dev, S. Manandhar, Y. H. Lee, and S. Winkler, "Detecting rainfall onset using sky images," in *Proc. TENCON 2016 - 2016 IEEE Region 10 Conference*, 2016.

[7] S. Dev, B. Wen, Y. H. Lee, and S. Winkler, "Ground-based image analysis: A tutorial on machine-learning techniques and applications," *IEEE Geoscience and Remote Sensing Magazine*, vol. 4, no. 2, pp. 79–93, June 2016.