GPS Derived PWV for Rainfall Nowcasting in Tropical Region

Shilpa Manandhar, *Student Member*, *IEEE,* Yee Hui Lee, *Senior Member, IEEE*, Yu Song Meng, *Member, IEEE*, Feng Yuan, *Student Member*, *IEEE*, and Jin Teong Ong

Abstract— In this paper, a simple algorithm is proposed to perform the nowcasting of rainfall in the tropical region. The algorithm applies Global Positioning System (GPS) derived Precipitable Water Vapor (PWV) values and its second derivative for the short-term prediction of rainfall. The proposed algorithm incorporates the seasonal dependency of PWV values for the prediction of a rain event in the coming 5 minutes based on the past 30 minutes of PWV data. This proposed algorithm is based on the statistical study of 4 - year PWV and rainfall data from a station in Singapore and is validated using 2 - year independent data for the same station. The results show that the algorithm can achieve an average true detection rate and a false alarm rate of 87.7 % and 38.6 % respectively. To analyze the applicability of the proposed algorithm, further validations are done using 1 - year of data from one independent station from Singapore and 2 - years of data from one station from Brazil. It is shown that the proposed algorithm performs well for both the independent stations. For the station from Brazil, the average true detection and false alarm rates are around 84.7 % and 37 % respectively. All these observations suggest that the proposed algorithm is reliable and works well with a good detection rate.

Index Terms—Global Positioning System (GPS), Precipitable Water Vapor (PWV), Rainfall prediction, Rainfall nowcasting

I. INTRODUCTION

THE water-vapor stored in a column of unit cross section of the atmosphere is the Precipitable Water Vapor (PWV). The PWV is an important indicator of water vapor climatology and variability in the lower troposphere and related climate processes [1, 2]. PWV values can be determined through the measurement data using radiosondes, microwave-radiometers and satellite-based instruments. However, these technologies have limitations in capturing good resolution diurnal variations as they have low spatial-temporal resolutions [3, 4]. To overcome the drawbacks of

these systems, Global Positioning System (GPS) signal is now extensively being used to retrieve the PWV values. With the rapid deployment of GPS monitoring stations in local, regional, and global scales, ground-based GPS meteorology is widely used for the remote sensing purposes, with the precision improvements of GPS baseline and signal delay estimations [5].

Moreover, ground-based GPS meteorology offers much improved spatial and temporal resolutions for different parameters in sensing the atmosphere and ionosphere [6]. The GPS derived PWV is found to be useful in the study of different meteorological phenomenon like rainfall, typhoons, floods and thunderstorms [7-11] as this parameter is strongly linked to the hydrological cycle and dynamical processes especially in regions such as the tropics where the overall PWV is high [12].

A. Past Progress

Many researchers are studying the PWV values and its usefulness in prediction of rainfall. Takiguchi et al. in [13] reported that the onset of monsoon season in Thailand is followed by rapid growth in PWV values. In our previous study [14], we have analyzed the trend in PWV before and after a rainfall event (expressed as with respect to [w.r.t] rainfall in this paper). The time series plot of the PWV and the rainfall data was presented. Generally, the PWV values increases before a rain event. We also presented the cumulative distribution function (CDF) plots of PWV for rainy and non-rainy days in Singapore. In general, it was shown that the PWV values are higher on rainy days compared to non-rainy days. Similarly, Shi et al. in [15] presented some severe rainfall cases and a series of moderate rainfall cases to indicate the feasibility of GPS derived PWV for rainfall monitoring in China. The general trend of PWV w.r.t rain has been highlighted in [15] and a fixed PWV value is taken as a threshold to detect a rain event. It has been mentioned that the fixed PWV threshold value may not be suitable for different scenarios and seasonal factors can be considered to make it robust.

Literature also highlights that anomaly in PWV might be useful in the detection of rainfall. Li et al. in [16] suggested that PWV anomaly values along with other meteorological parameters like temperature could possibly be used for predicting rainfall. Sharifi et al. in [17] proposed to use the relative humidity anomaly along with the PWV anomaly values to improve the prediction accuracy of rainfall. However, these literatures are limited in providing insights in

GPS PWV temporal evolution w.r.t rainfall. Few literatures like Neelin in [18] reported the temporal correlation between microwave radiometer measured column water vapor and rainfall data for a tropical island of Nauru. The paper concluded that the high PWV can be associated to a strong pickup of precipitation even 10-12 hours into the future.

Benevides et al. in [19] presented a study of rainfall and GPS PWV for a temperate station of Lisbon. The paper proposed a simple algorithm to forecast rainfall within 6 hours after the steep increase in PWV. The algorithm was found to produce reasonable forecasts of precipitation with a success rate of 75 % but with a substantial amount of false alarms of 60-70 %. Yao et al. in [20] presented the precipitation forecast results in 6 hours using the data from sub-tropical stations of China. The results indicated a 7 % improvement in the correct detection rate as compared to [19], but with a comparable false alarm rate of 66 %.

B. Outline of the Contribution in this Work

In this work, our research focus will be the rainfall nowcasting in a tropical region such as Singapore, where most of the rainfalls experienced are convective rain events which are frequent in number and last only for short durations. The time duration between two rain events is then much smaller, and thereby results in a rainfall prediction time window that is relatively small. Therefore, different from the works discussed above, in this paper we perform a short-term rainfall prediction by using GPS derived PWV in tropical climate. An algorithm to nowcast the rain event in next 5 minutes given the PWV data of 30 minutes is proposed in this paper.

The proposed algorithm adopts two threshold parameters; season dependent PWV values and double derivative of PWV values. Significant diurnal and seasonal variations of PWV are observed over most of the International Global Navigation Satellite Systems Service (IGS) stations [21, 22]. The range of PWV values in temperate region is typically 0-45 mm and in a sub-tropical region is 0-80 mm as reported in [19] and [20] respectively, while the PWV values in tropical climate is typically from 30-70 mm. The narrower range and mostly higher PWV values in tropical region is because, the PWV values have distinct seasonal and diurnal cycles and the seasonal fluctuations is very less in a tropical region with higher temperature and relative humidity. These inherent characteristics of PWV should be considered when studying the usefulness of PWV values for the prediction of severe weather conditions which might lead to possible false alarms [23]. The proposed algorithm in this paper addresses the

TABLE I
GPS AND WEATHER STATIONS DETAILS

GPS Station Identifier (Lat, Lon)	Provider	Weather Station (Lat, Lon)	Common Data Availability (Years)
NTUS (1.34 ºN, 103.68 ºE)	IGS	NTU (1.34 ºN, 103.68 ºE)	2010 – 2015
SNUS (1.29 ºN, 103.77 ºE)	SiReNT	NUS (1.29 ºN, 103.77 ºE)	2016
SALU (2.59 ºS, 44.21 ºW)	IGS	São Luis (2.59 ºS, 44.23 ºW)	2016-2017

seasonal characteristics of PWV values to nowcast rainfall in a tropical climate, which has not been reported previously.

In the following, we will present the time series analysis of PWV values and its diurnal and seasonal variations, study the temporal evolution of PWV w.r.t rainfall, and perform nowcasting of rainfall using the PWV information from tropical stations. A simple algorithm for performing nowcasting of rain events within the next 5 minutes given the PWV information of the past 30 minutes will be proposed, with its performance validated.

