A Simplified Model for the Retrieval of Precipitable Water Vapor from GPS Signal

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

***Abstract*— In this paper, a simplified latitude and day-of-year based model is proposed for the retrieval of Precipitable Water Vapor (PWV) from Global Positioning System (GPS) signal. Conventionally, PWV, the total amount of water in a vertical column of unit cross-section area, is estimated from the GPS signal delay and a dimensionless conversion factor *PI*. This *PI* value is found to rely on a water-vapor weighted mean temperature (T_m) value which varies widely across the day, month and year for different regions. It is therefore, both time and site specific. Analysis of the *PI* value and its effect on the retrieved PWV from the data obtained for tropical, sub-tropical and temperate regions show that although the *PI* value is time and site specific, the change in the median value of *PI* for different years is minimal and is dependent only on factors like the latitude coordinates of the particular site and the day of the year. Therefore, using the data obtained from 174 different sites, a latitude-coordinate and day-of-year based *PI* value model for the retrieval of PWV is proposed in this paper. The proposed model has been successfully validated using data from different databases: the IGS GPS NASA database, the IGS GPS GGOS database and the VLBI database. Results show a strong agreement between PWV values calculated using the proposed model and that calculated using the temperature dependent models with 99%, 98% and 93% of error within ±1 mm for IGS GPS NASA, IGS GPS GGOS and VLBI databases, respectively. Moreover, the proposed model allows for the ease of PWV retrieval, which is useful in meteorological studies and also applicable in satellite communications.**

Index Terms*—Global Positioning System (GPS), Precipitable Water Vapor (PWV), Radiosonde, T_m-T_s relation, Zenith Wet Delay (ZWD), GGOS, VLBI, *PI

I. Introduction

The water-vapor stored in the column of the atmosphere in terms of water budget is the Precipitable Water Vapor (PWV) [1]. The PWV is an important indicator of water vapor climatology and variability in the lower troposphere and related climate processes [2, 3]. This variable is strongly linked to the hydrological cycle and dynamical processes especially in regions where the overall PWV is high [4]. The PWV concentrations vary with high degree of spatial and temporal variability, depending upon the season, topography and other local and regional climatic conditions [5-7]. Thus, it is of importance to measure the PWV correctly and understand its variability.

Radiosondes, microwave-radiometers, satellite based instruments are a few conventional technologies that can be used to measure the PWV. Water vapor climatology is typically investigated using Radiosonde and satellite observations. Such observations have limitations in capturing good resolution diurnal variations as they have low spatial-temporal resolutions [8, 9]. Radiosondes are generally launched only twice a day.

To overcome the drawbacks of these systems, Global Positioning System (GPS) signal is extensively being used to retrieve the PWV values [10]. With the rapid deployment of GPS monitoring stations in local, regional, and global scales, ground-based GPS meteorology offers improved spatial and temporal resolutions for water vapor variations compared to traditional techniques [10, 11]. GPS is widely being used as an all-weather, low-cost remote sensing instrument for weather forecasting and climatology. Particularly in the field of remote sensing and weather forecasting, the GPS-derived PWV has been used for the analysis of severe weather conditions such as storms, floods [12-14], heavy rainfall events monitoring, rainfall now-casting, cloud microphysics and dynamics studies [15-17].

Estimation of the PWV values from GPS signal is mainly based on the delay incurred to the signals travelling from a GPS satellite to a receiver on the ground. This atmospheric delay of GPS signal is mostly caused by the ionosphere and troposphere of the earth. The total delay along the zenith path is called as Zenith Total Delay (ZTD), ΔL_T^o. The ZTD can be partitioned into two parts: Zenith Hydrostatic Delay (ZHD), ΔL_h^o and Zenith Wet Delay (ZWD), ΔL_w^o. ZHD mainly depends on the surface pressure (P_s) [18]; and the ZWD is a function of the atmospheric water vapor profile. ZWD is important for the retrieval of the PWV values from GPS signals. We use GPS Inferred Positioning SYstem-Orbit Analysis and Simulation Software (GIPSY-OASIS) [19-21] for the computation of ZWD in this paper.

Once ZWD is computed, PWV values, i.e., pwv (in mm, total amount of water in a vertical column of unit cross-section area) can be derived as [10, 22],

Manuscript received Mar 2017, revised 7 Jun 2017, accepted 30 Jun 2017. This work was supported in part by the Defence Science and Technology Agency, Singapore. S. Manandhar is financially supported by the Singapore International Graduate Award (SINGA) Scholarship.

S. Manandhar and Y. H. Lee are with the School of Electrical and Electronic Engineering, Nanyang Technological University, 50 Nanyang Avenue, Singapore 639798 (e-mail: shilpa005@e.ntu.edu.sg, eyhlee@ntu.edu.sg).

Y. S. Meng is with the National Metrology Centre, Agency for Science, Technology and Research (A*STAR), 1 Science Park Drive, Singapore 118221 (e-mail: ysmeng@ieee.org, meng_yusong@nmc.a-star.edu.sg).

J. T. Ong is with the C2N Pte. Ltd., Singapore 199098

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

where

$$PI = \frac{10^6}{\left(\frac{K_3}{T_m} + K_2'\right) R_v}. \quad (2)$$

Here, ρ_l is the density of liquid water (1000 kg/m^3). *PI* [23] is the dimensionless factor determined by eq. (2) where R_v is the specific gas constant for water vapor of 461.5181 J/kgK, K_3 and K_2' are the refractivity constants of (3.739 ± 0.012) ×10^3 K^2/P$_a$ and (22.1 ± 2.2) × 10^{-2} K/P$_a$, respectively [10]. T_m (in Kelvin) can be obtained from Radiosonde data as shown in eq. (3).

$$T_m = \frac{\int \frac{e}{T} dz}{\int \frac{e}{T^2} dz}. \quad (3)$$

where e is the water vapor pressure and T is the air temperature acquired from the Radiosonde sounding profiles.

