Deriving the Mean Tropospheric Temperature Model using AIRS and AMSU for GNSS Precipitable Water Vapour Estimation

Rata Suwantong*, Chalermchon Satirapod†, Panu Srestasathiern* and Chaiyaporn Kitpracha†

*Geo-informatics and Space Technology Development Agency (GISTDA)
†Department of Survey Engineering, Faculty of Engineering, Chulalongkorn University
Email: rata@gistda.or.th, chalermchon.s@chula.ac.th, panu@gistda.or.th, chaiyaporn.k@chula.ac.th

Abstract—Accurate estimation of the Precipitable Water Vapour (PWV) is hence crucial for weather forecast and climate studies. The Global Navigation Satellite System (GNSS) data is generally used for continuous, accurate, all-weather and real-time PWV. For this, one has to compute the mean tropospheric temperature which is generally modelled as a linear function depending on the air surface temperature. The parameters for a mean tropospheric temperature model are usually derived using radiosonde data which is expensive and covers only a small area. In this work, we propose to use temperature and water vapour profiles from satellite data to derive parameters of the mean tropospheric temperature models using the Extended Kalman Filter (EKF) supposing that the air surface temperature is periodic. The satellite data is obtained from the Atmospheric Infrared Sounder (AIRS) and the Advanced Microwave Sounding Unit (AMSU) board on NASA’s Aqua satellite. $T_r$ is supposed to be a triply-modulated sine function due to its periodic pattern and $T_m$ is supposed to be linearly dependent on $T_r$. The Extended Kalman Filter (EKF) is used to estimate the parameters of the local $T_m$ model using $T_m$ and $T_r$ from AIRS and AMSU as measurements.

Our study area is Bangkok, Thailand, where AIRS and AMSU data between November 2013 and November 2015 are used for the derivation of the local $T_m$ model. The GNSS data from a CORS own by the Department of Public Works and Town & Country Planning of Thailand (DPT) in 2015 are used for the GNSS PWV derivation. The accuracy of the GNSS PWV using the global and the local $T_m$ models will be compared using the PWV from AIRS and AMSU, which have been shown to provide high accuracy compared to PWV from Microwave Radiometer in the literature.

This paper is organized as follows. In section II, the interests of precipitable water vapour estimation will be elaborated. In section III, the methods for PWV estimation using GNSS observations will be discussed. In section IV, a method for deriving a local $T_m$ relation using AIRS and AMSU satellite data. The estimation results of the parameters for the local $T_m$ relation and the accuracy of the GNSS PWV derived using the global and local models will be studied in section V. The paper ends with concluding remarks.

I. INTRODUCTION

Precipitable Water Vapour (PWV) is an important factor for weather forecast and climate studies. It is currently monitored in many parts of the globe using GNSS observations from Continuously Operating Reference Stations (CORS) network, due to its high coverage both spatially and temporally. The PWV is computed from the Zenith Total Delay (ZTD) of the GNSS observations obtained using the Precise Point Positioning (PPP) technique. A model on the water vapour weighted mean tropospheric temperature $T_m$ in function of the air surface temperature $T_r$ which is measured at or near the GNSS receiver, must be used for the PWV computation. A global $T_m$ model is chosen in many existing works. However, it has been shown that a local $T_m$ model could provide GNSS PWV with better accuracy. Local models are generally derived using radiosonde data in the literature but sending radiosondes is expensive and cannot be done frequently for some areas as near the poles, in the ocean or in developing countries.

This paper therefore proposes to derive a local $T_m$ model using temperature and humidity profiles from satellite data which are from the Atmospheric Infrared Sounder (AIRS) and the Advanced Microwave Sounding Unit (AMSU) board on NASA’s Aqua satellite. $T_r$ is supposed to be a triply-modulated sine function due to its periodic pattern and $T_m$ is supposed to be linearly dependent on $T_r$. The Extended Kalman Filter (EKF) is used to estimate the parameters of the local $T_m$ model using $T_m$ and $T_r$ from AIRS and AMSU as measurements.