II. DATABASE

A. PWV Data

Estimation of the PWV values from GPS signal is based on the delay incurred to the signals travelling from a GPS satellite to a ground receiver. The total delay caused by troposphere of the earth along the zenith path is known as Zenith Total Delay (ZTD), which can be divided into two parts: Zenith Hydrostatic Delay (ZHD) and Zenith Wet Delay (ZWD). ZWD is a function of the atmospheric water vapor profile, is an important parameter for the retrieval of PWV values from GPS signals and is represented as ΔL_w^o.

PWV values (in mm) are then calculated using the ZWD values as shown in eq. (1). Here, ρ_l is the density of liquid water (1000 kg/m^3), and *PI* is the dimensionless factor which theoretically explains the variations in PWV values that can be introduced by temperature. Typically, *PI* values can be calculated using mean temperature values determined from radiosonde profiles [24]. However, this method has a poor temporal resolution due to the limited availability of radiosonde data. An alternative empirical method, eq. (2), to calculate the *PI* values with improved temporal resolution was therefore proposed in [25].

$$pwv = \frac{PI \cdot \Delta L_w^o}{\rho_l},\qquad(1)$$

where,

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

In eq. (2), L_a is the latitude, DoY is the day-of-year, h_{fac} is 1.48 for stations from northern hemisphere and 1.25 for stations from southern hemisphere, and $f = -2.38 \cdot 10^{-6}\,H$ where H is the station height. The performance of eq. (2) has been successfully verified using 4 years of data from 174 different radiosonde stations including tropical, sub-tropical and temperate regions in [25].

For this paper, we processed the data for three tropical GPS stations. The details of these stations are summarized in Table I. There are two stations from Singapore and one from Brazil. Out of the two stations from Singapore, GPS station NTUS is located at Nanyang Technological University

(NTU) and is under International GNSS Service (IGS), and another GPS station, SNUS is located at National University of Singapore (NUS) and is under Singapore Satellite Positioning Reference Network (SiReNT) under Singapore Land Authority (SLA) [26]. The station from Brazil, SALU is located at São Luis and is under IGS. The IGS data are downloaded from NASA [27], and for SiReNT the data is requested from SLA [26].

The ZWD values for these stations are processed using GIPSY OASIS II software and recommended scripts [28]. PWV values for respective GPS stations are then calculated using (1) and (2), at a temporal resolution of 5 minutes. Since wet delay values in the zenith direction are used for estimating the PWV values, the PWV values at any time stamp reflects the average of 5 to 6 slant-path PWVs depending on the number of satellites visible at the given elevation mask. For this paper an elevation mask of 10° is used [15].

B. Weather Station Data

Rain data collected by the weather stations which are collocated with those GPS stations have been used for the analysis in this work. The locations of the weather stations are summarized in the Table 1. All these weather stations use tipping bucket rain gauges to record the rainfall rate. The rainfall data in NTU is recorded every 1 minute and the resolution of the tipping-bucket rain gauge is 0.2 mm/tip. The weather station at NUS has a temporal resolution of 5 minutes and the resolution of the tipping-bucket rain gauge is 0.2 mm/tip [29]. The rain gauge at São Luis has a resolution of 0.1 mm/tip and the data is informed in every 10-minutes slot if the rain is registered else the data is sent in an hour interval [30].

Here it is important to define a rain event and the time duration considered to distinguish it. It is well-noted that the rainfall cases in the temperate region are mainly stratiform rain and are commonly associated with low rainfall rates, whereas the rainfall in the tropical region are mainly convective rain and are commonly associated with high rainfall rates [31]. Studies show that more than 80% of the rain events in Singapore are convective rain events [31]. In general, these convective rainfall events are short lived with a lifespan of less than 30 minutes [32]. Therefore, in the following sections, all rainfall cases which are separated by a duration of 30 minutes are considered as a separate rain event. The rainfall rate recorded by the weather station is used in verifying the onset of a rain event.

In the following, data from the GPS station, NTUS, are used for the proposal of the rainfall nowcasting algorithm, and data from the GPS stations, SNUS and SALU, are used for the validation purposes.

III. PWV AND RAINFALL

In this section we first study the trend in PWV values from the NTUS GPS station with diurnal and seasonal variations and then we analyze the change in PWV values before a rainfall event starts. NTUS GPS station is located at

Singapore where there are four different seasons namely, North-East Monsoon (NE) generally during Nov-Dec and Jan-Mar, First-Inter Monsoon (FI) generally during Apr-May, South-West Monsoon (SW) generally during Jun-Oct and Second-Inter Monsoon (SI) generally Oct-Nov. The period of these seasons varies a little from year to year which is updated in yearly weather report provided by National Environment Agency (NEA), Singapore [33].

A. Representation of PWV and Rainfall data

To study the diurnal and seasonal variations in PWV, 4 years' (2010 - 2013) of PWV data from the NTUS GPS station are first divided according to the seasons; NE and SW monsoons and inter-monsoons (FI and SI). Then from each of these seasons, PWV values are separated based on rainy and non-rainy days. In this classification, any day with at least one rain event is defined as a rainy day. In four years, there is a total of 250 rainy and 275 non-rainy days in NE, 222 rainy and 357 non-rainy days in SW and 191 rainy and 166 non-rainy days in inter-monsoon seasons.

The PWV data of all the rainy days are composited w.r.t the time of the day. All the PWV values at each time interval is then represented by a boxplot. PWV values are available every 5 min but to make the plot readable, the boxplots in Fig. 1 (a-c) are shown at every hour of the day. The red line in the boxplot represents the median, the upper bounding box represents the 75th percentile and the lower bounding box represents the 25th percentile of all the PWV values at the given hour of the day. The mean of all the rainy days' PWV values at the given time is also shown in cyan color with circle marker in Fig. 1 (a-c). The PWV values of all the non-rainy days of the respective seasons are averaged at each time stamp and are shown in Fig. 1 (a-c) by a solid curve in green.

The occurrences of a rain event at different time of the day for NE, SW and inter-monsoon seasons are plotted in Fig. 1 (d-f) respectively. There is a total of 399 rain events in NE, 280 rain events in SW and 234 rain events in inter-monsoon seasons.

B. Seasonal and Diurnal Variations in PWV

From Fig. 1 (a-c), it can clearly be observed that the average PWV values for rainy days are higher than that of the non-rainy days for all the monsoon and inter-monsoon seasons. It is interesting to note that although there is a clear shift in the amplitude of average PWV values, the average PWV values tend to follow similar trend based on the time of the day for both rainy and non-rainy cases for the respective seasons, which is mainly affected and dominated by the diurnal variations of ZTD values [34].