Here, it is noted that the PWV values are derived by multiplying ZWD by a dimensionless conversion factor *PI*. This factor *PI* is calculated using water-vapor weighted mean temperature (T_m) as shown in eq. (2), which is calculated using Radiosonde data with eq. (3). Due to the low temporal resolution of Radiosonde data, T_m is generally predicted using surface temperature (T_s) data. The relation between T_m and T_s is found to be site-specific and varies from one region to another. Thus, a new (T_m - T_s) relationship has to be derived for different observation sites and this is often time consuming. Meanwhile, there is also a practice of using *PI* value as a constant of 0.15 as a rule of thumb, which often results in inaccuracy in results [24]. Therefore, it is of great interest for us to propose an alternative and efficient way for retrieving the PWV values from GPS signal accurately.

In the rest of this paper, Section II presents an overview of some existing ($T_m - T_s$) relationships and the relationships that we have studied using 4 years of data from 174 different stations consisting of tropical, sub-tropical and temperate regions. Section III proposes a new model and highlights the novelty and contribution of this work. Section IV presents comparison results to verify and validate the proposed model. Finally, Section V concludes this paper.

II. Mean Temperature (T_m) and Surface Temperature (T_s) Relationships

A. Review on Existing ($T_m - T_s$) Relationships

The most commonly used ($T_m - T_s$) relationship was reported by Bevis *et al*. [10] as,

$$T_m = a + b\, T_s\,, \quad (4)$$

where a = 70.2 and b = 0.72.

It is noted that eq. (4) was derived based on analysis of 8718 Radiosonde profiles spanning approximately a two-year period from sites within the United States with a latitude range of 27° to 65° and an altitude range of 0 to 1.6 km [10]. This equation has been used by many researchers working in this field [25-27], notwithstanding researchers from the tropics and Antarctica regions. But studies have shown that the accuracy of the Bevis relationship in eq. (4) is found to be latitude dependent. It tends to overestimate T_m by up to 5 K at mid and high latitudes and tends to underestimate T_m by up to 6 K at low latitudes [28]. Therefore, instead of using eq. (4), some researchers were aware that the coefficients a and b in eq. (4) are region and season specific, and thus proposed their own ($T_m - T_s$) relationships [29-32] as summarized in Table I.

TABLE I
($T_M - T_S$) RELATIONSHIP FOR SPECIFIC STATIONS FROM LITERATURE

Ref.	Coordinate	Region	a	b
[29]	45.95° N, 13.64° E	Temperate	39.94	0.83
[30]	30.53° N, 117.12° E	Sub-Tropical	44.05	0.81
[31]	4.21° N, 101.98° E	Tropical	182	0.35
[32]	1.35° N, 103.68° E	Tropical	182.56	0.3432

As shown in Table I, for the temperate city of Nova Gorica, Slovenia (45.95° N, 13.64° E) [29], the coefficients a and b are found to be 39.94 and 0.83; while for the sub-tropical city of Anqing, China (30.53° N, 117.12° E) [30], they are found to be 44.05 and 0.81. Moreover, for the tropical regions of Malaysia (4.21°N, 101.98° E) [31] and Singapore (1.35° N, 103.68° E) [32], the values of a and b are found to be close to each other, i.e., 182 and 0.35, and 182.5 and 0.3432, respectively. As can be seen from Table I, when latitude increases, the coefficient a decreases and b increases. It is also clear that the coefficients are distinctly different from those of Bevis equation.

B. Derivation of ($T_m - T_s$) Relationships

For our analysis, we have also derived the ($T_m - T_s$) relationships for different regions. The database used for this purpose is downloaded from the database of Wyoming University [33], which provides Radiosonde data for any chosen station two times a day, generally at 00:00 UTC and 12:00 UTC. The surface observation data (T_s) are obtained from METAR reports of Wyoming University. METAR, which is a format for reporting the weather information, is available online for different stations in similar manner as the Radiosonde data [33]. The METAR reports show that the surface observation is made for different stations at different

TABLE II
DERIVED $T_M - T_S$ RELATIONSHIPS FOR DIFFERENT REGIONS

Region	Number of stations	Latitude Range (absolute values)	a	b
Temperate	59	36.41°- 90°	67.12	0.73
Sub-Tropical	62	23.80-35.33°	106.36	0.60
Tropical	53	0.05°- 22.31°	129.13	0.52

Fig. 1. Location of 174 Radiosonde stations. Markers numbered from 1-59 (in red color) represent temperate stations, numbered from 60-121 (in green color) represent sub-tropical stations and those from 121-174 (in pink) represent the tropical stations. [Best viewed in color]

timings. The temporal resolution of MEATR data is generally 30 minutes.

For this paper, 4 years (2012-2015) of data from 174 different Radiosonde stations (59 stations of temperate region, 62 stations from sub-tropical region and 53 stations from tropical region) is used for model derivation and analysis purposes. Locations of 174 stations in the study are illustrated in Fig. 1.

Fig.2 presents the scatter plot of T_m vs T_s and the linear regression between the two for the temperate, sub-tropical and tropical regions. The relationship for all the regions are in the form of eq. (4). The estimated coefficients a and b with the total number of stations used for derivation and the latitude ranges are summarized in Table II.

From Table II, it can be seen that the gradient term, b, for temperate region (0.73) is higher and decreases in the sub-tropical (0.60) and the tropical (0.52) regions. This trend is consistent to those derived by other literature as shown in Table I. This indicates that the correlation between T_m and T_s is highest for temperate and lowest for tropical, and also verified by the estimated correlation coefficient values which are 0.90, 0.77 and 0.5 for temperate, sub-tropical and tropical regions respectively. This result can also be distinctly analyzed from Fig.2, which shows data from all region at once. Such variations amongst the 3 regions are mainly due to seasonal changes. For example, temperate regions have four distinct seasons, spring, summer, autumn and winter, and therefore, the surface temperature fluctuates over a large range from 230 K to 320 K. For the tropical region, temperatures remain high all the year round, and therefore, the surface temperature fluctuates over a much smaller range of 289 K to 310 K as compared to the temperate region.

