The following publication He, J., & Liu, Z. (2020). Refining MODIS NIR atmospheric water vapor retrieval algorithm using GPS-derived water vapor data. IEEE Transactions on Geoscience and Remote Sensing, 59(5), 3682-3694 is available at https://doi.org/10.1109/TGRS.2020.3016655.

Refining MODIS NIR Atmospheric Water Vapor Retrieval Algorithm Using GPS-derived Water Vapor Data

Jia He, Zhizhao Liu

Abstract— **A new algorithm of retrieving atmospheric water vapor from MODIS near-infrared (NIR) data by using a regression fitting method based on GPS-derived water vapor is developed in this work. The algorithm has been used to retrieve total column water vapor from MODIS satellites both Terra and Aqua under cloud-free conditions from solar radiation in the NIR channels. Water vapor data estimated from GPS observations recorded from 2003 to 2017 by the SuomiNet GPS network over the western North America are used as ground truth references. The GPS stations were classified into six subsets based on the surface types adopted from MCD12Q1 IGBP legend. The differences in surface types are considered in the regression fitting procedure, thus different regression functions are trained for different surface types. Thus, the wet bias in the operational MODIS water vapor products has been significantly reduced. Water vapor retrieved from each of the three absorption channels and the weighted water vapor of combined three absorption channels are analyzed. Validation shows that the weighted water vapor performs better than single-channel results. Compared to the MODIS/Terra water vapor products, the RMSE has been reduced by 50.78% to 2.229 mm using the two-channel ratio transmittance method and has been reduced by 53.06% to 2.126 mm using the three-channel ratio transmittance method. Compared to the MODIS/Aqua water vapor products, the RMSE has been reduced by 45.54% to 2.423 mm using the two-channel ratio transmittance method and has been reduced by 45.34% to 2.432 mm using the three-channel ratio transmittance method.**

Index Terms— **MODIS, GPS, PWV, Land cover**

I. INTRODUCTION

Water vapor is one of the most important components of Earth's atmosphere that influences many atmospheric processes, providing latent heat, affecting thermal structure of the atmosphere and the energy balance between surface and atmosphere [1]. It has a significant impact on the hydrological cycle [2], weather formation [3], and climate change [4]. Water vapor also influences the environment as it affects the size, composition, optical properties of aerosols [5]. For instance, water vapor is an important contributor in moist intrusions in the arctic region which result in sea ice decline and climate warming [6]–[9]. Lee et al. [6] show that the increased water vapor in the Arctic leads to increased downward infrared radiation, which is responsible for the Arctic surface air temperature trend. Luo et al. [7] find that the water vapor in Barents and Kara Seas (BKS) plays a major role in the BKS warming and sea ice reduction. Yao et al. [8] proved that the water vapor content and lower tropospheric temperature control the downward IR channels, which in turn are coupled with surface air temperature. Screen et al. [9] suggest that with significant changes predicted in total column water vapor in the southern polar region, one would expect changes in the frequency of extreme cyclones.

Moreover, water vapor is an effective parameter in calculating the land surface temperature using remote sensing technique [10]. It is essential for satellite measurements that use long-wavelength signals [11]–[13].

Observation of the spatial and temporal variations of water vapor with high precision is crucial for studies of climate change and global warming [14]. As a consequence, the Global Climate Observation System (GCOS) declared the Essential Climate Variables (ECV) requirement on satellite-derived water vapor for climate observation be 5% measurement uncertainty and stability of 0.3% per decade [15]. Unfortunately, current water vapor observation techniques are usually a trade-off between accuracy, coverage, and the temporal extent [16]–[18]. For instance, the ground-based radiosonde provides a relatively long record of water vapor observation [19]. However, it only provides observation twice per day (at 0000 and 1200 UTC) or once daily, and the general inhomogeneity of radiosonde sensor types may introduce uncertainties in long-term climate trend retrieval [20]. Global Navigation Satellite Systems (GNSS) / Global Positioning Systems (GPS) [11] operate in continuous and almost all-weather conditions with root-mean-squares error (RMSE) of 1~2 mm [21]. The sun photometer observes water vapor through radiation attenuation with RMSE of 2.53 mm [22]. However, these ground-based observation methods do not

The support from the Key Program of the National Natural Science Foundation of China (No. 41730109) is acknowledged. The supports by the Hong Kong Research Grants Council (RGC) (No. 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 and 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 and the Research Institute for Sustainable Urban Development, Hong Kong Polytechnic University, Hung Hum, Hong Kong (e-mail: lszzliu@polyu.edu.hk).

provide data either over the oceans or the unreachable land areas, thus they have limited spatial coverage.