In case of NE and SW monsoon seasons, it can be observed that the average PWV for both seasons starts with higher PWV values, which drops in the early morning and then starts to rise again with the temperature as the sun shines. For the SW monsoon the peak of PWV is at around mid-afternoon, 12:00 hour, after which the PWV values falls until late evening. Whereas for the NE monsoon, the peak of PWV value is during the afternoon period of around 15:00 hour. A possible reasoning for such behavior in PWV is because, for the NE monsoon season, late afternoon rainfall
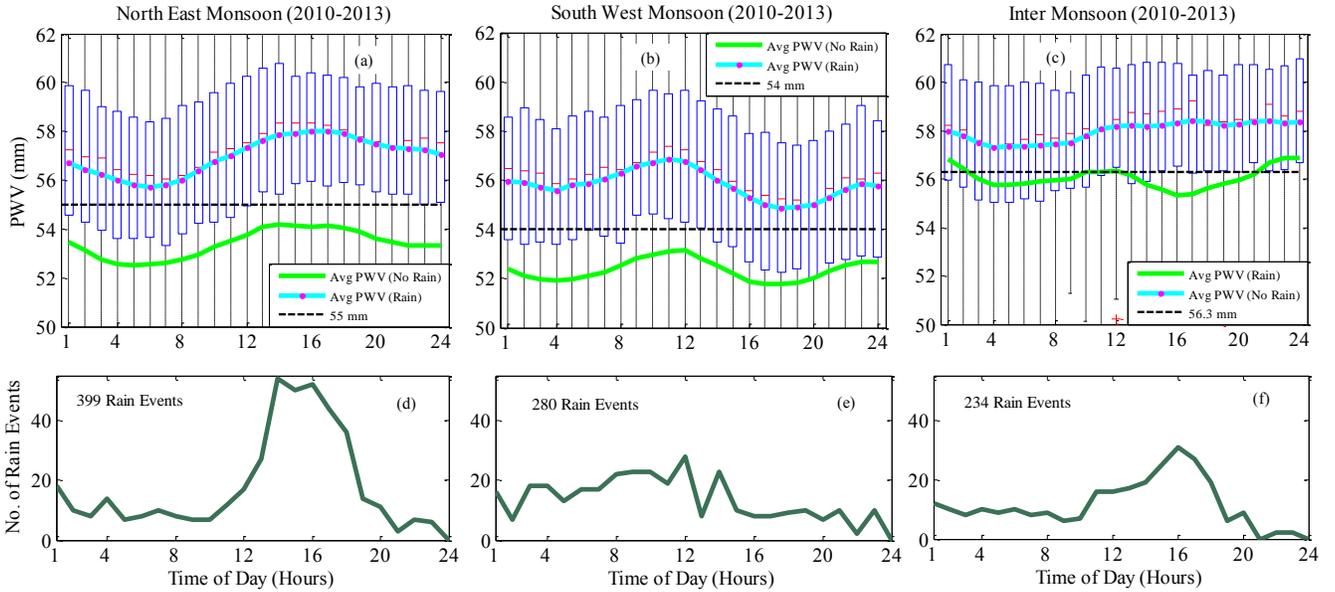
Fig. 1: Diurnal and seasonal variation in PWV and rainfall data for NTUS GPS station (2010 - 2013). The boxplots of PWV values with average of rainy (cyan line with circle markers) and non- rainy days (solid light green line) (a) NE, (b) SW and (c) inter-monsoons. Dashed black lines represent a threshold value of (a) 55 mm for NE, (b) 54 mm for SW and (c) 56.3 mm for inter-monsoons. The total number of rain events occurring at specific hours of the day (d) NE, (e) SW and (f) inter-monsoon seasons. The time for both GPS and weather station for NTUS station is recorded in local time format (UTC+8).

is often experienced, and thus the PWV values tend to increase around this time of the day. For the SW monsoon season, morning downpours are often experienced and thus the PWV values are higher in the morning. This is also illustrated in Fig. 1 (d-e), which show the total number of rain events of a specific season as a function of time of the day; Fig. 1 (d) represents the NE monsoon season and Fig. 1 (e) represents the SW monsoon season. Most of the rain events for the NE monsoon season occur during late afternoon whereas most of the rain evens for the SW monsoon season occur during morning and early afternoons.

As shown by Fig. 1 (c), PWV values during the inter-monsoon season are relatively higher compared to the monsoon seasons. There are several possible reasons as investigated. Firstly, in general, the inter-monsoon season has a higher temperature compared to that of the monsoon seasons. Inter-monsoon season can experience hot afternoons with maximum temperature going beyond 32 ℃ [33]. As the temperature is higher, it can relatively hold more water vapor increasing the overall PWV values. Secondly, it is observed that the rate of occurrence of rain events in inter-monsoon season is significantly higher compared to that in the monsoon seasons. Lastly, it needs to be considered that the NE monsoon season consists of both dry and wet phases. The driest month, generally February, falls under the NE monsoon season, which slightly lowers the overall PWV values of the NE monsoon compared to that of inter-monsoon season.

C. PWV Threshold Values

It is noted that a proper PWV threshold value plays an important role in indicating a rainy condition [15, 20]. In [15], a constant PWV threshold value of 50 mm is defined to indicate a rainy situation, and in [20] a range of PWV values from 23-59 mm is used as threshold values for different

months. It is clear that these threshold values are not suitable for a tropical climate, since the PWV values in tropical region can be much higher than these threshold values even for a non-rainy day. Thus, in this section it is of great interest for us to propose a method to define a PWV threshold value using the data from a tropical climate.

From Fig. 1 (a-c), a clear separation between average PWV values of rainy and non-rainy days can be observed. A constant PWV threshold value to indicate a rainy situation exists. Clearly, the PWV threshold value must change w.r.t the seasons as the inter-monsoon seasons have generally higher PWV values compared to the monsoon seasons. And amongst the monsoon seasons, in general, the NE monsoon experiences higher PWV values as compared to the SW monsoon. The average PWV values of rainy and non-rainy days also shows variations w.r.t the time of the day.

It is interesting to note from Fig.1 (a-c) that the maximum inter-quartile range for NE monsoon is 6.72 mm, for SW monsoon is 6.78 mm and for the inter-monsoon is 5.8 mm only. This indicates that the 50 % of the total number of PWV values at any instant of time is distributed within a small range of 6.4 mm in average. Since the 50 % of total number of PWV values are distributed in a small interquartile range, the rainfall prediction results can be very sensitive to the chosen PWV threshold value. A slightly higher PWV threshold value leads to a case where rain events can be undetected and therefore affect the true detection rate. Similarly, a slightly lower PWV threshold value can give rise to many false alarms.