From both Table I and Table II, it can clearly be noted that the relation between T_m and T_s is site-specific which varies from one region to another and the coefficients of only

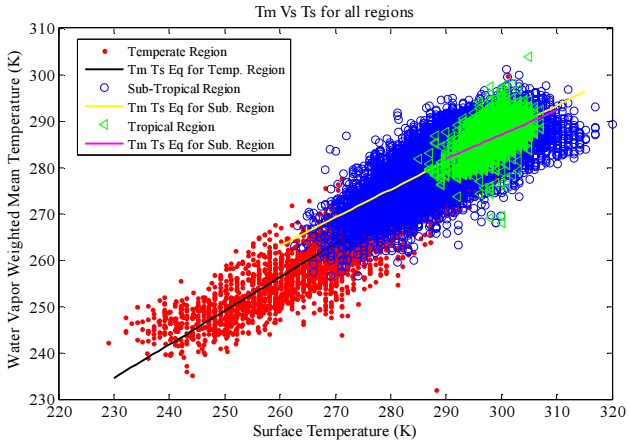

Fig. 2. T_m vs T_s for stations from temperate region (in red dots), from sub-tropical region (in blue circles) and from tropical region (in green triangles) with linearly regressed equations represented in black, yellow and magenta for the respective regions. [Best viewed in color]

temperate region matches well with that of the Bevis equation [10]. Thus, it can be concluded that the Bevis equation is only suitable for stations in the temperate region and a new (T_m - T_s) relationship needs to be derived for each station/region, which becomes complex and tedious especially when the data sources for T_m and T_s are not collocated.

Therefore, it is significant to develop an alternative methodology with good accuracy that can directly predict the PI values for different stations, instead of using long term site specific parameters like temperature which might be difficult to come by.

III. Modeling and Proposal of a New *PI* Function

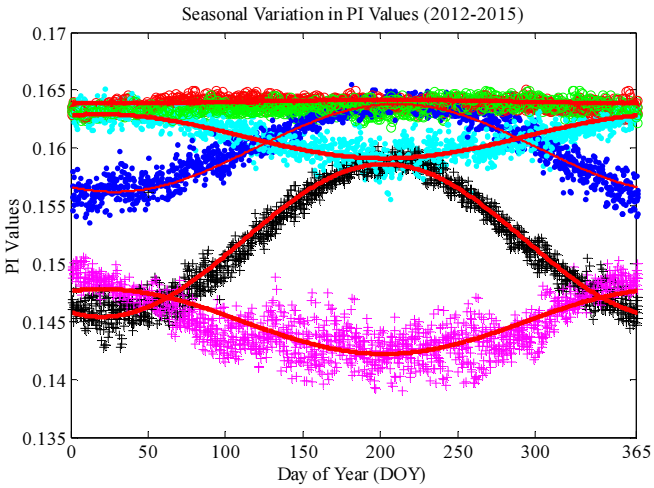

Fig 3. Plot of daily mean PI values against Day of Year (DoY), for 4 years of data from 2012 to 2015. The red and green circles are for North and South Tropical stations, blue and cyan dots are for North and South Sub-Tropical stations and black and magenta crosses are for North and South Temperate stations respectively. The solid red line represents the best fit line for respective regions. [Best viewed in color]

Following the discussion in Section II, an alternative method to predict the PI values is presented in this section.

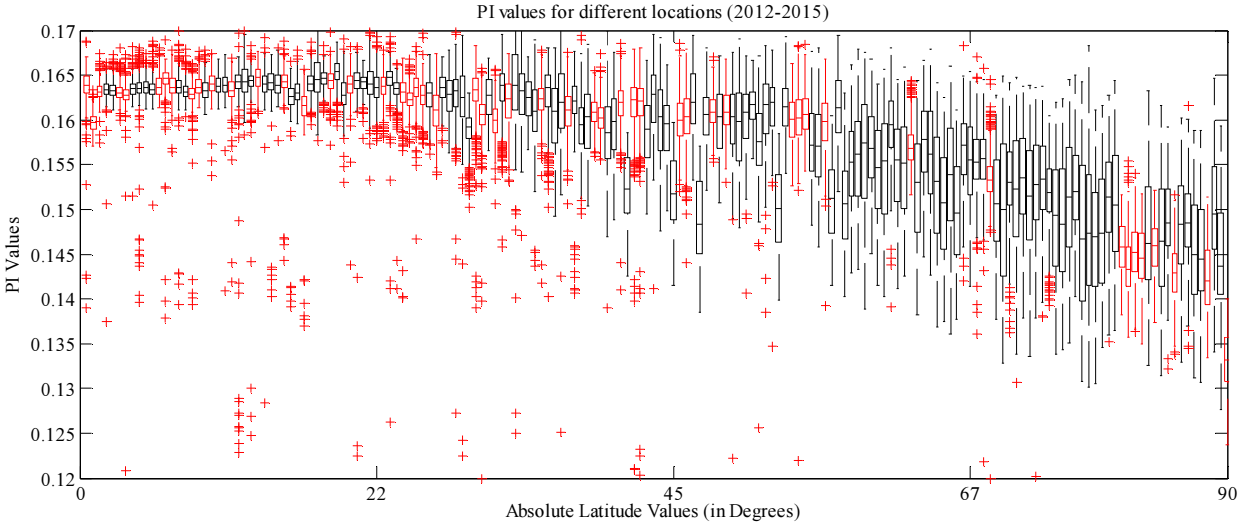

Fig 4. Boxplot of *PI* values for 174 Radiosonde stations for 4 years (2012-2015). Red color represents the stations from negative latitude and black from positive latitude. [Best viewed in color]

Four years (2012-2015) of *PI* values calculated from all the 174 Radiosonde stations (shown in Fig.1) using eq. (2-3) are studied against independent parameters like station coordinates, day of year (DoY), and station altitude. A temperature independent model is then proposed. Predicted *PI* values with the new proposed model will be used in eq. (1) to derive the PWV.