II. INTERESTS OF PRECIPITABLE WATER VAPOUR ESTIMATION

Precipitable Water Vapour (PWV) is the height of the amount of liquid water that would be produced if all the water vapour in an atmospheric column of a unit area, i.e. of 1 $m^2$, was condensed. The amount of PWV is generally close to 50 $mm$ at the equator and less than 5 $mm$ near the poles (Brutsaert, 2005). Atmospheric water vapour provides the mean by which moisture and latent heat are transported (Awange, 2012). Hence, it plays a significant role in weather forecasting and climate studies at local, regional and global scales. Errors in PWV estimation can be the major source of errors in a short-term (0-24h) precipitation forecast (Bevis et al., 1992). It is also proposed to use PWV as an additional input for Numerical Weather Prediction (NWP) models for better accuracy (Baker et al., 2001) and for flash flood prediction (Awange and Fukuda, 2003). It has also been shown that a high level of PWV occurs before extreme rainfall events.
Concerning the climate studies, water vapour is the largest greenhouse gas contributing to global warming, and therefore must be used for climate change monitoring. Water vapour also affects the amount and type of cloud clover and therefore its variation should be used for a robust solar power forecast system (Freedman et al., 2012).

Traditionally, radiosondes and ground-based Microwave Radiometers (MWR) are used for PWV estimation. However, they are expensive and could provide only a sparse spatial coverage of PWV data at low temporal frequency (Satriapod et al., 2010). To achieve better coverage both temporally and spatially at smaller cost, the Global Navigation Satellite System (GNSS) data from Continuously Operating Reference Stations (CORS) have been intensively used (Bevis et al., 1992; Sapucci, 2014; Alshawaf et al., 2015; Wilgan et al., 2015). In the US, a university-based GNSS network, called SuomiNet, provides continuous, accurate, all-weather, real-time GNSS-derived PWV data used for research in weather forecasts, severe weather, cloud dynamics, regional climate and hydrology (Ware et al., 2000).

III. PWV Estimation using GNSS Observations

A. Precipitable Water Vapour Computation from the ZTD

The Zenith Total Delay (ZTD) obtained using Precise Positioning Point (PPP) techniques can be decomposed into two parts: the zenith hydrostatic delay (ZHD) and the zenith wet delay (ZWD). The ZHD is contributed by dry air and the nondipole component of water vapour (Bevis et al., 1992). It can be computed if the air surface pressure (P_s) at the GNSS receiver is available using

\[
ZHD = \frac{2.2768P_s}{1 - 0.00266\cos(2\phi) - 0.00000028H} \tag{1}
\]

where \(P_s\) is the air surface pressure in bar, \(\phi\) is the latitude of the station and \(H\) is the height above the ellipsoid in m.

The ZWD is contributed only by the dipole component of the atmospheric water vapour, hence, can be used to derive information concerning the atmospheric water vapour. It is computed using

\[
ZWD = ZTD - ZHD \tag{2}
\]

Then, the Precipitable Water Vapour (PWV) is computed using

\[
PWV = \frac{10^{-6}}{R_w} \cdot \frac{1}{k_2 + k_3/T_m} \cdot ZWD \tag{3}
\]

where every quantity is described in SI units, i.e. PWV in m, \(k_2 = 0.221\ K/\text{Pa},\ k_3 = 3739\ K/\text{Pa}^2,\ R_w = 461.525\ J/(\text{kg} \cdot \text{K})\) is the specific gas constant for water vapour and \(T_m\) is the water vapour-weighted mean tropospheric temperature, shortly called the mean tropospheric temperature in K.

B. Existing Tropospheric Mean Temperature Models

The Tropospheric Mean Temperature (\(T_m\)) varies with location, height, season and weather. (Bevis et al., 1992) empirically derived

\[
T_m = 0.72T_s + 70.2 \tag{4}
\]

for the region between Alaska and Florida but (Schueler et al., 2001) showed that it can be applied globally. The unit of \(T_m\) and \(T_s\) in all the formulations in this section is in K.