Therefore, in this paper, a season dependent constant PWV threshold value is proposed to ensure high detection rate with a low false alarm rate, which satisfies the following two criteria;

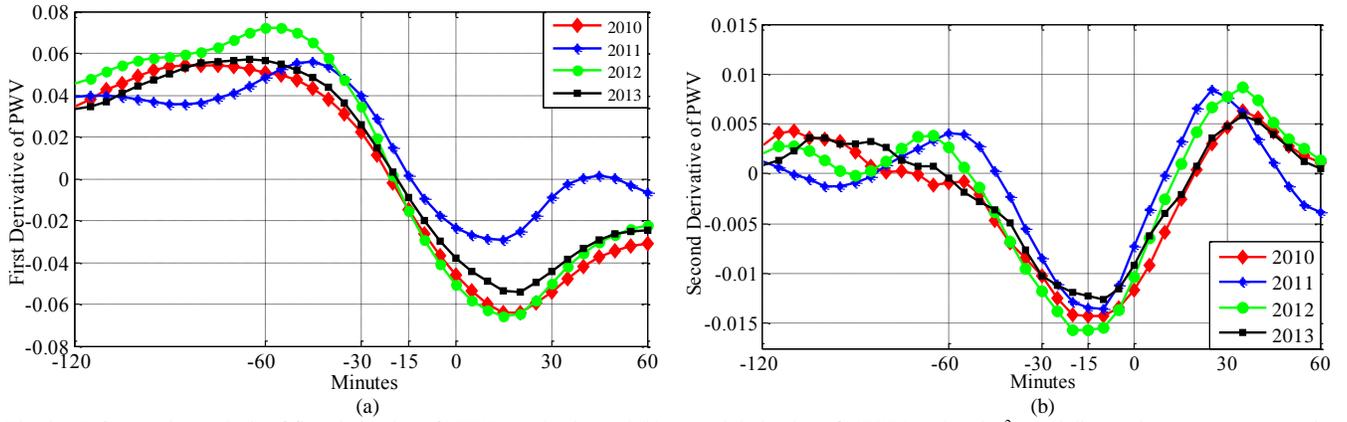*The criteria for choosing the threshold values,*

Fig. 2: (a) Composite analysis of first derivative of PWV (mm/5min) and (b) second derivative of PWV (mm/(5min)^2). Red diamonds represent mean values of PWV derivatives for 208 rain events of year 2010, blue stars for 234 rain events of 2011, green circles for 219 rain events of 2012 and black squares for 252 rain events of 2013. [Best viewed in color]

1) the threshold should be higher than the average PWV values of non-rainy days to make the false alarm as low as possible, and

2) the threshold should be lower than the 25th percentile of rainy days' PWV values at the specific hours (during which the probability of rain occurrence is higher) of respective seasons to maximize the true detection rate.

According to the second criteria, a PWV threshold value is chosen in such a way that it gives priority to a specific time-of-day when the probability of occurrence of a rain event is higher. Such specific hours differ according to the seasons. From Fig. 1 (d-f) and also from yearly weather reports for Singapore [33], it is clearly noted that, in NE monsoon, afternoon and early evening rainfalls are very frequent. In SW monsoon, Sumatra squalls are experienced during pre-dawn to midday, and short-lived rain falls are often experienced in the afternoon. During inter-monsoon seasons, afternoon to early evening rain events are common.

It should be noted that the threshold value varies from one location to another, depending on the local seasonal and climatic conditions [21]. Therefore, it is important to carefully study the distribution of PWV at a given location to choose a proper threshold value. In the next subsection we will carefully analyze the characteristics of PWV values w.r.t the rainfall initiation time and apply the findings in the rain prediction algorithm.

D. Composite Analysis of PWV Derivatives w.r.t Rainfall Initiation

In this section we consider all the rain events for the GPS station NTUS and study the PWV values w.r.t rainfall start time. To do so, first and second derivative of PWV values will be analyzed using the composite analysis method in [22].

1) First Derivative of PWV Values: The first derivative (FD) is calculated from the PWV values before and after the start time of each rain event. It is then composited in time, i.e., the mean of the first derivative values at every time stamp before and after the start of the rain event corresponding to all the rain events are calculated. The mean of the first derivative values are plotted in Fig. 2(a). The x-axis of the figure represents the time in minutes, where x=0 is the start of the rain, x<0 indicates minutes before the start

of the rain and x>0 indicates minutes during and after the start of the rain. Here, a 2-hour (120 minutes) time frame before the start of the rain and a 1-hour (60 minutes) time frame after the start of the rain are presented. The red diamonds represent the mean of the derivatives of the PWV values for all 208 rain events in the year 2010, blue stars represent the mean of the derivatives of the PWV values for all 234 rain events in the year 2011, green circles represent the mean of the derivatives of the PWV values for all 219 rain events of in the year 2012, and the black squares represents the mean of the derivatives of the PWV values for all 252 rain events in the year 2013.

From Fig. 2(a), it can be observed that the first derivative of PWV is positive until about 15 minutes before the start of the rain. This indicates that the PWV is accumulated in the atmosphere before the rain event. A significant increase in rate can be noticed after x= -120, which indicates in general for tropical regions, the rate of PWV accumulation increases around 2 hours before a rain event. It can be observed that the zero crossing of the first derivative occurs within 30 minutes before the start of the rain, after which, the PWV values start to decrease and derivative of the PWV values becomes negative. Here, the position of zero crossing of the derivative of the PWV values corresponds to the peak PWV values. Statistically, it is seen that for most of the rain events, the peak PWV values occur at around 15 minutes before the start of the rain. It is noted that the rate of change of first derivative of PWV values is sharp within 30 mins before the start of the rain. Thus, rate of change of first derivative (i.e., second derivative) of PWV before the start of the rain can be an important feature for short term rain prediction and will be the focus of our study in the next section.

2) Second Derivative (SD) of PWV Values: Like that of first derivative, the second derivative (SD) of the PWV values are calculated from the PWV values before and after the start of a rain event. The composite of the second derivatives are then plotted w.r.t time. Fig. 2(b) shows the composite analysis of the second derivative of PWV values.

In tropical climate, PWV values are generally higher and show very little day-to-day variations. That is the reason the fluctuations induced in PWV values before a rain event starts are prominent and are easy to be distinguished. This is also
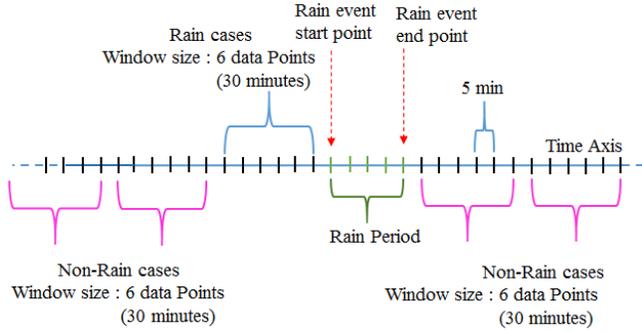
Fig. 3. Illustration of data points considered for rainy and non-rainy cases. A window size of 6 data points (30 min) is used for both the cases.

reflected in Fig. 2(a) and 2(b). In general, the rate of change of PWV (first derivative) does not vary much which results in a small or negligible amount of variation in the second derivative of PWV values. Fig. 2(b) shows that generally the second derivative of PWV values fluctuate around 0. A prominent feature in Fig. 2(b) is noticed when the rain starts, the second derivative values of PWV reaches a minimum value. There are significant fluctuations in the second derivative values 30 minutes before and 30 minutes after the start time of the rain event. During this one-hour period, the second derivative values are always less than 0.

This feature of second derivative value attaining a minimum value within 30 minutes before the rain start time will be used as a criterion for rain prediction.

IV. RAINFALL PREDICTION

In this section, we propose a simple algorithm to perform nowcasting of the start of a rain event. PWV data from year 2010 to year 2013 for GPS station NTUS are used in the proposal the empirical algorithm.