A. Trend of PI Values with Day of Year

Fig. 3 presents 4-years' daily averaged *PI* values plotted against the DoY for all 174 stations. The red and green circles are for northern (31 stations) and southern (22 stations) tropical stations, blue and cyan dots are for northern (42 stations) and southern (20 stations) sub-tropical stations and black and magenta crosses are for northern (49 stations) and southern (10 stations) temperate stations respectively. In general, Radiosonde stations are sparsely distributed in the southern hemisphere. And most of the southern area on earth is covered by sea. Thus the number of southern temperate stations compared to the northern hemisphere is small. Almost all the available stations from southern temperate regions have been included in the study. It can be noted from Fig. 1 that the chosen stations are well distributed.

From Fig. 3, it can be observed that the *PI* values for tropical stations are almost constant throughout all years. However, the *PI* values for other two regions show sinusoidal trends with the change of day number (DoY). The trend of change in *PI* values for stations from north and south alternates for the respective region due to the alternate seasons. The amplitude of the sinusoid is higher for stations from northern hemisphere compared to that for stations from the southern. This phenomenon is especially observed for stations from southern temperate region. As previously mentioned, the number of stations considered from southern temperate region is limited and the chosen stations are all near to the sea (from coastal areas). Thus the variation in *PI* values does not follow very distinct pattern as that of northern stations. The sinusoidal trend of *PI* values with DoY can be mathematically modelled with eq. (5). Eq. (5) follows the same pattern as the Global Pressure Temperature (GPT) model to calculate surface temperature and mean weighted temperature in [34]. Through the regression technique, the values of c and d in eq. (5) for different regions can be estimated and are tabulated in Table III, and the best fit line is plotted in solid red in Fig. 3.

TABLE III
VALUES OF c AND d FOR (5)

Region	c		d	
	North	South	North	South
Temperate	-0.0066	0.0028	0.152	0.145
Sub-Tropical	-0.0038	0.0019	0.160	0.161
Tropical	-0.00016	0.00010	0.164	0.163

$$PI = c\cos\left(\frac{DoY - 28}{365.25}2\pi\right) + d \tag{5}$$

From Table III, it is found that values of d decreases as we move from the tropical stations to the temperate stations for both northern and southern regions and the absolute value of parameter c increases from tropical stations to the temperate stations. Physically, parameter d, represents the long term annual statistics of *PI* values which can be derived based on the station location. Parameter c indicates the seasonal variations, which is the highest in the temperate region, low in the sub-tropical region and the least in the tropical region. This trend can be seen from c values in Table III. Note that the trend is more dominant in the northern hemisphere as compared to the southern. This parameter c is therefore location dependent taking into the consideration whether the station is from the northern or southern hemisphere.

B. Study of latitude dependency

1) Modelling Parameter d

Four years' of data is used to derive the long term annual median *PI* values to model the parameter d. Fig. 4 shows the boxplot of *PI* values for all the 174 stations. The red colored boxes are for stations from negative latitude and black for the

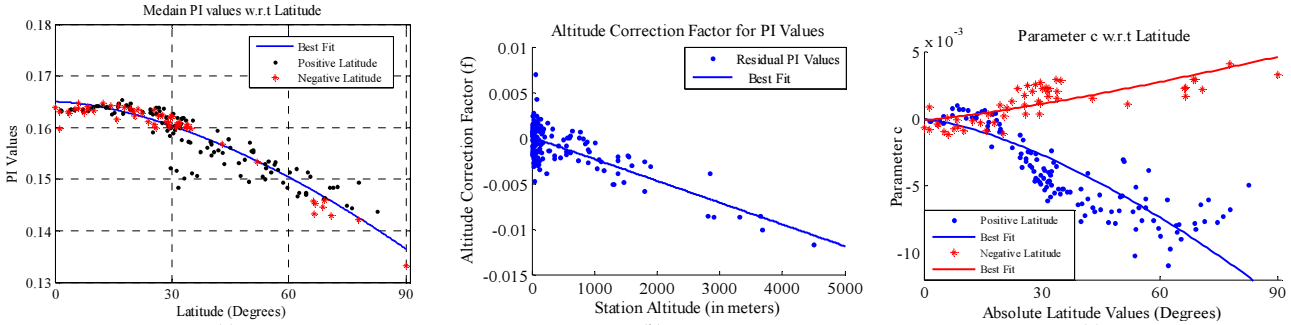

Fig. 5 (a) Study of median *PI* values with respect to absolute latitude values. (b) Altitude correction for the estimated *PI* values in (a). (c) Variation in parameter *c* with respect to the absolute latitude values. [Best viewed in color]

stations from positive latitude region. From Fig. 4, it is clearly observed that when the absolute value of latitude increases, the median *PI* values decrease whilst its variation increases. For stations from tropical region (with absolute latitude between 0.05˚ and 22.31˚), the median *PI* value is the highest at around 0.162 and its variation is also the smallest. A significant variation in *PI* values is observed as the stations move farther from the equator. Such differences can again be linked to the seasonal changes that occur in sub-tropical and temperate regions, which cause larger variation in temperature and hence effects are observed in the *PI* values.

Moreover, as can be inferred from Fig.4, the median *PI* values of a given station could be predicted using its latitude information. The median *PI* values is found to have a good and clear correlation with latitude of the station as shown in Fig. 5(a). The median *PI* values for stations from negative latitude are represented by red stars and from positive latitude are represented by black dots. The solid blue line is the best fit curve as shown in eq. (6). Given L_a, the latitude value of any given station, eq. (6) calculates the median *PI* value for that station.

$$d = 0.165 - \left(1.7 \cdot 10^{-5}\right) |L_a|^{1.65}. \quad (6)$$

While modeling *PI* values, it was observed that the station altitude (H) might also affect the accuracy of the modeled *PI* values. Therefore, the effects of station altitude, H, on the predicted *PI* values is also assessed by using a residual parameter which is the difference between predicted median *PI* value, d using eq. (6) and actual median *PI* value. Fig. 5(b) shows the residual plot. It clearly shows a linearly decreasing trend of residual with reference to the station altitudes. The solid blue line in Fig. 5(b) is the linearly regressed equation which gives the altitude correction factor defined as $f = -2.38 \cdot 10^{-6}\, H$. It is found that the change in *PI* values is independent of altitude when stations are from lower altitude region where H < 1000 m. The effect of H becomes significant when the stations are from higher altitude region where H > 1000 m. Therefore, it is advisable to use the altitude correction factor for stations from higher altitude region where H > 1000 m.

