The following publication He, J., & Liu, Z. (2020). Water vapor retrieval from MODIS NIR channels using ground-based GPS data. IEEE Transactions on Geoscience and Remote Sensing, 58(5), 3726-3737 is available at https://doi.org/10.1109/TGRS.2019.2962057.

Water Vapor Retrieval from MODIS NIR Channels Using Ground-based GPS Data

Jia He, Zhizhao Liu

***Abstract*— A novel algorithm for water vapor retrieval from MODerate resolution Imaging Spectroradiometer (MODIS) Near Infrared (NIR) Channels is proposed in this research. In contrast to conventional retrieval algorithms based on radiative transfer methods, this algorithm uses the empirical regression functions to calculate precipitable water vapor (PWV). In this study, water vapor data observed from January 1st, 2003 to December 31st, 2017 from 464 GPS stations situated in western North America serve as reference data to determine the relationship between the transmittance of the water vapor absorption channels and atmospheric water vapor content. The model is trained on different subsets of the training data through the bootstrap resampling method. Validation results against PWV observations during the period 2010-2017 from 5 globally distributed GPS stations illustrate that the algorithm can significantly improve the accuracy of MODIS NIR water vapor data, with root-mean-squares error (RMSE) reduction of 22.48% from 7.670 mm to 5.946 mm for 2-channel ratio method and 21.69% from 7.670 mm to 6.006 mm for 3-channel ratio method for MODIS/Terra satellite data, and RMSE reduction of 16.42% from 7.191 mm to 6.010 mm and 15.26% from 7.191 mm to 6.094 mm for PWV derived from 2-channel and 3-channel ratio method from Aqua, respectively, for MODIS/Aqua satellite data.**

***Index Terms*— MODIS, GPS, PWV, Retrieval**

I. Introduction

Water vapor is the most important natural greenhouse gas in the atmosphere [1]. It has a significant impact on the process of the hydrological cycle, weather formation [2] and climate change [3]. Observations of atmospheric water vapor using remote sensing techniques have been widely accepted as the most cost-effective approach to estimate precipitable water vapor (PWV) at a global scale [4]–[6]. It has the widest range of monitoring categories [7] and the development of geostationary satellites makes it possible to continuously observe water vapor distribution with a coarser spatial resolutions [8].

The Global Climate Observing System (GCOS) declared that the Essential Climate Variables (ECV) requirement for quantifying climate observation on satellite-derived water vapor be with 5% measurement uncertainty and stability of 0.3% per decade [9]. Water vapor mapping using remote sensing technologies has benefited from instruments with better resolutions, and advances in computational storage and processing capabilities [10]–[12]. Infrared (IR) provides observation results in both daytime and nighttime based on split window technique [13]. Research shows that mean root-mean-squares error (RMSE) for MODIS IR product against radiosonde is 6.02 mm during daytime and 5.81 mm during nighttime, respectively [14]. Other IR water vapor retrieval using Advanced Infra-Red WAter Vapour Estimator (AIRWAVE) was proposed for Along-Track Scanning Radiometer (ATSR) [15]. Validation for ATSR using AIRWAVEv2 shows that the RMSE is 4.69 mm against Special Sensor Microwave/Imager (SSM/I) and is 6.13 mm against Analyzed RadioSounding Archive (ARSA) [16]. NIR channels are more sensitive to precipitable water vapor in the boundary layer, where most water vapor resides [17]. Water vapor calculated from NIR channels is based on the variations of the transmittance of absorption bands and the nearby window channel [18], [19]. The RMSE for NIR observations of POLarization and Directionality of the Earth's Reflectances (POLDER) onboard Advanced Earth Observing Satellite (ADEOS) is 3.1 mm [20]. The microwave (MW) can penetrate most of the dense clouds and provide more information than traditional infrared-visible retrieval schemes [21], [22]. The microwave radiance shows high correlation with water vapor under either 22.235 GHz or 183.3 GHz [23]. Validation analysis for microwave derived PWV from Advanced Microwave Scanning Radiometer - Earth Observing System (AMSR-E) onboard Aqua is reported to have an RMSE around 3.0 mm [23], and the RMSE for AMSR2 is 4.7 mm [24]. The water vapor retrieved from Atmospheric Infrared Sounder (AIRS) uses the combined IR and MW radiances. Evaluation of AIRS PWV against GPS PWV shows good agreement (within 5%) and the mean bias is less than 2 mm [25].

MODIS onboard the Terra and Aqua satellite platforms is the first space instrument to obtain PWV with NIR bands as well as the traditional IR bands. It measures in 36 spectral bands, covering the spectral bands from 0.4 to 15 μm [17]. Five of these bands in NIR are used for water vapor retrieval. Three absorbing channels are centered at 905 nm, 936 nm, and 940

The support from the Key Program of the National Natural Science Foundation of China (project No.: 41730109) is acknowledged. The supports by the Hong Kong Research Grants Council (RGC) (projects B-Q52W PolyU 152149/16E and B-Q61L PolyU 152222/17E) are acknowledged. The Emerging Frontier Area (EFA) Scheme of Research Institute for Sustainable Urban Development (RISUD) of the Hong Kong Polytechnic University (No. 1-BBWJ) is also acknowledged.