A. Rainfall Nowcasting Algorithm

Two different criteria will be used to perform nowcasting of the start of a rain event.

1) PWV value criteria: Season-dependent PWV threshold values are proposed by using the two criteria described in Section III (C).

Following the two criteria, for NTUS GPS station, for the NE monsoon season, a threshold value of 55 mm is chosen as the constant PWV threshold as shown by the dashed black line in Fig. 1 (a), and for the SW monsoon season, a threshold value of 54 mm is chosen as shown by the dashed black line in Fig. 1 (b). Similarly, for the inter-monsoon season, a threshold value of 56.3 mm is chosen as shown by the dashed black line in Fig. 1 (c) so that both the criteria are fulfilled.

2) Second Derivative (SD) value criteria: In Section III (D)-2, it was observed that a minimum SD value is reached within 30 minutes before a rain starts. We therefore use this minimum SD criteria to predict the start of a rain. For NTUS GPS station it has been noted that the second derivative values of PWV falls to a minimum of about -0.015 mm/(5min)^2 (ref. Fig 2 (b)).

TABLE II
TRUE DETECTION AND FALSE ALARM RATE WHEN USING THE PROPOSED ALGORITHM FOR RAIN PREDICTION FOR NTUS

Year	Total No. of Rain Events (NE + SW + Inter-Monsoon Rain Events)	Using Equation (3)	
		True Detection (%)	False Alarm (%)
2010	208 (76 + 74 + 58)	88.9	49.1
2011	234 (126 + 55 + 53)	85.9	37.2
2012	219 (89 + 72 + 58)	88.5	36.4
2013	252 (108 + 79 + 65)	84.9	43.7
Validation Results (Independent Years, NTUS)			
2014	231 (65 + 102 + 64)	89.1	34.6
2015	222 (92 + 63 + 67)	89.1	31.0
Avg		**87.7**	**38.6**

With both the PWV value and SD criteria, the rain prediction, P_{NTUS} for NTUS GPS station becomes true upon satisfying the given criteria as,

$$P_{NTUS} = \begin{cases} 1, & SD \leq -0.015 \ and \ PWV \geq 55 \ \ (NE) \\ & \qquad\qquad OR \\ & SD \leq -0.015 \ and \ PWV \geq 54 \ \ (SW) \\ & \qquad\qquad OR \\ & SD \leq -0.015 \ and \ PWV \geq 56.3 \ (FI \,\&SI) \\ 0, & Otherwise \qquad\qquad\qquad\qquad\qquad (3) \end{cases}$$

B. Evaluation Metrics

True detection rate and the false alarm rate are the two-evaluation metrics that is used to evaluate the proposed nowcasting algorithm.

1) True detection rate: Six PWV data points before the start of each rain event (corresponding to 30 minutes of PWV values) are considered. An illustration of this prediction process is given in Fig. 3. Now for NTUS GPS station, eq. (3) is applied on these 6 PWV data points. Firstly, the maximum of these PWV values is recorded. Then as per eq. (3), we check if this maximum PWV value exceeds the threshold of 55 mm for NE or 54 mm for SW or if this maximum PWV value exceeds threshold of 56.3 mm for the inter-monsoons. If the condition is satisfied, second derivative values are calculated using the same 6 PWV data points and the minimum second derivative value is recorded. Now if the second derivative value is less than or equal to -0.015 then the start of a rainfall event is treated as predicted. This will then be classified as a true detection. If the maximum PWV value and the minimum second derivative criteria are not satisfied together, the start of the rain event will be classified as no detection.

2) False alarm rate: All the non-rain days' PWV data points and all the data points excluding the data points within and 30 minutes before the start of the rain event are considered for non-rainy cases. All these non-rainy data points are then grouped using window sizes of 6 PWV data points such that each window sizes of 6 PWV data points such that each window corresponds to 30 minutes of PWV values. This can also be seen from the illustration diagram in Fig. 3. For each window, the maximum PWV value and the minimum second derivative values are calculated, i.e. eq. (3) is applied. If the criteria are satisfied, it is considered as a false alarm.
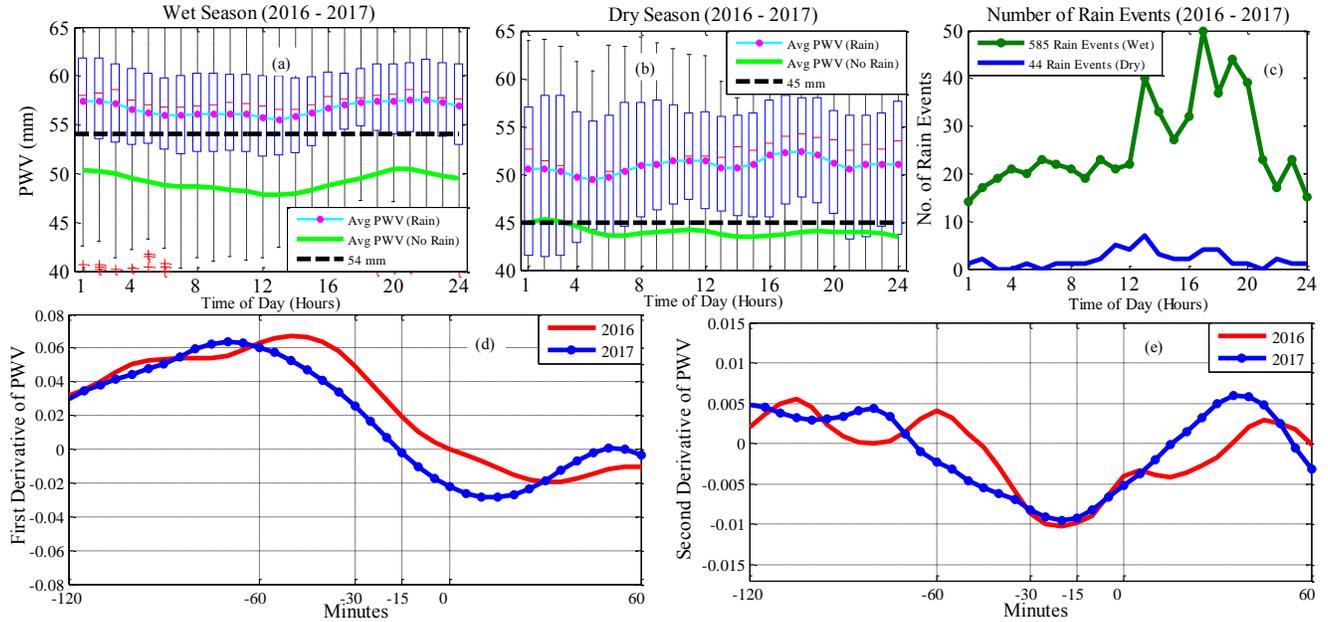
Fig. 4: Diurnal and seasonal variation in PWV and rainfall data for SALU GPS station (2016 - 2017). The boxplots of PWV values with average of rainy (cyan line with circle markers) and non- rainy days (solid light green line) (a) Wet, (b) Dry Seasons. Dashed black lines represent a threshold value of (a) 54 mm for dry and (b) 45 mm for wet season. (c) total number of rain events occurring at specific hours of the day for wet and dry seasons (d) First derivative values of 2 years. (e) Second derivative values of 2 years. The time for both GPS and weather station for SALU station is recorded in UTC format.