2) Modelling Parameter c

It is noted that the parameter c reflects the seasonal variations and specific values of c (as shown in Table III) show latitude and hemisphere dependency of *PI* values. The plot in Fig. 5(c) shows the relation of parameter c with respect to latitude. The red stars are for stations from negative latitude and the blue dots are for stations from positive latitude. The solid red and blue lines represent the best fit curve for parameter c as shown by eq. (7). Fig. 5(c) and typical values of parameter c from Table III, show that the parameter c also captures the phenomenon of higher *PI* values for southern stations compared to the northern stations with alternating signs.

$$c = -1 \cdot sgn(L_a) \cdot 1.7 \cdot 10^{-5} |L_a|^{h_{fac}} - 0.0001 \quad (7)$$

In eq. (7), a new factor, named as h_{fac}, is introduced to account for the differences between northern and southern hemispheres. The regression results show that h_{fac} is 1.25 for stations from southern hemisphere and 1.48 for stations from northern hemisphere. A term $sgn(L_a)$ denotes whether the station is from the northern hemisphere or the southern hemisphere. $sgn(L_a)$ is 1 for stations from the northern hemisphere and -1 for stations from the southern hemisphere.

C. Proposal of A simplified PI model

Therefore, to calculate the *PI* values for a given station at a specific day, we need the station latitude information in degrees, L_a, altitude of the station in meters, H, day of the year for which the *PI* values are to be calculated, DoY. Depending on whether L_a values are positive (from north) or negative (from south), h_{fac} can be selected accordingly. h_{fac} for the northern hemisphere is 1.48 with $sgn(L_a) = 1$ and for the southern hemisphere is 1.25 with $sgn(L_a) = -1$. Using these information, the proposed *PI* model is given in eq. (8).

$$\begin{aligned} PI = {} & \left[-1 \cdot sgn(L_a) \cdot 1.7 \cdot 10^{-5} |L_a|^{h_{fac}} - 0.0001\right] \\ & \times \cos\left(\frac{DoY - 28}{365.25} 2\pi\right) \\ & + \left[0.165 - \left(1.7 \cdot 10^{-5}\right) |L_a|^{1.65}\right] \\ & + f \end{aligned} \quad (8)$$

Where, f is negligible for stations from low altitude (H <1000 m) region and $f = -2.38 \cdot 10^{-6}\, H$ for stations from high altitude (H >1000 m) region.

IV. Performance Comparison and Validation

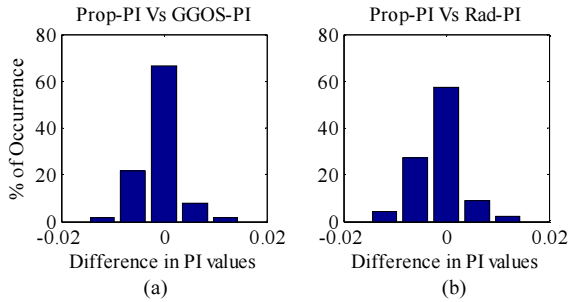

Fig. 6. (a) Histogram of the difference between Prop-PI and GGOS-PI (b) Histogram of difference between Prop-PI and Rad-PI. With bin size of 0.006 *PI* values. (4 years of data from 174 stations (Fig. 1))

A. Performance Comparison

For assessing the performance of our proposed temperature independent model, *PI* values calculated from our proposed method in eq. (8) are compared with two different temperature dependent methods: 1) *PI* values calculated using T_m derived from Radiosonde data and 2) *PI* values calculated using T_m database published by GGOS Atmosphere. GGOS database uses data from numerical weather models from European Centre for Medium-Range Weather Forecasts by operational analysis (ECMWF) [35]. Four years of *PI* values calculated for all 174 locations (shown in Fig. 1) using both Radiosonde and GGOS data are used for this purpose. In the following, *PI* values calculated using eq. (8) is named as Prop-PI, *PI* values from Radiosonde data (using eq. (2-3)) is named as Rad-PI, and *PI* values calculated from GGOS T_m database (using eq. (2)) is named as GGOS-PI.

The relative error in PWV, from eq. (1), due to error in *PI* values can be calculated using eq. (9) [28].

$$\frac{\Delta\, PWV}{PWV} = \frac{\Delta\, PI}{PI} \tag{9}$$

On the basis of eq. (9), for *PI* values ranging from 0.12 – 0.17 (Fig. 3), the 1% and 2% accuracies in PWV requires errors in *PI* values of less than 0.0014 and 0.003 on average, respectively. Fig 6(a) and 6(b) show the histograms of difference between Prop-PI and GGOS-PI and difference between Prop-PI and Rad-PI respectively. Almost 67 % of differences between Prop-PI and GGOS-PI and 58 % of differences between Prop-PI and Rad-PI are within ±0.003 *PI* values. Therefore, the results suggest that the predicted *PI* values from the proposed method can be used to predict PWV values with high accuracy as compared to the existing methodologies. This is also reflected in the next section.

B. Further Validation of Proposed Model

The proposed model is further validated using different databases; the IGS GPS GGOS database, IGS GPS NASA database, Very-Long-Baseline Interferometry (VLBI) wet delay data, and the Radiosonde data.