Jia He is with the Department of Land Surveying and Geo-Informatics, also with the Research Institute for Sustainable Urban Development, Hong Kong Polytechnic University, Hung Hum, Hong Kong (e-mail: gail.he@connect.polyu.hk).

Zhizhao Liu (correspondence author) is with the Department of Land Surveying and Geo-Informatics, also with the Research Institute for Sustainable Urban Development, Hong Kong Polytechnic University, Hung Hum, Hong Kong (e-mail: lszzliu@polyu.edu.hk).

nm while the two window channels are centered at 865 nm and 1240 nm. The operational products of MODIS NIR channels (MOD05 for Terra and MYD05 for Aqua) were derived based on a priori look-up table, which is estimated through High-resolution Transmission (HITRAN) 2000 [18], [26] or MODerate resolution atmospheric TRANsmission (MODTRAN) [27]. The employment of these atmospheric transmittance models requires a priori data of the atmosphere as input during the simulation process [28]. The common problem for this method is that it underestimates the transmittance variations at the majority of the absorption bands [29]. Validation analysis shows that the MOD05 data overestimated PWV against radiosonde with a scale factor from 1.14 to 1.20, while it overestimated against GPS by 7 to 14% [30]. Inter-comparisons among multi-source water vapor products over other places also show overestimation of MOD05 product [21], [31]–[33], indicating that a better model is needed for water vapor retrieval in MODIS NIR channels. Further development of retrieval algorithm of MODIS was performed using either a pre-calculated look-up table, regression method or an artificial neural network [17], [28], [34], [35]. Empirical correction coefficients were introduced for transmittance calculation to eliminate the wet bias of MODIS NIR products and the RMS deviations were between 0.9 and 2 mm against ground-based observations [36]. Optimization of water vapor measurement from MODIS was performed to improve the retrieval accuracy at a local scale. For instance, research indicated that the RMSE was reduced to 2.702 mm in western Iran [37].

Although improvements on MODIS NIR water vapor retrieval have been made with the employment of different algorithms, uncertainties in surface spectral reflectance, sensor calibration, atmospheric profile, channel shift, and mixed pixels remain in the retrieval process [26]. The quality of MODIS NIR water vapor product remains to be improved [35], [38], [39]. In this research, a novel PWV retrieval algorithm based on regression fitting with GPS-derived high accuracy PWV is proposed for MODIS NIR channels. Unlike the existing retrieval methods, which rely on theoretical modelling, this approach is empirically determined and fits well with the real scenario based on simple but accurate parameterization. Ensemble analysis is performed based on independent subsets of training data, which produce a robust and accurate model.

Section 2 gives a detailed description of the data used in this research. In section 3, the improved methodology of water vapor retrieval through MODIS NIR is discussed. In section 4, independent validation analysis against additional GPS observations is performed. Finally, the conclusions are drawn in section 5.

II. Data Description

Four pairs of MODIS data products from both Terra and Aqua satellites are utilized, including surface reflectance observations (MOD021KM/MYD021), geolocation data (MOD03/MYD03), cloud mask product (MOD35/MYD35), and level 2 NIR water vapor product (MOD05/MYD05) for comparative analysis. The detailed characteristics of the data are presented in Table 1. As the NIR wavelengths cannot penetrate cloud, water vapor estimation results have a poor accuracy under cloudy conditions [38], [39]. The cloud-mask MOD35/MYD35 is used as quality control flag. Only confident clear pixels are used in the algorithm development and validation. The data from January 1st, 2003 to December 31st, 2017 are used for the development of the new algorithm.

The GPS data used in this research is hourly water vapor data derived from SuomiNet GPS network by University Corporation for Atmospheric Research (UCAR) (http://www.suominet.ucar.edu/data.html). Phase delays of GPS signals can be converted into integrated water vapor. Thus GPS stations provide continuous, accurate, all-weather and all-time water vapor measurement over these stations [40]. The absolute errors for this data are less than 2 millimeters [41]. PWV observations during the period from January 1st, 2003 to December 31st, 2017 from 469 ground-based GPS stations from Continental United States (CONUS) sites are used in this research for model development. PWV data for the period 2010 to 2017 from 5 globally distributed GPS stations are used for validation. To reduce the impact of temporal discrepancies between GPS and MODIS remote sensing water vapor observations, the allowable time difference between the two data sources is 30 minutes.

In short, spatially and temporally collocated GPS and MODIS data collected under the cloud-free condition are used for both model construction (119,417 pairs for Terra and 121,800 pairs for Aqua) and model validation (1,527 pairs for Terra and 1,396 for Aqua).

III. Methodology for MODIS NIR Water Vapor Retrieval

Atmospheric water vapor is related to the transmission in the spectral channel. Existing differences in all 5 bands imply different spectral response functions (SRF) characteristics between the Terra and Aqua platforms. Therefore, regression fitting for individual MODIS sensor is expected. Detailed descriptions of the five NIR channels of MODIS are given in Table 2.

Before we began developing the new model for water vapor retrieval, the current retrieval algorithm for MOD05 product is examined. In the operational MODIS NIR PWV product, the relation between the measured radiance ratio and water vapor was calculated using a radiative transfer model (RTM) for a large variety of different atmospheric profiles [18], [28], [38]. The relationship can be expressed by an exponential formula written as:

$$T_w = \exp\,(\alpha - \beta\sqrt{W^*}) \tag{1}$$

where T_w is the transmittance of a water vapor, α and β are determined by the surface type, and the W^* is water vapor along the sun-surface-sensor (slant) path [26].