(Mendes et al., 2000) proposed another formula based on a linear regression using around 32500 radiosonde profiles from 50 dispersed stations covering a latitude range of 62°S – 83°N for the year 1992 which is

\[
T_m = 0.789T_s + 50.4 \tag{5}
\]

(Schueler et al., 2001) used global \(T_m\) and \(T_s\) from the National Center of Environmental Prediction (NCEP/NOAA) from July 1999 up to July 2001 validated using profiles from 6 radiosonde stations in Germany from February to July 2001 to derive a linear global relation

\[
T_m = 0.647T_s + 86.9 \tag{6}
\]

and a harmonic relation

\[
T_m = \bar{T}_m + \tilde{T}_{m\cos} \left(\frac{2\pi DoY - DoY_w}{365.25}\right) + q_T T_s \tag{7}
\]

where \(T_m\) is the average of \(T_m\), \(\bar{T}_m\) is the amplitude of the annual cycle of \(T_m\), DoY is the day of year, DoY_w is the day of maximum winter, equal to 28 for the Northern Hemisphere and to 211 for the Southern Hemisphere. \(q_T\) is an amplifier-weighting term for the \(T_s\), \(\bar{T}_m\), \(\tilde{T}_{m\cos}\) and \(q_T\) are specific for each station.

Despite the availability of these global \(T_m\) relations, it has been shown in (Sapucci, 2014) that the GNSS derived PWV are more accurate when a local \(T_m\) relation is used. Moreover, it is important to accurately estimate \(T_m\) seeing that it is the largest source of errors in PWV estimation (Awange, 2012). In the literature, the coefficients of each meteorological parameter in the \(T_m\) relations are generally derived using radiosonde data. However, sending a radiosonde is expensive and cannot cover a large area.

To cope with this issue, it is proposed in this paper to use satellite data from the Atmospheric Infrared Sounder (AIRS) and the Advanced Microwave Sounding Unit (AMSU), board on Aqua satellite to derive a local \(T_m\) model. This method should be interesting for an area where radiosonde data are not frequently available as near the poles, in the ocean, or in developing countries.

IV. Mean Tropospheric Temperature Retreival from Satellite Data

A. AIRS and AMSU Data

The satellite Aqua has a global coverage with a frequency of 2 times a day which enables the use of AIRS and AMSU data for any region on Earth. The data used here is the combined AIRS and AMSU Level 2 version 6 Standard product, called...
AIRX2RET, which can be downloaded for free from (NASA 2013). The files are in hdf format which can be read in MATLAB using the hdread function.

The data when Aqua is above the region of interest limited by the minimum and the maximum latitude and longitude will be retrieved. Our area of interest here is Bangkok and the AIRX2RET data from November 2013 to November 2015 will be used. Aqua passes above Bangkok at around 02 and 13 local time (LT). Two local \( T_m \) models, one for day time and one for night time separately.

AIRS and AMSU collect temperature profile at 28 standard pressure levels and water vapour profile at 15 pressure levels from bottom atmosphere to top. The pressure levels on which temperature profile are reported in pressStd product. The pressure levels on which moisture products are reported in pressH2O product. The first 15 levels of pressStd are the same as the pressure levels in pressH2O. It has been shown that AIRS provides tropospheric retrieval of temperature with an accuracy close to 1-K in 1-km layers and water vapour retrievals with an uncertainty close to 15% in 2-km layers over sea and land (Divakarla et al. 2006).

B. Computing Mean Tropospheric Temperature from AIRS and AMSU

Denote \( T_i \) and \( e_i \) the air temperature and the water vapour partial pressure at the pressure level \( i \) respectively. The mean tropospheric temperature \( T_m \) can be retrieved using

\[
T_m = \frac{\sum e_i T_i}{\sum e_i} \quad \text{(8)}
\]

where \( T_i \) is the retrieved atmospheric temperature at the standard pressure level \( i \).

It is proposed in this paper to use \( T_i \) from TAirStd product of AIRX2RET in \( K \) and to compute \( e_i \) for each level from the mixing ratio at level \( i \) \((w_i)\) retrieved from H2OMMRLevStd product of AIRX2RET. The mixing ratio is the ratio of the mass of water vapour to the mass of dry air. The model between \( e_i \) and \( w_i \) is (Wallace and Hobbs 2006)

\[
e_i = \frac{w_i}{w_j + \varepsilon} p_i \quad \text{(9)}
\]

\( \varepsilon = M_w/M_d = 0.622 \) where \( M_w \) is the molecular weight of water and \( M_d \) is the apparent molecular weight of dry air. \( p_i \) is the the total pressure level where the mixing ratio is measured, which can be taken from PressH2O product. Note that the mixing ratio from AIRX2RET is in g/kg dry air. Hence, the factor of \( 10^{-3} \) must be multiplied to the product before being used in (9).