C. Results

Using the given procedure, the true detection rate and the false alarm rate have been calculated for all cases from the year 2010 to the year 2013 for NTUS GPS station. The results are summarized in Table II. It is observed that for different years, the true detection rate is very high, around 88.9 %, 85.9 %, 88.5 % and 84.9 % for 2010, 2011, 2012 and 2013 respectively. The corresponding false alarm rate is only 49.1 %, 37.2 %, 36.4 % and 43.7 % for the same years respectively.

Here, it is important to note that the certain percentage of the total false alarm rate might be the result of comparing the point rainfall rate recorded at NTU with the PWV values which is an average over a cone of 5 to 6 satellite paths. Therefore, some false alarms might not be false alarm but the detection of rainfall that might be occurring at a location some distance away from NTU along one or more of the satellite paths within the cone causing an increase in the average zenith PWV value. Thus, in such cases, the average zenith PWV values might satisfy the given criteria, indicating a buildup of rain clouds but the rain is not at the point location of NTU, resulting in a false alarm.

D. Performance Validation Using Data from Independent Years at the Same Station

For validation purposes, the proposed algorithm is applied on a set of independent PWV data for NTUS station. Data from year 2014 and 2015 which were not used in previous threshold derivation are used for performance validation here. It is found that the proposed prediction algorithm works well with a true detection rate of 89.1 % for both 2014 and 2015 and with a false alarm rate of only 34.6 % and 31.0 % for 2014 and 2015 respectively.

The six years' result show that the proposed algorithm has an average true detection rate of 87.7 % and false alarm rate of 38.6 % only. The true detection rate of 87.7 % suggests that the proposed algorithm can detect most of the rain events without significant misses. The true detection rate and the false alarm rate reported in this paper shows a significant improvement over the true detection and false alarm rates of 75 % and 60-70 % respectively as reported in [19] and the true detection and false alarm rates of 80 % and 66 % respectively as reported in [20].

Thus, the comparison results show that the proposed prediction method is reliably with better performance. It is also simple for implement and easy to use and is applicable for the nowcasting of the start of a rain event.

E. Performance Validation Using Data from Another Station in Singapore

In this subsection, one-year (2016) data from SNUS GPS station is processed for performance validation of the proposed method. Since both the stations, NTUS and SNUS are geographically close to each other (16 km apart), SNUS station shall experience similar weather variations as that of NTUS station. Therefore, we apply the thresholds in eq. (3) on the SNUS data where a total of 195 rain events were recorded in the year 2016, a true detection rate of 90.7 % and a false alarm rate of 50.2 % are observed. This is a meaningful improvement over recent works by Zhao et al. [35], where their proposed method with site-dependent thresholds in a similar scale give a true detection rate of at most 87.1 % and a false rate of at least 60.5 %. Therefore, the proposed threshold values of eq. (3) looks widely applicable (site-independent) within a small-scale region (validated with 16-km apart NTUS and SNUS stations).

Although, better results are obtained as compared to those in the literature, a slightly higher false alarm is observed for

SNUS station using the threshold values of eq. (3) from NTUS station. From our analysis, it is found that the PWV values, which is an average over a cone is generally slightly higher at the SNUS station as it is located much closer to the coastal area compared to the NTUS station which is at the western inland region of Singapore and backed by land (Malaysia) on its northern side. Therefore, when applying the threshold values of eq. (3) from NTUS station (with lower PWV values), the false alarm at SNUS becomes slightly higher.

V. PERFORMANCE VALIDATION OF THE PROPOSED METHODOLOGY FOR A STATION FROM BRAZIL

In this section, the proposed rainfall nowcasting algorithm is further validated using two-year data (2016-2017) at a GPS station from Brazil. This station, SALU, is located at São Luis of Brazil (another tropical region) with very different climatic conditions compared to that of Singapore. Therefore, the thresholds in eq. (3) which were derived from NTUS data in Singapore is not applicable here. Thus, in the following, we will first present the analysis of the properties of PWV and rain in São Luis, and then derive the new threshold values.

From the data analysis, the weather conditions in São Luis are broadly divided into dry seasons and wet seasons. The dry season lasts for very short duration which is during September to November, and the rest of the months generally falls under the wet season. April is the wettest month and October is the driest month.

Fig. 4 (a-b) show the distribution of two years (2016 - 2017) of PWV values for rainy and non-rainy days of wet and dry seasons respectively. The wet season has higher PWV values compared to the dry season. For wet seasons, the mean rainy PWV values are consistently higher, and there is an increment in the mean rainy PWV values of wet season in the evening hours where there is higher probability of rain occurrence as shown by Fig. 4 (c). The dry season, which lasts only for 3 months in a year has very less number of rain events. The PWV values are consistently lower for the dry season.

Fig. 4 (d-e) show the mean first and second derivative values of PWV before and after a rain starts for all the rain events of 2016 and 2017. Both Fig. 4 (d) and (e) have similar trends as those observed in Fig. 2 (a) and (b) respectively. The minimum value of second derivative occurs within 30 minutes before the start of a rain event. The minimum value is reported to be -0.01 mm/(5min)^2.

After applying the proposed rainfall nowcasting algorithm, a PWV threshold value of 54 mm (ref. to dashed line in Fig. 4(a)) and 45 mm (ref. to dashed line in Fig. 4(b)) are selected for the wet and dry seasons respectively, and the SD threshold value of -0.01 mm/(5min)^2 is chosen. It is interesting to note that the PWV threshold value for the rainy season in Brazil is similar to that for the monsoon season in Singapore, and similar observation also holds for SD threshold value.

Therefore, with both the PWV value and SD conditions, the rain prediction, P_{SALU} for SALU GPS station becomes true upon satisfying the given criteria as,

TABLE III
VALIDATION OF PROPOSED ALGORITHM FOR A STATION FROM SINGAPORE AND BRAZIL

Year	Number of Rain Events	True Detection (%)	False Alarm (%)
Station from Singapore: SNUS			
2016	91 (NE) + 51 (SW) + 53 (Inter)	**90.7**	**50.2**
Station from Brazil: SALU			
2016	232 (Wet) + 16 (Dry)	82.6	38.5
2017	309 (Wet) + 28 (Dry)	86.9	35.5
Avg		**84.7**	**37**

$$P_{SALU} = \begin{cases} 1, & SD \leq -0.01 \ and \ PWV \geq 54 \ \ (Wet) \\ & OR \\ & SD \leq -0.01 \ and \ PWV \geq 45 \ \ (Dry) \\ 0, & Otherwise \end{cases} \quad (4)$$

The eq. (4) is applied on PWV and rain data from SALU (2016 & 2017) and the results are tabulated in Table III. From Table III, it can be observed that, the rain events are predicted with a true detection rate of 82.6 % and a false alarm rate of 38.5 % for 2016 and a true detection rate of 86.9 % and a false alarm rate of 35.5 % for 2017.