As suggested in previous research, the current MODIS NIR PWV products tend to systematically overestimate water vapor values [42], [43]. The critical step to improve the retrieval algorithm is to mathematically describe the relationship between the transmittance and atmospheric water vapor content in a more accurate way. Therefore, regression analysis using the least squares curve fitting method is performed, where the GPS water vapor data are treated as the ground truth.

The flow chart of this retrieval scheme is illustrated in Figure 1. Firstly, the ratio method is utilized to reduce the effect of ground surface type while calculating the transmittance. Cloud mask products from MODIS are employed as quality control flag. Only confident clear pixels are selected for water vapor retrieval in this research. The collocated data points were resampled into independent training and testing subsets using the bootstrap method. Ensemble functions were then generated from 10 separate resampling training sets. The GPS data are used to quantify the relationship between water vapor and the transmittance from three absorption channels. By employing the regression functions, the water vapor could be estimated from the MODIS Level 1B data.

A. Transmittance Calculation Using Differential Absorption Techniques

Observation of transmittance is one of the most crucial steps in atmospheric remote sensing. The earth surface varies from location to location. Therefore, the surface reflectance becomes a main source of uncertainty in the estimation of water vapor absorption. Based on the theory of molecular physics, the reflectance of the earth atmosphere is affected by the aero-physical characteristics of the molecules [44]. Furthermore, only molecules with asymmetric-top could affect the transmission of solar radiation [45]. Solar radiation between 860 nm and 1240 nm on the slant path is subject to atmospheric water vapor absorption, atmospheric aerosol scattering, and surface reflection [26], [39]. Water vapor (H_2O), a non-symmetric molecular, contributes the most to the transmission decrease of channels between 930 to 950 nm, while the transmission of symmetric molecular (O_3) remains the same among MODIS NIR channels [46].

As the water vapor transmittance cannot be observed directly, interpolation of surface reflectance between two or three channels around water vapor absorption is conducted to estimate the transmittance. In this case, the ratio technique is employed to calculate the transmittance of atmospheric water vapor. This technique is built on the differential absorption method, which assumes that the transmittance of solar energy can be estimated through the reflectance ratio between one absorption channel and one or two window channels [17].

For most types of ground surface cover, the reflectance varies linearly with wavelength. The ratio partially eliminates the impact of surface reflectance on different wavelengths, and it is approximately equal to the atmospheric water vapor transmittance [34]. The functions using a two-channel ratio method to calculate the transmittances of T_{17}, T_{18}, and T_{19} are defined as:

$$T_i \cong R_i = \frac{L_i}{L_2} \tag{2}$$

where T_i is the transmittance of channel i, which is approximately equal to the reflectance ratio R_i. L_i is the reflectance in absorption channel i, i = 17, 18 and 19. L_2 is the reflectance in window channel 2 (centred at 865 nm).

For a complex land surface with variable reflectance spectra, more window channels are required to estimate the transmittance in the water vapor absorption channel. Therefore, an additional window channel 5 (centred at 1240 nm) is also included, and a three-channel ratio method is employed to estimate the transmission, which can be calculated as:

$$T_i \cong R_i = \frac{L_i}{[C_1 L_2 + C_2 L_5]} \tag{3}$$

where the coefficients C_1 and C_2 are prescribed as 0.8 and 0.2, respectively. It is assumed that the reflectance ratio around 1 μm remains the same, or the reflectance ratio varies linearly [38].

B. Water Vapor Retrieval from MODIS NIR Channels

The key procedure in this algorithm is to accurately model the relationship between water vapor concentration and transmittance from each absorption channel. Previous studies on atmospheric transmission variation at different water vapor levels using MODTRAN show that the total atmospheric transmission decreases with the increase of water vapor [18], [26]. The largest decrease in transmittance occurs at Band 18, the strongest absorption band. Band 17 is the least sensitive band to water vapor variation among the three absorption bands. Band 19 has a moderate sensitivity. On the other hand, the transmission in the window channels has a weak dependence on water vapor concentration. In conclusion, the transmittance in the NIR absorption channels can represent the magnitude of radiance attenuation caused by water vapor.

The selection of an exponential function is based on the examination of the numerical relationship between the MODIS transmittance and water vapor content. The functions from RTM are developed based on simplified assumptions in order to reduce the complexity of band transmittance calculation. This will lead to large transmittance errors, which in turn lead to spectrally dependent flux and heating rate errors [29]. For a complex and variable land surface, the α in equation (1) is unlikely to be zero. As displayed in Figure 2, the relationship between transmittance T_W and sun-surface-sensor optical path (slant path) water vapor can be well characterized by an exponential function. After studying the properties of many different functions, the best results are obtained by expressing the transmittance as:

$$T_i = a \exp\,(b\, W_i^*) + c \exp\,(d\, W_i^*) \tag{4}$$

where T_i is the transmittance from MODIS NIR channel i; W_i^* is the slant path water vapor content at the channel i; a, b, c and d are the coefficients to be determined. The vertical total precipitation water vapor (W) is written as:

$$W^* = W\left(\frac{1}{\cos\theta} + \frac{1}{\cos\theta_0}\right) \tag{5}$$

where θ is the view zenith angle and θ_0 is the solar zenith angle [26]. Nevertheless, the results may be affected for observations with large view and solar zenith angle because of the stronger atmospheric effect of aerosol scattering through the longer optical path [26].