1) IGS GPS GGOS Database

Two years (2013-2014) of ZWD values and corresponding T_m values for 384 IGS stations are downloaded from GGOS Atmosphere website [35]. The PWV values are then derived for each IGS site using T_m and GGOS ZWD (eq. (1-2)) and is named as PWV-GGOS-Tm. The IGS site location information (altitude and latitude) and DoY is used in eq. (8) for the proposed model and PWV values are calculated in eq. (1) using predicted *PI* values and GGOS ZWD, named as PWV-GGOS-Prop.

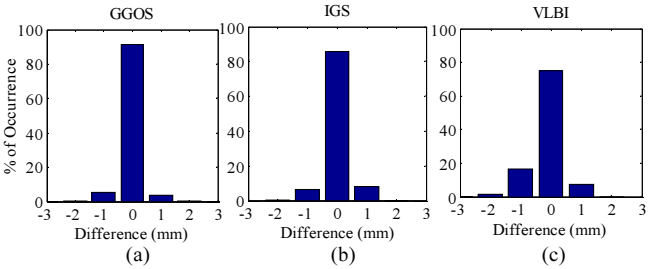

Fig. 7. (a). Histogram of difference between PWV-GGOS-Prop and PWV-GGOS-Tm. (2 years of data from 384 IGS stations) (b) Histogram of difference between PWV-IGS-Prop and PWV-IGS-Tm. (2 years of data from 24 IGS stations). (c) Histogram of difference between PWV-VLBI-Prop and PWV-VLBI-Tm. (2 years of data from 15 VLBI stations).With bin size of 1 mm.

Fig. 7(a) shows the histogram plot of difference between PWV-GGOS-Prop and PWV-GGOS-Tm. It is calculated that almost 98% of difference is within ±1 mm, and the root mean square, RMS of difference is around 0.31 mm. The RMS values suggest that the PWV values calculated using the proposed method works with similar accuracy as compared to the existing temperature dependent methodology [24].

2) IGS GPS NASA Database

Two years (2013-2014) of RINEX files are downloaded from NASA's website [37] for 24 IGS stations (8 from temperate, 8 from sub-tropical and 8 from tropical regions) and ZWD is processed using GIPSY OASIS software with the recommended scripts. All the chosen stations are collocated to the corresponding Radiosonde stations used in the derivation of the model. For each IGS station, corresponding (T_m - T_s) relationship is selected from Table II to predict the T_m values. The predicted T_m values are then used in eq. (1) and eq. (2) to derive the PWV values which is named as PWV-IGS-Tm. Each station location information (altitude and latitude) and DoY is then used in the proposed model, eq. (8), to get the corresponding *PI* values. Finally the predicted *PI* values and calculated ZWD values are used in eq. (1) and PWV values are calculated, named as PWV-IGS-Prop.

Fig. 7(b) shows the histogram plot of difference between PWV-IGS-Prop and PWV-IGS-Tm. The distribution of histogram is similar to Fig. 7(a) with almost 99% of the differences within ±1 mm with an RMS value of 0.33 mm. This result again strongly suggest that the proposed model can calculate PWV values as accurately as using the traditional T_m - T_s relationship approach [24].

3) VLBI Database

The proposed model is further validated using the VLBI database. The CDDIS archive of VLBI products consists of solutions derived by IVS analysis centers from the analysis of VLBI experiment files [38]. These products include the troposphere parameters like ZTD and ZWD. These tropospheric parameters with meteorologically important

surface parameters such as temperature, pressure and humidity, is downloaded from IVS data center [39]. In this analysis, 2 years (2013-2014) of ZWD values and corresponding surface temperature values for 15 VLBI stations (Algo, Bada, Fort, Hart, Hoba, Kath, koke, Mate, Nyal, Onsa, TSKB, West, Wett, Yebe, Zele) are used.

For each VLBI station, corresponding (T_m - T_s) relationship is selected from Table II and T_m values are then predicted. The predicted T_m values and VLBI wet delay values are then used in eq. (1) and eq. (2) to calculate the PWV values which is named as PWV-VLBI-Tm. Each station location information (altitude and latitude) and DoY is then used in the propose model, eq. (8), to get the corresponding PI values, which are used in eq. (1) with VLBI ZWD to calculate the PWV values, and named as PWV-VLBI-Prop.

Fig. 7(c) shows the histogram plot of difference between PWV-VLBI-Prop and PWV-VLBI-Tm. The distribution shows that a large portion (almost 93%) of the differences is very small (within ±1 mm) with an RMS value of 0.38 mm. Therefore, similar conclusions as presented in earlier subsection can be drawn. Through validations using different datasets, it is shown that the proposed model works well with high accuracy.

4) Independent Station Validation

To have a comprehensive comparison, a GPS station from Japan (ID: TSKB), which is collocated with the VLBI station (ID: TS) and Radiosonde station (No.: 47646) is used for independent validation purpose. The Radiosonde data from this station was deliberately left out (not one of the 174 stations) during the derivation of the proposed equation.

Based on the good data availability from all databases, two years (2011 and 2012) of ZWD data are downloaded from the respective sites of GPS station (using NASA database and GGOS database) and VLBI station. Corresponding PWV values are then calculated using the proposed model. PWV values are also calculated for the co-located Radiosonde station. The calculation procedure of the PWV value from Radiosonde data is given in [40]. PWV calculated for GPS station; using NASA database is named as IGS PWV; using GGOS database is named as GGOS PWV; using VLBI database is named as VLBI PWV and using Radiosonde database is named as Radiosonde PWV.

Fig. 8 shows the scatter plot of PWV values from different databases against DoY. The green circles represent the IGS PWV, black dots represent the GGOS PWV, red stars represent the VLBI PWV and blue triangles represent the Radiosonde PWV. From Fig. 8, it is observed that the PWV values calculated using GPS data and VLBI data which uses the proposed model follows the same pattern over different years and matches well with the Radiosonde PWV. Although the pattern is the same, absolute values of PWV from GPS, VLBI and Radiosonde have some differences since different instruments were used for the measurement of the delays and for calculation of the PWV. Furthermore, all these stations are not exactly at the same location. There is approximately a 10 km distance between the GPS and Radiosonde stations.