The space-based remote sensing observation is an effective way to observe water vapor at a global scale but with large uncertainty compared to ground-based approach [17], [18], [22]. Infrared (IR) observations could provide measurement in both daytime and nighttime. For example, the MOderate Resolution Imaging Spectroradiometer (MODIS) observes water vapor in the IR channels with RMSE of 6.02 mm during daytime and 5.81 mm during nighttime compared against radiosonde [23]. IR water vapor product from Along-Track Scanning Radiometer (ATSR) using Advanced Infra-Red Water Vapor Estimator (AIRWAVE) provide observation with RMSE of 4.69 mm against Special Sensor Microwave/Imager (SSM/I) and 6.13 mm against Analyzed Radio Sounding Archive (ARSA) [24]. The microwave (MW) is less affected by the cloud and can provide more information than traditional IR and Near-Infrared (NIR) retrieval schemes [25]. Observation from Advanced Microwave Scanning Radiometer – Earth Observing System (AMSR-E) of Aqua satellite has an RMSE around 3 mm [26], while the RMSE for AMSR-2 water vapor is 4.7 mm [25]. Water vapor estimated from combined radiances of IR and MW of Atmospheric Infrared Sounder (AIRS) showed a 5% error and a mean bias less than 2 mm [27]. However, the spatial resolution for IR and MW observation are coarser than NIR observation. NIR water vapor observations are more sensitive in the boundary layer where most of the atmospheric water vapor resides [28]. It provides water vapor product with spatial resolution of 1 ~ 5 km. The RMSE for MODIS NIR water vapor product is 5.48 mm against GPS observations [29]. The Medium Resolution Imaging Spectrometer (MERIS) of Envisat provides a NIR water vapor product with an RMSE of 1.22 mm against GPS [30]. The RMSE of the NIR water vapor product from POLarization and Directionality of the Earth's Reflectances (POLDER) onboard of Advanced Earth Observing Satellite (ADEOS) is 3.1 mm [31]. Compared to other water vapor data such as those derived from GPS, further improvement of the water vapor accuracy from NIR observations is desired.

The MODIS is the first space instrument that observes water vapor through both IR and NIR channels, covering spectral range between 0.4 and 14.4 µm. Five of the bands in the NIR region between 0.8 and 1.3 µm are used for water vapor retrieval with a nadir spatial resolution of 1 km [28]. The band 2 (centered at 865 nm) and band 5 (centered at 1,240 nm) are window channels, which are hardly affected by water vapor. The band 17 (centered at 905 nm), band 18 (centered at 936 nm) and band 19 (centered at 940 nm) are water vapor absorption channels, as shown in Table 1.

MODIS onboard the Terra satellite started to provide operational water vapor products (MOD05) with global coverage since 1999 and the other MODIS sensor onboard the Aqua satellite started to produce a similar product (MYD05) since 2002. Validation reveals that the conventional algorithm used by the MODIS satellites overestimates the water vapor and an uncertainty of around 20% is observed in the MOD05 and MYD05 products [29], [32], which cannot meet the requirement (5%) for climate research. Various algorithms to improve water vapor retrieval accuracy from MODIS NIR channels have been proposed by using forward model [32]–[35] and artificial neural network [36], [37]. Localized implementation was conducted through optimizing the coefficient of transmittance [38].

In this research, we adopted the regression fitting method and considered the land surface types while relating the water vapor to transmittance. This empirically determined approach uses a large number of high accuracy GPS water vapor observations to fit the real land surface scenario. This method is better than the conventional model-based simulations because the consideration of surface types reduces the uncertainty caused by the inhomogeneous surface characteristics.

II. DATA

Five sets of MODIS products collected during the period from 2003 to 2017 are utilized in this work, including surface reflectance observations (MOD02/MYD02), geolocation data (MOD03/MYD03), cloud mask product (MOD35/MYD35), level 2 NIR water vapor product (MOD05/MYD05) for comparative analysis, and level 3 land cover type product (MCD12Q1) for surface classification. Their characteristics are presented in Table 2. In addition, the description of GPS estimated water vapor data collected from 464 GPS sites in the SuomiNet GPS network is also included.