Now that we could extract \( T_m \) from AIRX2RET, it will be used along with the air surface temperature \( T_s \) from TSurfAir product to derive a local \( T_m \) model. For further information on the AIRX2RET products, the readers are referred to (Olsen et al. 2013).

\[\text{Fig. 1: Air surface temperature in time in year (left) and correlation between the mean tropospheric temperature and the air surface temperature (right), both for Bangkok from November 2013 to November 2015.}\]

C. Mean Tropospheric Temperature and Air Surface Temperature Models

Consider the evolution of \( T_s \) in time and the correlation between \( T_m \) and \( T_s \) for Bangkok from November 2013 to November 2015 in figure 1. We observe that \( T_s \) is periodic with a period of 1 year approximatively and that \( T_m \) seems to be linear in function of \( T_s \). We, therefore, propose that it is \( T_s \) that is periodic itself and that \( T_m \) is a linear function of the periodic \( T_s \) unlike what is proposed in (7) by Schueler et al. (2001) where \( T_m \) itself was considered periodic. To summarize, we propose that

\[
T_{sk} = \mu_k + A_k \sin(2\pi \frac{\text{DoY}}{365.25} + \phi_k) \quad \text{(10)}
\]

where \( \mu_k, A_k, \phi_k \) are the mean, the amplitude and the initial phase of the sine function at instant \( k \) respectively. This function is called a triply modulated sine function. It is inspired by the model of single crop rice in (Chumkesornkulkit et al. 2013).

Thanks to the correlation of \( T_m \) and \( T_s \) in figure 1, \( T_m \) is modelled by

\[
T_{mk} = a T_{sk} + b \quad \text{(11)}
\]

The parameters left to be estimated are therefore \( a \) and \( b \) to have a local \( T_m \) model for the area of interest.

V. DERIVING THE MEAN TROPOSPHERIC TEMPERATURE MODEL USING THE EXTENDED Kalman Filter

A. State and Measurement Definitions

We implement the Extended Kalman Filter (EKF) to estimate the parameters \( a \) and \( b \) for a local \( T_m \) model for the area of interest using \( T_m \) and \( T_s \) derived from AIRS and AMSU as measurements. Define \( x_k \) the state vector containing the parameters to be estimated at instant \( k \)

\[
x_k = \begin{bmatrix} a_k & b_k & \mu_k & A_k & \phi_k \end{bmatrix}^T \quad \text{(12)}
\]

where each component of the state is one of the parameters of \( T_s \) and \( T_m \) models in (10) and (11). Each instant \( k \) represents the instant when AIRS and AMSU measurements are available.

Assuming that the parameters do not vary much from one instant to another. The state is propagated using

\[
x_{k+1} = f(x_k) + w_k = x_k + w_k \quad \text{(13)}
\]
where \( w_k \sim \mathcal{N}(0, Q) \) is a zero-mean white Gaussian process noise with covariance matrix \( Q \). The function \( f(\cdot) \) is called the state equation.

Define \( y_k \) the measurement vector at instant \( k \) as

\[
y_k = \left( T_{m_k} \ T_{s_k} \right) + v_k \tag{14}
\]

where \( v_k \sim \mathcal{N}(0, R) \) is a zero-mean white Gaussian measurement noise with covariance matrix \( R \). \( T_{s_k} \) and \( T_{m_k} \) are the measured air surface temperature and the mean tropospheric temperature retrieved from AIRS and AMSU at instant \( k \) respectively.

The measurement vector is described in function of the state as

\[
y_k = h(x_k) + v_k = \left( \mu_k + A_k \sin(2\pi \frac{\text{DoY}_{k}}{365.25} + \phi_k) \right) + v_k
\]

The function \( h(\cdot) \) is called the measurement equation, which is nonlinear for this system. Hence, the nonlinear version of the Kalman Filter, the Extended Kalman Filter (EKF), will be used.