For SALU GPS station, the wet season has a total of 541 rain events; 232 in 2016 and 309 in 2017, whereas the dry season has only 44 rain events; 16 in 2016 and 28 in 2017. This is different from Singapore, where all the monsoon seasons and inter-monsoon seasons experience equitable number of rainfall events. Therefore, it is also interesting to access the performance based on different seasons separately.

When the algorithm is applied onto only wet seasons of 2016 and 2017, a true detection rate of 84.4 % and a false alarm rate of 26.8 % were determined. Similarly, when it was applied only onto the dry seasons from 2016 and 2017, a true detection rate of 89.7 % and a false alarm rate of 42.7 % were observed. The proposed algorithm shows better performance for the wet seasons with constantly lower false alarm rate as compared to that of dry season. From rainfall prediction point of view, wet seasons are more important as most of the rain events fall under wet seasons.

The results show that the proposed algorithm works with good accuracy for a tropical station from Brazil as well, and thus it is widely useful and applicable for tropical stations.

Here it is important to point out that the PWV threshold values clearly vary from seasons to seasons and the values will change for different locations (large scale, ref. to eq. (3) and (4)). The proposed rainfall nowcasting algorithm can help to derive a priori value of threshold in a location for a particular season, which can be used to select the threshold for other events, which makes the process to forecast easier [20].

VI. CONCLUSION & FUTURE WORKS

In this paper, a detailed study of the PWV values w.r.t rainfall is performed and a simple, yet effective algorithm is proposed for prediction of the start of a rain event in the tropical region.

The proposed prediction algorithm implements a season dependent PWV threshold value, which is optimized by considering the time-of-day when most of the rain events occur for different seasons and a second derivative threshold, to predict the start of the rain within the next 5 minutes using data from the past 30 minutes in a tropical region. The algorithm is derived based on the data from Singapore, NTUS GPS station. The results show that the proposed algorithm works well with a true detection of 87.7 % and a false alarm of 38.6 % in average for NTUS. The proposed algorithm is validated using the data from two more tropical stations SNUS and SALU. The proposed algorithm shows good accuracy for both the independent stations as well. Hence, the proposed prediction algorithm shows that the use of PWV and its second derivatives are good features for rainfall nowcasting.

In the near future, an extended work will be carried out to study and characterize the PWV values according to different types of rain; convective or stratiform. It is also interesting to study the use of PWV data for rainfall prediction in temperate region since the PWV values for temperate regions vary significantly over different seasons.

REFERENCES

[1] S. G. Jin, Z. Li, and J. Cho, "Integrated water vapor field and multiscale variations over China from GPS measurements," *J. Appl. Meteor. Climatol.*, vol. 47, pp. 3008-3015, Nov. 2008, doi: 10.1175/2008JAMC1920.1.

[2] S. G. Jin, J.U. Park, J.H. Cho, and P. H. Park, "Seasonal variability of GPS-derived Zenith Tropospheric Delay (1994-2006) and climate implications," *J. Geophys. Res.*, vol. 112, no. D9, Jun. 2007, doi: 10.1029/2006JD007772.

[3] K. Gui, H. Che, and Q. Chen, et. al, "Evaluation of radiosonde, MODIS-NIR-Clear, and AERONET precipitable water vapor using IGS ground-based GPS measurements over China," *Atm. Res.,* vol. 197, pp. 461 – 473, Nov. 2017.

[4] S. Abbasy, M. Abbasi, J. Asgari, and A. Ghods, "Precipitable water vapour estimation using the permanent single GPS station in Zanjan, Iran," *Meteorol. Appl.,* vol. 24, no. 3, pp. 415 – 422, Jul. 2017, doi: 10.1002/met.1639.

[5] S. G. Jin, O. Luo and C. Ren, "Effects of physical correlations on long-distance GPS positioning and zenith tropospheric delay estimates," *Adv. Space Res.,* vol. 46, no. 2, pp. 190 – 195, Jul. 2010

[6] S. G. Jin, G. P. Feng, S. Gleason, "Remote sensing using GNSS signals: Current status and future directions," *Adv. Space Res.,* vol. 47, no. 10, pp. 1645 – 1653, May 2011.

[7] P. Basili, S. Bonafoni, V. Mattioli, P. Ciotti, and N. Pierdicca, "Mapping the atmospheric water vapor by integrating microwave radiometer and GPS measurements," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 8, pp. 1657–1665, Aug. 2004.

[8] D. Seob Song and D. A. Grenjner-Brzezinska, "Remote sensing of atmospheric water vapor variation from GPS measurements during a severe weather event," *Earth Planets Space*, vol. 61, pp. 1117-1125, 2009.

[9] M. Muhammad, M. Abdullah, M.J. Singh, W. Suparta, M.T. Islam, F. Tangang, "Characterization of GPS PWV during flooding event over Keningau, Sabah," in *Proc. IconSpace*, Malaysia, pp. 429-433, Jul. 2013.

[10] Y. Liou and C. Y. Huang, "GPS observations of PW during the passage of a typhoon," *Earth Planets Space*, vol. 52, pp. 52 709-712, 2000.

[11] W. Suparta, J. B. Adnan and M. A. B. M. Ali, "The Impacts of GPS PWV during Heavy Thunderstorms over Kuala Terengganu, Malaysia," in *Proc. 2012 International Conference in Green and Ubiquitous Technology*, Jakarta, Indonesia, Jun. 2012.

[12] G. G. Amenu and P. Kumar, "NVAP and reanalysis-2 global precipitable water products: intercomparison and variability studies," *Bull. Amer. Meteor. Soc.*, vol. 86, no. 2, pp. 245–256, 2005.

[13] H. Takiguchi, T. Kato, H. Kobayashi, and T. Nakaegawa, "GPS observations in Thailand for hydrological applications," *Earth Planets Space*, vol. 52, pp. 913-919, 2000.

[14] S. Manandhar, Y. H. Lee and S. Dev, "GPS derived PWV for rainfall monitoring," in *Proc. 2016 IEEE Int. Geosci. Remote Sens. Symp (IGARSS),* Beijing, pp. 2170-2173, Jul. 2016.

[15] J. Shi, C. Xu, J. Guo and Y. Gao, "Real-Time GPS precise point positioning-based precipitable water vapor estimation for rainfall monitoring and forecasting," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 6, pp. 3452-3459, Jun. 2015.

[16] P. W. Li., X. Y. Wang, Y. Q. Chen and S. T. Lai, "Use of GPS signal delay for real-time atmospheric water vapour estimation and rainfall nowcast in Hongkong," *in Proc. First Int. Symp Cloud-prone & Rainy Area Rem. Sens.*, Hong Kong, Oct. 2005.

[17] M. A. Sharifi, A. S. Khaniani, M. Joghataei, "Comparison of GPS Precipitable water vapor and meteorological parameters during rainfall in Tehran," *Meteorol. Atmos. Phys.*, vol. 127, pp. 701-710, 2015.

[18] J. D. Neelin, "Temporal relations of column water vapor and tropical precipitaion," *J. Atmospheric Sci.,* vol. 67, pp. 1091-1105, Apr 2010.