Water vapor observation using ground-based GPS has been widely recognized as the reference data for validation purpose because of its high precision [11], [18], [27]. SuomiNet GPS data (http://www.suominet.ucar.edu/data.html) are closely related to the U.S. Weather Research Program (USWRP), the international Global Energy and Water Cycle Experiment (GEWEX) and the National Space Weather Program (NSWP). It provides continuous, all-weather, real-time observations on atmospheric water vapor with absolute error less than 2 mm [21]. Therefore, the hourly precipitable water vapor (PWV) data from the Continental United States (CONUS) sites within the SuomiNet GPS network were employed in this research. To reduce errors caused by the temporal discrepancies between GPS and remote sensing satellites, only water vapor data with a time difference less than 30 minutes were considered.

The land cover type product (MCD12Q1) is used to flag the surface characteristics of each GPS station (https://doi.org/10.5067/MODIS/MCD12Q1.006). This product is a combined land cover result from MODIS Terra and Aqua satellites. It provides global maps of land cover at 500-meter spatial resolution every year for six different land cover legends. The maps are created from MODIS Terra and Aqua reflectance using the supervised decision-tree classification method. Additional post-processing with prior knowledge and ancillary information were also used to further refine the classification results [39], [40]. We used the International Geosphere-Biosphere Programme (IGBP) legend, which was classified using the decision tree algorithm that ingested a full year of 8-day MODIS Nadir BRDF-Adjusted Reflectance. This scheme identifies 17 classes of land cover, including 11 natural vegetation classes, 3 human-altered classes, and 3 non-vegetated classes. Based on this classification, we have classified the land surfaces of the GPS stations into 14 types, with 9 natural vegetation classes, 3 human-altered classes, and

2 non-vegetated classes. The distribution map of the GPS stations used in this research is shown in Figure 1.

As the regression model is highly data-dependent, the selection of training data is crucial for the performance of model output. The NIR wavelength cannot penetrate cloud [41], [42]. Therefore, cloud mask data (MOD35/MYD35) are used as a quality control flag. Pixels flagged as confident clear are retained for further analysis. Additionally, since the change of surface type can interfere with the modeling, it is reasonable to use GPS stations with consistent surface types for the observation period (2003~2017) in the training process. Moreover, the bootstrap method is used to divide the collocated data pairs into independent training and testing subsets [43]. To be specific, the data pairs are firstly classified into six categories based on the surface types (Table 4) of the GPS stations. Data pairs from each category are then divided into independent training (around 70%) and testing subset using the bootstrap method.

III. METHODOLOGY FOR MODIS NIR WATER VAPOR RETRIEVAL

A. Physics Background

Based on the theory of molecular physics, the reflectance of the earth atmosphere is affected by the aero-physical characteristics of the molecules, asymmetrical molecules like H_2O could affect the transmission of solar radiation [44], [45]. Therefore, the water vapor amount can be derived as it has various absorption features in the solar and terrestrial spectrum. The spectral range of MODIS NIR bands is suited for daytime, cloud-free retrieval of water vapor.

As shown in Figure 2, the solar radiation between 860 nm and 1,200 nm on the sun-surface-sensor path is subjected to water vapor absorption [46], [47]. The radiance at the sensor [45], [48] can be written as:

$$L_{sensor}(\lambda) = L_{Path}(\lambda) + [\cos(\theta_s)\, E_0(\lambda)/\pi] T(\lambda)\rho(\lambda) \quad (1)$$

where $L_{sensor}(\lambda)$ represents the radiance at the sensor; λ is wavelength; $L_{Path}(\lambda)$ is the path scattered radiance; θ_s is the solar zenith angle; $E_0(\lambda)$ is the extra-terrestrial solar flux; $T(\lambda)$ is the total atmospheric transmittance; and $\rho(\lambda)$ is the surface reflectance.

If the scattering process from the photon path is neglected, the transmittance T can be written as:

$$T = exp(-\tau/\mu) \quad (2)$$

where τ is the optical depth and air mass $\mu = 1/\cos(\theta_s)$, following the Beer-Lambert law [32]. The regression coefficients are derived by inverting results of the radiative transfer simulations and dependent on observation geometry.

B. Model Development

Atmospheric water vapor observation from remote sensing satellite is based on its relationship to the transmission in the spectral channel. Therefore, the most crucial step in the retrieval algorithm is to accurately describe this relationship. In this research, the regression fitting using least square curve fitting method is employed and the GPS water vapor are used as ground truth.