C. Optimization of Channel Selection

Water vapor values can be estimated from individual absorption channel. The absorption channel at 936 nm is more sensitive to water vapor variation under dry conditions, while the absorption channel at 905 nm is more sensitive to water vapor under humid conditions. To get a more accurate water vapor value, the weighted mean PWV from the three absorption channels are calculated as total column water vapor as follow:

$$W = f_1 W_{17} + f_2 W_{18} + f_3 W_{19} \quad (6)$$

where W_{17}, W_{18} and W_{19} are water vapor estimated from band 17, band 18 and band 19, respectively; f_1, f_2 and f_3 are the corresponding weighting parameters [18], [39].

The coefficients f_1 , f_2 and f_3 are calculated from normalized values of sensitivity in each simulation band:

$$f_i = \frac{\eta_i}{\eta_1 + \eta_2 + \eta_3} \quad (7)$$

where η_i is the slope of the graph of transmission versus water vapor for each water vapor absorption band of MODIS. It represents the sensitivity of transmission in each absorption band:

$$\eta_i = \left|\frac{dT_w}{dW_i}\right| \quad (8)$$

D. Training of Regression Algorithm

A total of 119,417 pairs of MODIS-GPS collocated water vapor data points for Terra and 121,800 pairs for Aqua have been observed for the western part of the North America continent (Figure 3), covering the area from 15 °N to 50 °N and 90 °W to 130 °W. They are used for model development because this is a large dataset and the geographic region covers a diversity of climates. However, a common problem with empirical regression model is that the selection of the training data might influence the outcome, as the model is highly data-dependent [47]. An ensemble-based algorithm is introduced to solve this potential imbalance problem in the training dataset. The multiple classifiers could have a better answer than a single one as it can average prediction errors and reduce the bias and variance of errors [48]. In this research, the bootstrap resampling method [48] is applied to balance class distribution. It is a resampling technique used to estimate statistics of a population by sampling a dataset [49], [50]. By resampling the collocated datasets into 10 independent training and testing subsets, the obtained regression functions are expected to be bias-robust [50].

To reduce the effect of random sampling errors generated from the bootstrap procedure itself, and minimize the sensitivity of possible channel drifting in the channel position over the years, and find a proper sampling number for training sets, a series of tests using different numbers of samples (Table 3) have been performed in this research. Results show that the standard deviation for the ensemble members of water vapor from the test datasets is smaller than the observation error of GPS PWV (1~2 mm), indicating that this empirical regression model is bias-robust. To take both quantity and variability of the subsets into account, around 70% of total training data (66,500 pairs for Terra and 62,500 pairs for Aqua) are used in the ensemble analysis. The least squares fitting parameters are listed in Table 4 and Table 5.

E. Verification of the Ensemble Analysis

With the above retrieval procedure, the new sets of ensemble members of MODIS NIR PWV from both Terra and Aqua can be recalculated. To evaluate the performance of the proposed new water vapor retrieval scheme, verification results of the corresponding ensemble test subsets against GPS PWV observation are discussed.

Statistical metrics used to evaluate the performance of the proposed new algorithm are the coefficient of determination (R^2), mean bias (MB), and root mean squares error (RMSE). The metrics are written as:

$$R^2 = \left[\frac{\sum_{i=1}^{n}\left(PWV_{R_i} - \overline{PWV}_R\right)\left(PWV_{O_i} - \overline{PWV}_O\right)}{\sqrt{\sum_{i=1}^{n}\left(PWV_{R_i} - \overline{PWV}_R\right)^2\left(PWV_{O_i} - \overline{PWV}_O\right)^2}}\right]^2 \quad (9)$$

$$MB = \frac{1}{N}\sum_{i=1}^{n}\left(PWV_{R_i} - PWV_{O_i}\right) \quad (10)$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left|PWV_{R_i} - PWV_{O_i}\right|^2} \quad (11)$$

where the PWV_{R_i} is the reference PWV from GPS, $\overline{PWV}_R$ is the mean PWV from GPS, PWV_{O_i} is the observed PWV from MODIS. The R^2 shows the relationship strength between the calibrated water vapor and reference PWV datasets of GPS; the MB indicates the systematic difference between two PWV datasets; the standard deviation quantifies the dispersion of the relative differences; and the RMSE measures the overall agreement between calibrated PWV observations and the reference PWV data.

As shown in Figure 4, the calibrated water vapor has improved accuracy for each ensemble member. The operational products of MOD05 and MYD05 have a wet bias of 22.0% and 21.9%, respectively. For MODIS/Terra, the RMSE has reduced 48.12% to 2.362 mm for PWV calculated from 2-channel ratio method and has reduced 50.74% to 2.243 mm for 3-channel ratio method derived PWV. Meanwhile, the MB has reduced from 3.281 mm to 0.054 mm and -0.001 mm for 2-channel ratio method and 3-channel ratio method, respectively. For MODIS/Aqua, the RMSE has reduced 42.54% to 2.562 mm and has reduced 42.99% to 2.541 mm for PWV retrieved from 2-channel and 3-channel ratio method, respectively. The MB has reduced from 2.920 mm to 0.115 mm and 0.055 mm for 2-channel ratio method and 3-channel ratio method, respectively.