Moreover, it is also observed that the PWV from VLBI and the PWV from GPS have the same pattern but the PWV from VLBI is generally less than the PWV from GPS.

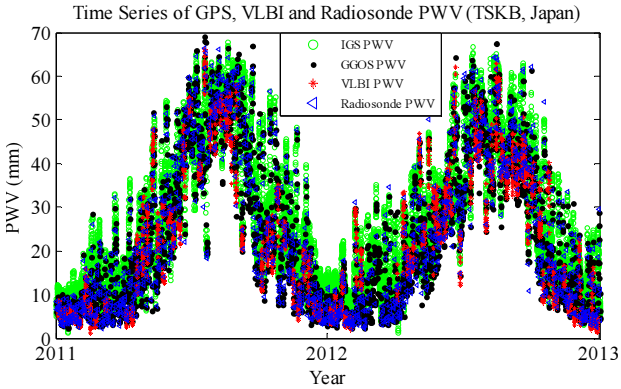

Fig. 8. Plot of PWV Vs DoY for different years. Green circle represent PWV from GPS station using IGS database, black dots represent PWV from GPS station calculated from GGOS database, red stars represent PWV from VLBI data and blue triangles represent PWV from Radiosonde data. 2 years of data from independent station of Japan, TSKB (36.06N, 140.05E). [Best viewed in color]

Similar observations have been reported in the literature [38]. Small differences can be observed between PWV calculated for the same GPS station but using different sources (from NASA and from GGOS). This is due to the error induced by different parameters considered during the processing of the wet delay by using different processing softwares.

Despite the small differences which can be accounted for in the above, all 4 PWV matches very well. Therefore, it can be concluded that, by using an independent station for performance evaluation, using three independent databases, GPS, VLBI and Radiosonde, the proposed model is found to be both accurate and reliable.

V. Conclusion

In this paper, a simplified, latitude and day-of-year based PI value model is proposed for the retrieval of PWV values from GPS signals. The PI value is conventionally determined using the water-vapor weighted mean temperature T_m which is found to have a linear relationship with the surface temperature T_s. This method is site-specific and its' use is limited by the poor temporal and spatial resolution of Radiosonde data.

In the analysis of different regions: temperate, sub-tropical and tropical using data of four years from 174 different stations, it was concluded that the median PI value varies with latitude and variations about the median PI value are the result of seasonal changes that a particular station goes through. Stations from the tropics have very little variation whereas stations from the temperate has the highest variation. PI value residuals were studied with respect to station altitude and was concluded that the altitude of a station plays a significant role when the stations are from higher altitude compared to lower altitude. For simplified model it is recommended that the altitude factor should only be used when the altitude of a given station is higher than 1000 m. Thus, by examining the relationship of PI with respect to its geographic coordinates and considering the seasonal effects, a latitude and DoY based PI model was proposed.

The accuracy of the proposed PI model for estimating the GPS-based PWV is compared against the PI values calculated using T_m data from Radiosonde and from GGOS

database using 4 years of data from 174 stations. Results show that the *PI* values estimated using the proposed model have similar accuracy to the temperature dependent model. PWV values were calculated using temperature dependent models and compared against the proposed model using 3 databases; GPS, VLBI and Radiosonde. It is shown that the error in PWV is small and lies within ±1 mm. This accuracy is high enough for different geodetic applications [24].

The proposed model has the advantage of being computationally efficient and simple to use without requiring high temporal and spatial resolution resources, and thus could be applied universally.

ACKNOWLEDGMENT

The authors would like to thank the anonymous editor and reviewers for their constructive comments and suggestions to improve this paper, and would also like to thank our team member Mr. Yuan Feng for his valuable discussion.

REFERENCES

[1] B. Fontaine, P. Roucou, and S. Trzaska, "Atmospheric water cycle and moisture fluxes in the West African monsoon: Mean annual cycles and relationship using NCEP/NCAR reanalysis," *Geophys. Res. Lett*., vol. 30, no. 3, Feb. 2003.

[2] J. Wang, L. Zhang, A. Dai, T. V. Hove and J.V. Baelen, "A near-global, 2-hourly data set of atmospheric Precipitable water from ground-based GPS measurements," *J. Geophys. Res*., vol. 112, no. D11, Jun. 2007, doi: 10.1029/2006JD007529.

[3] 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.

[4] 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.

[5] S. G. Jin., Z. Li, and J. Cho, "Integrated water vapor field and multi-scale variations over China from GPS measurements," *J. Appl. Meteorol. Clim*., vol. 47, no. 11, pp. 3008-3015, Mar. 2008.

[6] S. G. Jin., O.F. Luo, and S. Gleason, "Characterization of diurnal cycles in ZTD from a decade of global GPS observations," *J. Geodesy*, vol. 83, no. 6, pp. 537-545, 2009.

[7] 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.

[8] D. J. Gaffen, W. P. Elliott, and A. Robock, "Relationships between tropospheric water vapor and surface temperature as observed by radiosondes," *Geophys. Res. Lett*., vol. 19, no. 18, pp. 1839–1879, 1992.

[9] W. R. Elliott, R. J. Ross, and D. J. Gaffen, "Water vapor trends over North America," *in Proc. 6th Symp. Global Change Studies*, Dallas, TX, pp. 185–186, 1995.

[10] 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.

[11] 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.

[12] K. Zhang, T. Manning, S. Wu, W. Rohm, D. Silcock, and S. Choy, "Capturing the signature of severe weather events in Australia using GPS measurements," *IEEE Trans. Geosci. Remote Sens*., vol. 8, no. 4, pp. 1839–1847, Apr. 2015.

[13] 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.

[14] S. Choy, C. Wang, K. Zhang, and Y. Kuleshov, "GPS sensing of precipitable water vapour during the March 2010 Melbourne storm," *Adv. Space Res*., vol. 52, no. 9, pp. 1688–1699, Nov. 2013.