[19] P. Benevides, J. Catalao, and P. M. A. Miranda, "On the inclusion of GPS Precipitable water vapour in the nowcasting of rainfall", *Nat. Hazards Easrth Syst. Sci.*, vol. 15, pp. 2605-2616, 2015.

[20] Y. Yao, L. Shan and Q. Zhao, "Establishing a method of short term rainfall forecasting based on GNSS-derived PWV and its application," *Scientific Reports,* vol. 7, no. 1, 2017, DOI:10.1038/s41598-017-12593-z.

[21] S. G. Jin and O. F. Luo, "Variability and climatology of PWV from global 13-year GPS observations," *IEEE Trans. on Geosci. Remote Sens.*, vol. 47, no. 7, pp. 1918–1924, Jul. 2009.

[22] G. Li, F. Kimura, T. Sato, and D. Huang, "A composite analysis of diurnal cycle of GPS precipitable water vapor in central Japan during Calm Summer Days," *Theor. Appl. Climatol.*, vol. 92, pp. 15-29, 2008. DOI 10.1007/s00704-006-0293-x.

[23] S. Bonafoni, R. Biondi, "The usefulness of the Global Navigation Satellite Systems (GNSS) in the analysis of precipitation events," *Atm. Res.,* vol. 167, pp. 15-23, 2016.

[24] M. Bevis, S. Businger, S. Chiswell et al., "GPS meteorology: Mapping zenith wet delays onto precipitable water," *J. Appl. Meteor.*, vol. 33, no. 3, pp. 379–386, Mar. 1994.

[25] 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.

[26] Singapore Satellite Positioning Reference Network, last accessed on Mar. 2018. [Online]. https://sirent.inlis.gov.sg/

[27] NASA's Earth Science Data Systems, Crustal Dynamics Data Information System, last accessed on Mar. 2018. [Online]. Available: ftp://cddis.gsfc.nasa.gov/pub/gps/data/

[28] S. Desai, D. Kuang, and W. Bertiger, "GIPSY/OASIS (GIPSY) overview and under the hood," Dept. Geomatics, Univ. Newcastle Upon Tyne, Newcastle upon Tyne, U.K., Oct. 1996.

[29] Geography Weather Station, National University of Singapore, Singapore. last accessed on Mar. 2018. [Online]. Available: https://inetapps.nus.edu.sg/fas/geog/

[30] Cemaden's Observational Network for Natural Disaster Risk Monitoring. last accessed on Mar. 2018. [Online]. Available: http://www.cemaden.gov.br/mapainterativo/

[31] J. X. Yeo, Y. H. Lee and J. T. Ong, "Performance of site diversity investigated through Radar derived results", *IEEE Trans. Antennas Prop.*, vol. 59, no. 10, pp. 3890-3898, 2011.

[32] J. X. Yeo, Y. H. Lee and J. T. Ong, "Use of weather radar rain cell motion forecasting for site diversity system", *in Proc. MILCOM*, Maryland, USA, November 2016.

[33] National Environment Agency, last accessed on Oct 1, 2017. [Online]. Available:http://www.nea.gov.sg/trainingknowledge/publications/annual-weather-review.

[34] S. G. Jin, O. F. Luo, and S. Gleason, "Characterization of diurnal cycles in ZTD from a decade of global GPS observations," *J. Geod.*, vol. 83, no. 6, pp. 537-545, Jun. 2009, doi: 0.1007/s00190-008-0264-3.

[35] Q. Zhao, Y. Yao, and W. Yao, "GPS-based PWV for precipitation forecasting and its application to a typhoon event," *J. Atmospheric Sol.-Terr. Phys.*, vol. 167, pp. 124 - 133, Jan. 2018.

Shilpa Manandhar (S'14) received the B. Eng. (Hons.) degree in electrical and electronics engineering from the Kathmandu University, Nepal, in 2013. She is currently pursuing her Ph.D. degree in the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore.

Her current research interests are in Remote Sensing and Propagation including the study of GPS signals to predict meteorological phenomenon like rain and clouds.

Yee Hui Lee (S'96–M'02–SM'11) received the B. Eng. (Hons.) and M. Eng. degrees in School of Electrical and Electronics Engineering from the Nanyang Technological University, Singapore, in 1996 and 1998, respectively, and the Ph.D. degree from the University of York, York, U.K., in 2002.

Dr. Lee is currently Associate Professor of the School of Electrical and Electronic Engineering, Nanyang Technological University, where she has been a faculty member since 2002. Her interests are channel characterization, rain propagation, antenna design, electromagnetic bandgap structures, and evolutionary techniques.

Yu Song Meng (S'09–M'11) received the B.Eng. (Hons.) and Ph.D. degrees in electrical and electronic engineering from Nanyang Technological University, Singapore, in 2005 and 2010 respectively.

He was a Research Engineer with the School of Electrical and Electronic Engineering, Nanyang Technological University, from 2008 to 2009. He joined the Institute for Infocomm Research, Agency for Science, Technology and Research (A*STAR), Singapore, in 2009 as a Research Fellow and then a Scientist I, and later transferred to the National Metrology Centre, A*STAR in 2011. From 2012 to 2014, he was part-timely seconded to Psiber Data Pte. Ltd., Singapore, where he was involved in metrological development and assurance of a handheld cable analyser, under a national Technology for Enterprise Capability Upgrading (T-Up) scheme of Singapore. Currently, he is appointed as a Scientist III with the National Metrology Centre, A*STAR. Concurrently, he also serves as a Technical Assessor for the Singapore Accreditation Council – Singapore Laboratory Accreditation Scheme (SAC-SINGLAS) in the field of RF and microwave metrology. His current research interests include electromagnetic metrology, electromagnetic measurements and standards, and electromagnetic-wave propagations.

Dr. Meng is a member of the IEEE Microwave Theory and Techniques Society. He is a recipient of the Asia Pacific Metrology Programme (APMP) Iizuka Young Metrologist Award in 2017 and the national T-Up Excellence Award in 2015.

Feng Yuan (S'15) received the B. Eng. (Hons.) degree in electrical and electronics engineering from the Nanyang Technological University, Singapore, in 2013. He is currently working toward the Ph.D. degree in the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore.

His research interests include the study of the effects of cloud on performance of satellite communication and the mitigation technique to counteract rain fades.

Jin Teong Ong received the B.Sc. (Eng.) degree from London University, London, U.K., the M.Sc. degree from University College, London, and the Ph.D. degree from Imperial College London, London.

He was with Cable & Wireless Worldwide PLCfrom 1971 to 1984. He was an Associate Professor with the School of Electrical and Electronic Engineering, Nanyang Technological Institute (now Nanyang Technological University), Singapore, from 1984 to 2005, and an Adjunct Associate Professor from 2005 to 2008. He was the Head of the Division of Electronic Engineering from 1985 to 1991. He is currently the Director of research and technology of C2N Pte. Ltd.—a company set up to provide consultancy services in wireless and broadcasting systems. His research and consultancy interests are in antenna and propagation—in system aspects of satellite, terrestrial, and free-space optical systems including the effects of rain and atmosphere; planning of broadcast services; intelligent transportation system; EMC/I; and frequency spectrum management.

Dr. Ong is a member of the Institution of Engineering and Technology.