IV. Validation of the New Water Vapor Retrieval Algorithm

In order to evaluate the performance of this newly developed retrieval method, validation against additional GPS stations are discussed in detail. A total of 1,527 pairs for Terra and 1,396 pairs for Aqua under clear conditions during daytime obtained from 5 global stations in different climate zones are selected (Figure 5). These stations include Alice Springs, Australia (ALIC) of hot desert; Kiruna, Sweden (KIRU) of arctic region; Quezon City, Philippines (PIMO) of tropical monsoon region; Braunschweig, Germany (PTBB) from mid-latitude; and Salta, Argentina (UNSA) of sub-tropical highland. The detailed characteristics are listed in Table 6. These stations are selected as they have a relatively long period of GPS observations, and they are representative of the climatic regions.

Validation results in Figure 6 show that the operational PWV products from Aqua perform better than Terra, with higher correlation and smaller RMSE against GPS PWV. Moreover, the employment of new retrieval algorithm can significantly reduce the wet bias of operational products on a global scale. For MODIS/Terra, the RMSE has reduced 22.48% from 7.670 mm to 5.946 mm for PWV derived from 2-channel ratio method and has reduced 21.69% to 6.006 mm for 3-channel ratio derived PWV. For MODIS/Aqua, the RMSE has reduced 16.42% from 7.191 mm to 6.010 mm for 2-channel ratio method and 15.26% from 7.191 mm to 6.094 mm for PWV 3-channel ratio method.

The detailed validation result for each station is shown in Table 7. All stations have shown improvement in retrieval accuracy in terms of RMSE reduction compared to the MODIS operational PWV product. In particular, the calibrated PWV from ALIC using 2-channel ratio method has the smallest wet bias among these stations, with a reduced wet bias of 5.00% from Terra and 6.50% from Aqua. The PWV retrieved from Terra over UNSA station has the largest reduction rate in RMSE, which drops 45.23% to 4.517 mm using 2-channel ratio method and 44.88% to 4.546 mm using 3-channel ratio method. The decline of RMSE is also obvious over ALIC and PIMO stations, with over 10% decrease in all calibrated products, indicating that this empirical fitting algorithm is globally applicable and valid.

V. Conclusion

Water vapor can be estimated from remote sensing satellites through transmittance measurement of water vapor absorption channels. Accurately identifying the relationship between the transmittance and atmospheric water vapor content is the key step in retrieval algorithm. In MODIS NIR PWV operational products, this relationship is calculated through a radiative transfer model. This conventional method is based on several simplifying assumptions, which requires pre-calculated input parameters of atmospheric profiles. A systematic wet bias has been observed.

A new algorithm is proposed in this study, which aims to retrieve water vapor from MODIS NIR channels using regression functions derived from ground-based GPS PWV data. It provides an effective way to retrieve water vapor with significantly improved accuracy. The training data samples in the fitting procedure are constructed by GPS-based PWV observations collected in various environmental conditions by the SuomiNet in the western North America region. It is resampled into 10 subsets based on the bootstrap method. The regression functions trained by those independent subsets minimize the uncertainty in the model training and minimize the sensitivity of possible channel drifting. Therefore, the results of ensemble analysis are improved over the whole absorption channel.

Verification of the ensemble analysis shows that the RMSE for calibrated MODIS/Terra PWV data has reduced to 2.362 mm using 2-channel ratio method, and has reduced to 2.243 mm using 3-channel ratio method. For MODIS/Aqua, the RMSE of calibrated PWV has reduced to 2.562 mm and 2.541 mm using 2-channel and 3-channel ratio method, respectively. Validation against PWV from 5 global GPS stations shows that the RMSE of PWV data has reduced 22.48% from 7.670 mm to 5.946 mm using 2-channel ratio method and has reduced 21.69% from 7.670 mm to 6.006 mm for 3-channel ratio method. For MODIS/Aqua, the RMSE of calibrated PWV has reduced 16.42% from 7.191 mm to 6.010 mm using 2-channel ratio method and reduced 15.26% from 7.191 mm to 6.094 mm using 3-channel ratio method.

In summary, this empirical regression model can significantly reduce the wet bias and RMSE for most occasions. Although a large number of training data are employed in the model construction, the number of data points under extremely wet and arid conditions is however still limited, which may result in an underestimation of transmittance variation under these conditions. Analysis with more training data from these extreme conditions is likely to further improve the performance of the model.

Acknowledgment

This research used the SPECTRE High Performance Computing Facility at the University of Leicester, United Kingdom. The authors would like to thank the research staff in National Aeronautics and Space Administration (NASA) Goddard Earth Sciences Distributed Active Archive Center (DAAC) (https://search.earthdata.nasa.gov) for providing MODIS NIR reflectance data and water vapor data. They would like to thank the University Corporation for Atmospheric Research (UCAR) and SuomiNet project (http://www.suominet.ucar.edu/) for providing GPS water vapor data. The support to Jia He by Prof. John Remedios and Dr. Tim Trent from the National Centre for Earth Observation, University of Leicester, United Kingdom, is acknowledged.