**B. Extended Kalman Filter Equations and Tuning**

\( a) \) EKF Equations: Denote \( \hat{x}_k \) the a priori estimate at instant \( k \), \( P_k^a \) the a priori estimation error covariance matrix, \( \hat{x}_k \) the a posteriori estimate and \( P_k \) the a posteriori error covariance matrix. The algorithm of the Extended Kalman Filter (EKF) can be divided into two stages: prediction and update as follows (Suwantong 2014):

**Prediction:**

\[
\hat{x}_k = f(\hat{x}_{k-1}) \tag{16}
\]

\[
P_k^a = Q_{k-1} + \hat{F}_{k-1} P_{k-1} \hat{F}_{k-1}^T \tag{17}
\]

**Update:**

\[
\hat{x}_k = \hat{x}_k + K_{k}(y_k - h(\hat{x}_k)) \tag{18}
\]

\[
P_k = P_k^a - K_k S_k K_k^T \tag{19}
\]

where

\[
S_k = \frac{\partial h}{\partial x} \mid _{x_{k-1}} \frac{\partial h}{\partial \hat{x}} \mid _{\hat{x}_{k}} \tag{20}
\]

\[
K_k = P_k^a \frac{\partial h}{\partial x} \mid _{x_{k-1}} \tag{21}
\]

\( \hat{F}_{k-1} = \frac{\partial f}{\partial x} \mid _{x_{k-1}} \) is the Jacobian of \( f \) evaluated at \( x_{k-1} \) and \( \frac{\partial h}{\partial x} \mid _{x_{k}} \) is the Jacobian of \( h \) evaluated at \( \hat{x}_k \).

\( b) \) Tuning for this study: The a priori initial state \( \hat{x}_0 \) is chosen such that

\[
\hat{x}_0 = (0.789 \ 60.4 \ 5 \ -67.5 \cdot \pi/180)^T \tag{22}
\]

where \( \bar{T} \) is the mean of all the \( T \) measurements. The values 5 and \(-67.5^\circ\) come from our approximation of the initial amplitude and phase in figure [1] (left). Note that the a priori initial value for \( a \) and \( b \) are chosen equal to the parameters of the global model in (5).

The initial a priori error covariance matrix \( P_0^a \) which represents the confidence in the a priori initial state is chosen such that

\[
P_0^a = \text{diag} \begin{pmatrix} 1 & 20 & 5 & 5 & 5 \cdot \pi/180 \end{pmatrix} \tag{23}
\]

The process noise covariance matrix \( Q \) is chosen such that

\[
Q = \text{diag} \begin{pmatrix} 0 & 0 & 0.5 & 0.5 \cdot \pi/180 \end{pmatrix} \tag{24}
\]

The process noises on \( a \) and \( b \) are supposed to be zero since \( a \) and \( b \) are assumed to be constant as done for the global models in (4), (5) and (9).

The measurement noise covariance matrix is chosen as

\[
R = \text{diag} \begin{pmatrix} 0.5 & 0.5 \end{pmatrix} \tag{25}
\]

due to the fact that AIRS temperature retrieval has been shown to have an accuracy close to 1-K in 1-km layers (Divakarla et al. 2006).

The EKF estimates the state at each instant but the estimated \( a \) and \( b \) at the final instant will be used for the local \( T_m \) model assuming that they approach the real values as time goes by.

The estimation results for the parameters for the \( T_m \) and \( T_s \) model will be presented in the next section.

**VI. Estimation Results**

\( A. \) Estimated Parameters for \( T_m \) and \( T_s \) Models

The estimated air surface temperature \( \bar{T} \) and the estimated mean tropospheric temperature \( \bar{T}_m \) are shown in figure [2] in °C. We observe low \( \bar{T}_s \) and \( \bar{T}_m \) during winter and high \( \bar{T}_s \) and \( \bar{T}_m \) during summer. The values of \( \bar{T}_s \) and \( \bar{T}_m \) are generally lower at night than at day and the values of \( \bar{T}_s \) are in the range of usual values of \( T_s \) in Bangkok.

The estimated parameters for the \( T_m \) model \( \tilde{a} \) and \( \tilde{b} \) are shown in figure [3] Using the final values of \( \tilde{a} \) and \( \tilde{b} \), we derive the model for day time for Bangkok as

\[
T_m = 0.6066T_s + 113.2914 \tag{26}
\]

and for night time as

\[
T_m = 0.7938T_s + 57.4856 \tag{27}
\]
The $T_m$ and $T_s$ from AIRS and AMSU along with global $T_m$ models and the derived local $T_m$ models are plotted together in figure 4 where it can be observed that the local models fit more the AIRS and AMSU observations than the global models. The local $T_m$ models (26) and (27) will be used for the PWV estimation using GNSS observations. The GNSS PWV using the local models will be compared to the GNSS PWV using Mendes global model in (5) seeing that it is the most fit to the measurements compared to two other global models in figure 4.