1) Land cover classification

The earth's surface varies from location to location, with a variety of different surface types such as vegetation, water body, and soil. Water vapor observation over land is rather challenging because of the high heterogeneity of the surface characteristics. Before developing a new and more accurate water vapor retrieval model for MODIS NIR channels, the current water vapor MODIS product is examined against collocated GPS water vapor observations. A total of 81,374 pairs of data points from MODIS/Terra and 75,470 pairs from MODIS/Aqua are obtained in this research. The surface types of GPS stations used for the estimation of water vapor data, which are treated as reference value in the evaluation, are classified into 14 categories according to the MCD12Q1 IGBP legend. The validation results are listed in Table 3. It shows that the current MODIS NIR water vapor product tends to overestimate the water vapor value in most occasions, as the 'slope' for most of the data is larger than 1, and the 'offset' is a positive value in most cases. The RMSE of MODIS NIR water vapor product compared to GPS-estimated water vapor varies from 2.573 mm to 5.726 mm for Terra satellite, and from 2.472 mm to 6.379 mm for Aqua satellite. It clearly shows that the accuracy of the current MODIS NIR water vapor product is affected by their surface types. Thus the surface types should be considered in the process of developing a water vapor retrieval model.

A diagram of the nested classifications of land cover types of MCD12Q1 is displayed in Figure 3. The hierarchical nature of the classification allows us to create new legends as there are overlaps in some definitions. In this research, we re-classified the surface types into six major categories based on the similarities in the surface hydrology and land use classification. They are (1) Urban and Built-up Lands; (2) Water Bodies; (3) Barren; (4) Forests (including Evergreen Needleleaf Forests, Delicious Broadleaf Forests, Mixed Forests); (5) Shrublands (including Closed Shrublands, Open Shrublands, Woody Savannas, Savannas), and (6) Meadows (including Grasslands, Croplands, Cropland/Natural Vegetation, Permanent Wetlands).

The surface types of the GPS stations used in this research are classified into six categories. For each category, the collocated data points are further divided into independent training and testing datasets using bootstrap resampling method. To reduce the random sampling error and consider the quantity and variability of the subsets, about 70% of the data in each subset are used as training data in the model development procedure and the rest are used as testing data in the validation procedure. The details of the classification on the collocated datapoints of GPS and MODIS observations are listed in Table 4.

2) Transmittance

Solar radiation between 860 nm and 1,240 nm on the optical path (sun-surface-sensor) is subjected to atmospheric water vapor absorption, atmospheric aerosol scattering, and surface reflectance [28], [48]. Observation of transmittance is one of the most important steps in water vapor retrieval. However, water vapor transmittance cannot be observed directly. Instead, it is calculated from the ratio of surface reflectance between two or three channels using a differential absorption technique [48]. Because for most surface types the reflectance varies linearly with wavelength, the ratio will partially eliminate the impact of surface reflectance and is approximately equal to the atmospheric water vapor transmittance [28].

For homogeneous surfaces, the two-channel ratio method, which is defined in Eq. (3) as the ratio of the absorption channel to one window channel, is employed to calculate the transmittance. The transmittances of the three absorption channels (band 17, band 18, and band 19) in MODIS are written as:

$$T_i \cong \frac{L_i}{L_2} \tag{3}$$

where T_i is the transmittance of band i (i = 17, 18 and 19). The L_i is the reflectance of band i and L_2 is reflectance of window channel centered at 865 nm.

For complex land surfaces with variable reflectance spectrum, more window channels are required to estimate the transmittance in the water vapor absorption channel. Hence, three-channel ratio method is employed to calculate the transmittance:

$$T_i \cong \frac{L_i}{(C_1 L_2 + C_2 L_5)} \tag{4}$$

where the coefficients C_1 and C_2 are prescribed as 0.8 and 0.2, respectively [42]. Here L_2 and L_5 (centered at 1,240 nm) are window channels. We assume that the transmittance around 1 μm remains the same or it varies linearly [48].

3) Regression fitting

The key step in this retrieval algorithm is to accurately model the relationship between water vapor concentration and transmittance in the absorption channels. In the model development procedure, the high accuracy water vapor data obtained from GPS observations are used as the ground truth while the transmittance are calculated from the collocated MODIS absorption channels. For the absorption channels, the total atmospheric transmission decreases with the increase of water vapor on the sun-surface-sensor path [32], [48]. As displayed in Figure 4, band 18 is the strongest absorption band, with the largest decrease in the transmittance as the water vapor content increases. Band 17 is the weakest absorption band with the least variation in transmittance. Band 19 has a moderate transmittance. The regression model is based on the assumption that an exponential relation between the water vapor and transmittance exists. It reflects Lambert's law for an idealized non-scattering atmosphere, unsaturated absorption, and monochromatic radiation [33]. It is worth mentioning that several outliers have been observed at three bands in Figure 4. It might be caused by mixed pixels, clouds, impact of hazy conditions, or being observed over dark surfaces [48]. To reduce the model error caused by these outliers, points with distance to the model larger than three standard deviations have been excluded in the model training.