References

[1] I. M. Held and B. J. Soden, “Water Vapor Feedback and Global Warming,” *Annu. Rev. Energy Environ.*, vol. 25, no. 1, pp. 441–475, 2000.

[2] B. J. Soden, R. T. Wetherald, G. L. Stenchikov, and A. Robock, “Global Cooling After the Eruption of Mount Pinatubo: A Test of Climate Feedback by Water Vapor,” *Science*, vol. 296, no. 5568, pp. 727–730, Apr. 2002.

[3] T. R. Karl and K. E. Trenberth, “Modern Global Climate Change,” *Science*, vol. 302, no. 5651, pp. 1719–1723, Dec. 2003.

[4] C. Prabhakara, H. D. Change, and A. T. C. Chang, “Remote sensing of precipitable water over the oceans from Nimbus-7 microwave measurements,” *J. Appl. Meteorol.*, vol. 21, pp. 59–68, Apr. 1981.

[5] K. E. Trenberth and C. J. Guillemot, “Evaluation of the Global Atmospheric Moisture Budget as Seen from Analyses,” *J. Clim.*, vol. 8, no. 9, pp. 2255–2272, Sep. 1995.

[6] K. E. Trenberth, J. Fasullo, and L. Smith, “Trends and variability in column-integrated atmospheric water vapor,” *Clim. Dyn.*, vol. 24, no. 7–8, pp. 741–758, Jun. 2005.

[7] H. H. Aumann *et al.*, “AIRS/AMSU/HSB on the aqua mission: design, science objectives, data products, and processing systems,” *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 2, pp. 253–264, Feb. 2003.

[8] S. Bouffiès, F. M. Bréon, D. Tanré, and P. Dubuisson, “Atmospheric water vapor estimate by a differential absorption technique with the polarisation and directionality of the Earth reflectances (POLDER) instrument,” *J. Geophys. Res. Atmospheres*, vol. 102, no. D3, pp. 3831–3841, Feb. 1997.

[9] A. Belward, “The Global Observing System for Climate: Implementation Needs,” *Ref. Number GCOS-200*, vol. 315, p. 349, 2016.

[10] R. F. Hanssen, T. M. Weckwerth, H. A. Zebker, and R. Klees, "High-Resolution Water Vapor Mapping from Interferometric Radar Measurements," *Science*, vol. 283, no. 5406, pp. 1297–1299, Feb. 1999.

[11] R. Lindenbergh, M. Keshin, H. van der Marel, and R. Hanssen, "High resolution spatio-temporal water vapour mapping using GPS and MERIS observations," *Int. J. Remote Sens.*, vol. 29, no. 8, pp. 2393–2409, Apr. 2008.

[12] N. Levin, K. Johansen, J. M. Hacker, and S. Phinn, "A new source for high spatial resolution night time images — The EROS-B commercial satellite," *Remote Sens. Environ.*, vol. 149, pp. 1–12, Jun. 2014.

[13] D. Chesters, L. W. Uccellini, and W. D. Robinson, "Low-Level Water Vapor Fields from the VISSR Atmospheric Sounder (VAS) 'Split Window' Channels," *J. Clim. Appl. Meteorol.*, vol. 22, no. 5, pp. 725–743, May 1983.

[14] H. Liu, S. Tang, S. Zhang, and J. Hu, "Evaluation of MODIS water vapour products over China using radiosonde data," *Int. J. Remote Sens.*, vol. 36, no. 2, pp. 680–690, Jan. 2015.

[15] S. Casadio *et al.*, "Total column water vapour from along track scanning radiometer series using thermal infrared dual view ocean cloud free measurements: The Advanced Infra-Red WAter Vapour Estimator (AIRWAVE) algorithm," *Remote Sens. Environ.*, vol. 172, pp. 1–14, Jan. 2016.

[16] E. Castelli, E. Papandrea, A. D. Roma, B. M. Dinelli, S. Casadio, and B. Bojkov, "The Advanced Infra-Red WAter Vapour Estimator (AIRWAVE) version 2: algorithm evolution, dataset description and performance improvements," *Atmospheric Meas. Tech.*, vol. 12, no. 1, pp. 371–388, Jan. 2019.

[17] M. D. King, Y. J. Kaufman, W. P. Menzel, and D. Tanre, "Remote sensing of cloud, aerosol, and water vapor properties from the Moderate Resolution Imaging Spectrometer (MODIS)," *IEEE Trans. Geosci. Remote Sens.*, vol. 30, no. 1, pp. 2–27, 1992.

[18] B.-C. Gao and Y. J. Kaufman, "Water vapor retrievals using Moderate Resolution Imaging Spectroradiometer (MODIS) near-infrared channels," *J. Geophys. Res. Atmospheres*, vol. 108, no. D13, Jul. 2003.

[19] J. Fischer and R. Bennartz, "Retrieval of Total Water Vapor Content from MERIS Measurements," *Algorithm Theor. Basis Doc. PO-NT-MEL-GS-005*, 1997.