[15] P. Benevide, J. Catalao, and P.M.A Miranda, "On the inclusion of GPS precipitable water vapour in the nowcasting of rainfall", *Nat. Hazards Earth Syst. Sci*., vol. 15, no. 12, pp. 2605–2616, 2015.

[16] 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. 1918–1924, Jun. 2015.

[17] Z. Wang, J. French, G. Vali, P. Wechsler, et al., "Single aircraft integration of remote sensing and in situ sampling for the study of cloud microphysics and dynamics," *Bull. Amer. Meteor. Soc*., vol. 93, no. 5, pp., 653–668, 2012.

[18] G. Elgered, J. L. Davis, T. A. Herring, and I. I Shapiro, "Geodesy by radio interferometry: water vapor radiometry for estimation of the wet delay," *J. Geophys. Res*., vol. 96, no. B4, pp. 6541–6555, Apr. 1991.

[19] T. Gregorius, "GIPSY-OASIS II: How it works.?", California Institute of Technology, United States of America, 1996.

[20] S. Desai, W. Bertiger, "GIPSY/OASIS (GIPSY) Overview and Under the Hood", *Near Earth Tracking Systems and Applications Groups*, JPL, California Institute of Technology, March 3, 2014.

[21] S. Desai, W. Bertiger, "Static and Kinematic Precise Point Positioning with gd2p.pl", *Near Earth Tracking Systems and Applications Groups*, JPL, California Institute of Technology, March 4, 2014.

[22] O. T. Davies, and P. A. Watson, "GPS phase-delay measurement: technique for the calibration and analysis in millimetre-wave radio propagation studies," *IEE Proc. Microw. Antennas Propag*., vol. 146, no. 6, pp. 369–373, 1999.

[23] Y. Shoji, H. Nakamura, K. Aonashi, A. Ichiki, H. Seko, "Semi-diurnal and diurnal variation of errors in GPS precipitable water vapor at Tsukuba, Japan caused by site displacement due to ocean tidal loading", *Earth Planets Space*, vol. 52, pp. 685-690, 2000.

[24] F. Alshawaff, T. Fuhrmann , A. Knöpfler, X. Luo, M. Mayer, S. Hinz, B. Heck, "Accurate estimation of atmospheric water vapor using GNSS observations and surface meteorological data," *IEEE Trans. on Geosci. Remote Sens.*, vol. 53, no. 7, pp. Jul. 2015.

[25] X. Li, L. Xu, Y.Fang, Y. Zhang, J. Ding, H. Liu, and X. Deng, "Estimation of the precipitable water vapor from ground-based GPS with GAMIT/GLOBK," *in Proc. 2010 Sec. IITA Int. Conf. Geosci. Remote Sens,* China, pp. 210–214, Aug. 2010.

[26] W. Suparta, M. Alauddin, M. Ali, B. Yatim, N. Misran, "Remote sensing of antarctic atmospheric water vapour using ground-based GPS meteorology," *in Proc. 5th Student Conf. Research and Development (SCOReD)*, Malaysia, pp. 1-8, Dec. 2007.

[27] L. S. Kumar, F. Yuan, and Y. H. Lee, "Comparison of precipitable water vapor derived from GPS and radiosonde data for Singapore*," in Proc. 2014 Int. Symp Antennas Propag*, Taiwan, pp. 437-438, Dec. 2014.

[28] J. Wang, L. Zhang, and A. Dai, "Global estimates of water-vapor-weighted mean temperature of the atmosphere for GPS applications," *J. Geophys. Res*, vol. 110, no. D21, Nov. 2005.

[29] A. Sŭsnik, "GPS for Weather and SpaceWeather Studies," Ph. D thesis, University of Nova Gorica, 2013.

[30] Y. Cao, F. Zheng, Y. Xie, Y. Bi, "Impact of the weighted mean temperature on the estimation of GPS precipitable water vapor," *in Proc. 2008 Int. Conf. Microw. Millimeter Wave Technol*., China, pp.799–801, Apr. 2008.

[31] S. Amir and T.A.Musa, "GPS Meteorology Activities in the Malaysian Peninsula,"*in Proc. of 10th International Symposium & Exhibition on Geoinformation (ISG2011) & ISPRS Commission*., Malaysia, 27-29 Sept, 2011.

[32] S. Manandhar, Y.H. Lee, and S. Winkler, "Mean weighted temperature for PWV estimation," *in Proc. 2015 IEEE 4th Asia-Pacific Conf. Antennas Propag*, Bali, Indonesia, pp. 437-438, Jun.-Jul. 2015.

[33] Wyoming University, Department of Atmospheric Sciences, available at: http://weather.uwyo.edu/upperair/sounding.html (Sounding data) and http://weather.uwyo.edu/surface/meteorogram/ (METAR Report).

[34] Y. Yao, S. Zgu, S. Yue, "A globally applicable, season-specific model for estimating the weighted mean temperature of the atmosphere", *J. Geod.*, vol. 86, pp. 1125-1135, Apr. 2012.

[35] http://ggosatm.hg.tuwien.ac.at/DELAY/ETC/TMEAN/

[36] http://ggosatm.hg.tuwien.ac.at/DELAY/SITE/GPS/

[37] NASA's Earth Science Data Systems, Crustal Dynamics Data Information System , available at: ftp://cddis.gsfc.nasa.gov/pub/gps/data/

[38] S. G. Jin., O.F. Luo, and J. Cho, "Systematic Errors between VLBI and GPS Precipitable water vapor estimations from 5 year co-located

measurements," *J. Atm. Sol. Ter. Phy.*, vol. 71, no. 2, pp. 264-272, 2009.
[39] ftp://cddis.gsfc.nasa.gov/vlbi/ivsproducts/trop/
[40] Y. Liu and Y. Chen, "Precision of precipitable water vapour from radiosonde data for GPS solutions," *Geomatica*, vol. 54, no. 2, pp. 171–175, 2000.

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 was a recipient of the national T-Up Excellence Award in 2015.

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.