This exponential relationship for each channel could be perfectly described from an empirical correlation term written as:

$$T_i = a \exp\,(b\, W_i^*) + c \exp\,(d\, W_i^*) \tag{5}$$

where T_i is the transmittance from channel i; a, b, c and d are the coefficients determined from regression fitting; W_i^* is the water vapor content over its optical slant path.

The training data of MODIS reflectance are observed at various solar zenith angle θ_0 and view zenith angle θ, and the majority of the training data points are observed from off-nadir view due to the limitation in the location of GPS stations [48]. Therefore, the total column water vapor (W_i) in the vertical direction for channel i can be written as, considering the solar and observation geometry:

$$W_i^* = W_i\left(\frac{1}{\cos\theta} + \frac{1}{\cos\theta_0}\right) \tag{6}$$

The sensitivity of each absorption channel is different depending on the condition of water vapor. Band 18 is sensitive to water vapor variation under low humidity conditions, while band 17 is more sensitive to water vapor under more humid condition. Therefore, the weighted water vapor value from the three bands is expected to have a better performance than the observation from a single absorption channel:

$$W = f_1 W_{17} + f_2 W_{18} + f_3 W_{19} \tag{7}$$

where W_{17}, W_{18} and W_{19} are water vapor calculated from band 17, band 18, and band 19, respectively. The f_1, f_2 and f_3 are normalized corresponding weighting parameters, calculating from:

$$f_i = \frac{\eta_i}{\eta_1 + \eta_2 + \eta_3} \tag{8}$$

where η_i (i=1, 2, and 3) represents the sensitivity of transmittance in the absorption band i. And η_i is defined as:

$$\eta_i = \left|\frac{dT_w}{dW_i}\right| \tag{9}$$

An example of the regression fitting results is displayed in Figure 5. Figure 5(a) shows the transmittance of two-channel ratio (absorption channel / window channel) as a function of total column water vapor in the optical sun-surface-sensor path. Figure 5(b) shows the dependence of f_i on the total column water vapor for the three channels.

IV. Validation and Discussion

The above algorithm has been applied to independent testing dataset for validation purpose. Water vapor data calculated from each absorption channel and the weighted water vapor value of the three channels are discussed. The performance of these empirically determined datasets are evaluated by high precision GPS observation in terms of coefficient of determination (R^2), mean bias (MB), and root mean squares error (RMSE). The metrics are defined as:

$$R^2 = \left[\frac{\sum_{i=1}^{n}\left(PWV_{R_i} - \overline{PWV}_R\right)\left(PWV_{O_i} - \overline{PWV}_O\right)}{\sqrt{\sum_{i=1}^{n}\left(PWV_{R_i} - \overline{PWV}_R\right)^2\left(PWV_{O_i} - \overline{PWV}_O\right)^2}}\right]^2 \tag{10}$$

$$MB = \frac{1}{n}\sum_{i=1}^{n}\left(PWV_{R_i} - PWV_{O_i}\right) \tag{11}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}\left(PWV_{R_i} - PWV_{O_i}\right)^2}{n}} \tag{12}$$

where the PWV_{R_i} represents the independent, reference water vapor data observed from GPS; $\overline{PWV_R}$ is the mean value of GPS PWV; PWV_{O_i} is the water vapor retrieved from MODIS NIR channels.

A. Single Channel Observation

Firstly, the water vapor datasets calculated from single absorption channel using the new retrieval method are evaluated against independent GPS PWV. The validation results are summarized in Table 5. Both MB and RMSE have decreased. The retrieval accuracies from all the three absorption channels of MODIS have been improved. For MODIS/Terra satellite, band 18 has the best performance, with RMSE of 2.262 mm and 2.231 mm for datasets using two-channel and three-channel ratio transmittance, respectively. For MODIS/Aqua satellite, band 19 performs the best, with RMSE of 2.455 mm for two-channel ratio transmittance and 2.438 mm for three-channel ratio transmittance. For both Terra and Aqua satellites, the water vapor retrieved from band 17 has the worst accuracy among the three channels.