[20] M. Vesperini, F.-M. Breon, and D. Tanré, "Atmospheric water vapor content from spaceborne POLDER measurements," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 3, pp. 1613–1619, May 1999.

[21] J. Van Baelen, J.-P. Aubagnac, and A. Dabas, "Comparison of Near–Real Time Estimates of Integrated Water Vapor Derived with GPS, Radiosondes, and Microwave Radiometer," *J. Atmospheric Ocean. Technol.*, vol. 22, no. 2, pp. 201–210, Feb. 2005.

[22] D. L. Jackson, G. A. Wick, and J. J. Bates, "Near-surface retrieval of air temperature and specific humidity using multisensor microwave satellite observations," *J. Geophys. Res. Atmospheres*, vol. 111, no. D10, p. D10306, May 2006.

[23] B. Stefania, M. Vinia, B. Patrizia, C. Piero, and P. Nazzareno, "Satellite-Based Retrieval of Precipitable Water Vapor Over Land by Using a Neural Network Approach," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 9, pp. 3236–3248, Sep. 2011.

[24] C. Du, H. Ren, Q. Qin, J. Meng, and S. Zhao, "A Practical Split-Window Algorithm for Estimating Land Surface Temperature from Landsat 8 Data," *Remote Sens.*, vol. 7, no. 1, pp. 647–665, Jan. 2015.

[25] J. Roman, R. Knuteson, T. August, T. Hultberg, S. Ackerman, and H. Revercomb, "A global assessment of NASA AIRS v6 and EUMETSAT IASI v6 precipitable water vapor using ground-based GPS SuomiNet stations: IR PWV ASSESSMENT," *J. Geophys. Res. Atmospheres*, vol. 121, no. 15, pp. 8925–8948, Aug. 2016.

[26] Y. J. Kaufman and B. C. Gao, "Remote sensing of water vapor in the near IR from EOS/MODIS," *IEEE Trans. Geosci. Remote Sens.*, vol. 30, no. 5, pp. 871–884, Sep. 1992.

[27] D. Schläpfer, C. C. Borel, J. Keller, and K. I. Itten, "Atmospheric Precorrected Differential Absorption Technique to Retrieve Columnar Water Vapor," *Remote Sens. Environ.*, vol. 65, no. 3, pp. 353–366, Sep. 1998.

[28] P. Albert, R. Bennartz, R. Preusker, R. Leinweber, and J. Fischer, "Remote Sensing of Atmospheric Water Vapor Using the Moderate Resolution Imaging Spectroradiometer," *J. Atmospheric Ocean. Technol.*, vol. 22, no. 3, pp. 309–314, Mar. 2005.

[29] J. X. Warner and R. G. Ellingson, "A New Narrowband Radiation Model for Water Vapor Absorption," *J. Atmospheric Sci.*, vol. 57, no. 10, pp. 1481–1496, May 2000.

[30] Z. Li, J.-P. Muller, and P. Cross, "Comparison of precipitable water vapor derived from radiosonde, GPS, and Moderate-Resolution Imaging Spectroradiometer measurements," *J. Geophys. Res. Atmospheres*, vol. 108, no. D20, p. 4651, Oct. 2003.

[31] Z. Liu, M. S. Wong, J. Nichol, and P. W. Chan, "A multi-sensor study of water vapour from radiosonde, MODIS and AERONET: a case study of Hong Kong," *Int. J. Climatol.*, vol. 33, no. 1, pp. 109–120, Jan. 2013.

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

[33] J. Vaquero-Martínez *et al.*, "Inter-comparison of integrated water vapor from satellite instruments using reference GPS data at the Iberian Peninsula," *Remote Sens. Environ.*, vol. 204, pp. 729–740, Jan. 2018.

[34] K. B. Mao *et al.*, "Estimation of water vapor content in near-infrared bands around 1 μm from MODIS data by using RM–NN," *Opt. Express*, vol. 18, no. 9, pp. 9542–9554, Apr. 2010.

[35] K. Mao, X. Shen, Z. Zuo, Y. Ma, G. Liu, and H. Tang, "An Advanced Radiative Transfer and Neural Network Scheme and Evaluation for Estimating Water Vapor Content from MODIS Data," *Atmosphere*, vol. 8, no. 8, p. 139, Jul. 2017.

[36] H. Diedrich, R. Preusker, R. Lindstrot, and J. Fischer, "Retrieval of daytime total columnar water vapour from MODIS measurements over land surfaces," *Atmos Meas Tech*, vol. 8, no. 2, pp. 823–836, Feb. 2015.

[37] M. H. Merrikhpour and M. Rahimzadegan, "Improving the Algorithm of Extracting Regional Total Precipitable Water Vapor Over Land From MODIS Images," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 10, pp. 5889–5898, Oct. 2017.

[38] S. W. Seemann, E. E. Borbas, J. Li, W. P. Menzel, and L. E. Gumley, "MODIS atmospheric profile retrieval algorithm theoretical basis document," Univ. of Wis.–Madison, Madison, 2006.

[39] J. A. Sobrino, J. E. Kharraz, and Z.-L. Li, "Surface temperature and water vapour retrieval from MODIS data," *Int. J. Remote Sens.*, vol. 24, no. 24, pp. 5161–5182, Jan. 2003.