B. Accuracy of GNSS PWV using Local and Global Models

The accuracy of the PWV from GNSS using the local and the global models will be studied using the PWV values derived from AIRS and AMSU in the AIRX2RET products as the reference. This choice is legitimate seeing that the AIRS PWV retrievals have been shown to have a strong agreement with the Microwave Radiometer (MWR) PWV with a bias of 0.039 mm and rmse of 2.254 mm at day time and a bias of $-0.233$ mm and rmse of 2.103 mm at day time for 3 different ground truth sites at tropical, mid-latitude and Arctic from September 2002 through August 2008 (Bedka et al., 2010).

The ZTD from GNSS observation are processed using PANDA Software from Wuhan University. GNSS observations from the CORS own by the Department of Public Works and Town & Country Planning (DPT) of Thailand from January 2015 to November 2015 will be used for the PWV estimation.

Only GNSS data in 2015 is used seeing that it is the only year where meteorological data at the CORS is available.

The PWV derived from GNSS observations using the global and the local $T_m$ models are compared with the PWV from AIRS and AMSU in figure 5. Using the AIRS PWV as the reference, the GNSS PWV mean error and the rmse for both $T_m$ models are shown in table I in mm and in percent compared to the mean of the PWV from AIRS and AMSU.

We observe that the GNSS PWV using local and global $T_m$ models provide the same level of rmse. However, the local $T_m$ model provide GNSS PWV with less mean error. The global model provides a mean error of -1.06 mm while the local model provides a mean error of 0.20 mm. Seeing that AIRS PWV has been shown to have nearly 0 mm of bias compared to MWR PWV (Bedka et al., 2010). The GNSS PWV with local model should therefore have a nearly 0 mm of bias compared to MWR PWV as well.

VII. CONCLUDING REMARKS

This paper proposed to derive a local $T_m$ model using temperature and humidity profiles retrieved from satellite data: the Atmospheric Infrared Sounder (AIRS) and the Advanced Microwave Sounding Unit (AMSU) board on NASA’s Aqua satellite. Using $T_m$ and $T_s$ from AIRS and AMSU as measurements, the Extended Kalman Filter (EKF) is used to estimate the parameters of the local $T_m$ model. AIRS and AMSU data between November 2013 and November 2015 for Bangkok are used to derive a local $T_m$ model and the GNSS data from a CORS of the Department of Public Works and Town & Country Planning (DPT) of Thailand during 2015 is used for the GNSS PWV derivation.

Then, this paper compared the accuracy of the GNSS PWV using the local $T_m$ model to that of the GNSS PWV using the
global model. The PWV retrieved from AIRS and AMSU are used as the reference since they have high accuracy compared to the PWV from Microwave Radiometers. We showed that the GNSS PWV derived using the local \( T_m \) model have the same level of rmse as those derived using the global \( T_m \) model. However, the bias is reduced from \(-1.06 \) mm when the global model is used to \(0.20 \) mm when the local model is used. Therefore, it is important to derive a local \( T_m \) model for the area of interest to eliminate the bias in GNSS PWV estimation and satellite data, such as from AIRS and AMSU, could be efficiently used for local \( T_m \) model derivation.

The methods proposed in this study should emphasize the interests of using satellite data to derive local \( T_m \) model for better GNSS PWV accuracy at small cost and should be useful especially for areas where radiosonde data are not frequently available.

**ACKNOWLEDGEMENT**

The authors would like to thank the GNSS Research Center, Wuhan University of China for kindly providing us the PANDA Software for GNSS data processing. Also, we would like to express our gratitude to the Department of Public Works and Town & Country Planning of Thailand for GNSS observations from their CORS network. Also, we would like to thank Prof. Dr. Allen Huang from the Cooperative Institute for Meteorological Satellite Studies, Space Science and Engineering Center, University of Wisconsin-Madison, for his advice on water vapour retrieval from satellite data.

**REFERENCES**