[40] R. H. Ware *et al.*, "SuomiNet: A real–time national GPS network for atmospheric research and education," *Bull. Am. Meteorol. Soc.*, vol. 81, no. 4, pp. 677–694, 2000.

[41] J. Braun, C. Rocken, and J. Liljegren, "Comparisons of Line-of-Sight Water Vapor Observations Using the Global Positioning System and a Pointing Microwave Radiometer," *J. Atmospheric Ocean. Technol.*, vol. 20, no. 5, pp. 606–612, May 2003.

[42] A. K. Prasad and R. P. Singh, "Validation of MODIS Terra, AIRS, NCEP/DOE AMIP-II Reanalysis-2, and AERONET Sun photometer derived integrated precipitable water vapor using ground-based GPS receivers over India," *J. Geophys. Res.*, vol. 114, no. D5, p. D05107, Mar. 2009.

[43] N. S. Shantikumar, S. Jade, T. S. Shrungeshwara, and H.-J. Song, "Validation of water vapor retrieval from Moderate Resolution Imaging Spectro-radiometer (MODIS) in near infrared channels using GPS data over IAO-Hanle, in the trans-Himalayan region," *J. Atmospheric Sol.-Terr. Phys.*, vol. Vol. 137, pp. 76–85, Jan. 2016.

[44] R. S. Fraser and Y. J. Kaufman, "The Relative Importance of Aerosol Scattering and Absorption in Remote Sensing," *IEEE Trans. Geosci. Remote Sens.*, vol. GE-23, no. 5, pp. 625–633, Sep. 1985.

[45] R. E. Roberts, J. E. A. Selby, and L. M. Biberman, "Infrared continuum absorption by atmospheric water vapor in the 8–12-μm window," *Appl. Opt.*, vol. 15, no. 9, pp. 2085–2090, Sep. 1976.

[46] A. Berk, P. Conforti, R. Kennett, T. Perkins, F. Hawes, and J. van den Bosch, "MODTRAN6: a major upgrade of the MODTRAN radiative transfer code," in *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XX*, 2014, vol. 9088, p. 90880H.

[47] J. D. Barrow, S. G. Bhavsar, and D. H. Sonoda, "A bootstrap resampling analysis of galaxy clustering," *Mon. Not. R. Astron. Soc.*, vol. 210, no. 1, pp. 19P-23P, Sep. 1984.

[48] G. E. A. P. A. Batista, R. C. Prati, and M. C. Monard, "A Study of the Behavior of Several Methods for Balancing Machine Learning Training Data," *SIGKDD Explor Newsl*, vol. 6, no. 1, pp. 20–29, Jun. 2004.

[49] B. Efron, "Bootstrap Methods: Another Look at the Jackknife," *Ann. Stat.*, vol. 7, no. 1, pp. 1–26, Jan. 1979.

[50] C. F. J. Wu, “Jackknife, Bootstrap and Other Resampling Methods in Regression Analysis,” *Ann. Stat.*, vol. 14, no. 4, pp. 1261–1295, Dec. 1986.

Jia He received the B.Sc. degree in geography from Beijing Forestry University, Beijing, China, in 2013, and M.Sc. degree in Geo-Information Science from the Chinese University of Hong Kong, Hong Kong, China, in 2015. She is currently working toward the Ph.D. degree in the Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hong Kong, China. Her current research interests include water vapor retrieval with multi-sensor observations.

Zhizhao Liu received the B.Sc. degree in Surveying Engineering from Jiangxi University of Science and Technology, Ganzhou, China, in 1994; the M.Sc. degree in Geodesy from the Wuhan University, Wuhan, China, in 1997; and the Ph.D. degree in Geomatics Engineering from the University of Calgary, Calgary, Alberta, Canada, in 2004.

He is currently an Associate Professor with the Department of Land Surveying and Geo-Informatics, the Hong Kong Polytechnic University, Kowloon, Hong Kong. His research interests include new algorithm development for precise Global Positioning System (GPS) and Global Navigation Satellite System (GNSS), GPS/GNSS precise point positioning (PPP), ionosphere modeling and scintillation monitoring, tropospheric remote sensing and modeling, and GPS/GNSS meteorology. He has over 20 years of experience in GPS/GNSS research. His group developed China’s first GPS PPP-based Precipitable Water Vapor Real-time Monitoring System in the Pearl-River-Delta region in 2012. His research group established Hong Kong’s first GPS/GNSS-radiosonde water vapor sounding collocation system in 2013 in collaboration with Hong Kong Observatory. In 2012, his group developed Hong Kong’s first GPS/GNSS-based ionosphere scintillation monitoring system (two stations deployed in South and North Hong Kong) with his collaborators.

Dr. Liu was the recipient of the inaugural Early Career Award of the Hong Kong Research Grants Council (RGC), Hong Kong, in 2012 and was the recipient of the inaugural Best Conference Paper of the China Satellite Navigation Conference (CSNC), China, in 2013. In 2014, he was nominated by the Hong Kong Observatory for the World Meteorological Organization (WMO) “Norbert Gerbier-MUMM International Award for 2015” for his paper that has developed a method to evaluate the absolute accuracy of water vapor measurements, in addition to other numerous awards and honors.