The validation results in Table 5 also suggest that in most occasions, water vapor results from each single absorption channel using three-channel ratio transmittance have better accuracy compared to the data using two-channel ratio transmittance, except the band 18 of Aqua satellite.

B. Weighted Water Vapor Observation

Previous studies showed that weighted water vapor observations from three absorption channels have better accuracy than those from single channel [41], [48]. Therefore, the weighted water vapor value of three absorption channel calculated based on their sensitivity to water vapor variation is also validated against GPS observation. The normalized frequencies of occurrence for the comparison of the MODIS operation product and the weighted water vapor value calculated from both two-channel and three channel ratio transmittance to the ground-based GPS PWV data are displayed in Figure 6.

For MODIS/Terra satellite, the RMSE is 2.229 mm for two-channel ratio transmittance and is 2.126 mm for three-channel ratio transmittance, respectively. For MODIS/Aqua satellite, the RMSE is 2.423 mm for two-channel ratio method and is 2.432 mm for three-channel ratio method. The results indicate that the weighted water vapor PWV estimated from three absorption channels using the new algorithm have better accuracy than PWV from each single absorption channel.

The results show that the newly retrieved datasets have greatly reduced the wet bias that previously exists in the MODIS operational products. Compared to MODIS/Terra operational products, the RMSE reduction rate is 50.78% using two-channel ratio transmittance method and is 53.06% using three-channel ratio transmittance method. Compared to MODIS/Aqua operational products, the RMSE has been reduced by 45.54% and 45.34% using two-channel and three-channel ratio transmittance methods, respectively. It can be seen that the new algorithm performs better for data derived from MODIS/Terra than data from MODIS/Aqua.

The detailed validation results of the weighted water vapor for each one of the surface types are summarized in Table 6. The results further indicate that the water vapor data retrieved from the proposed new algorithm are robust and have improved accuracy for all types of land surface. For MODIS/Terra data, the RMSE is between 1.971 mm and 3.295 mm using two-channel ratio transmittance, and is between 1.935 mm and 3.748 mm using three-channel ratio transmittance. For MODIS/Aqua data, the RMSE is between 2.197 mm and 3.485 mm using two-channel ratio transmittance, and is between 2.243 mm and 4.021 mm using three-channel ratio transmittance.

The retrieval over shrubland has the best accuracy among all surface types using the new algorithm. For MODIS/Terra data, the RMSE has been reduced to 1.971 mm and 1.935 mm using two-channel and three-channel ratio transmittance method, respectively. For MODIS/Aqua, the RMSE has been reduced to 2.197 mm and 2.243 mm using two-channel and three-channel ratio transmittance method, respectively.

The retrieval over barren has the largest RMSE reduction rate for both MODIS/Terra and MODIS/Aqua satellites. For MODIS/Terra data, the RMSE has been reduced by 64.13% (from 5.843 mm to 2.096 mm) and 63.37% (from 5.843 mm to 2.140 mm) using two-channel and three-channel ratio transmittance, respectively. For MODIS/Aqua data, the RMSE has been reduced by 61.55% (from 5.904 mm to 2.270 mm) and 59.55% (from 5.904 mm to 2.388 mm) using two-channel and three-channel ratio transmittance, respectively.

It is also worth mentioning that the performance of water vapor retrieval over water bodies remain problematic. Only limited RMSE reduction is observed using the two-channel ratio transmittance method. When using the three-channel ratio method, the RMSE values are even getting larger. This is probably because the water bodies sometimes act as black surfaces that affect the signal of absorption channels. Therefore larger uncertainty in atmospheric scattering over water bodies is resulted [49].

V. Conclusion

Atmospheric water vapor can be retrieved from remote sensing satellites with the observation of transmittance in the NIR channels. The conventional water vapor retrieval algorithm uses radiative transfer model to simulate the relationship with simplified assumptions and pre-calculated atmospheric information. A systematic overestimation of water vapor value has been observed when compared to high accuracy ground-based water vapor measurements.

A new empirical regression algorithm is proposed in this research taking the land surface type into consideration. The new algorithm takes advantage of the high precision water vapor data estimated from GPS to derive empirical regression function. A large amount of GPS water vapor data collected over a period of 15 years from the western North America are employed in the model training process. The MODIS land cover product (MCD12Q1, IGBP Legend) is used for land surface type classification.

Retrievals from single absorption channel and the weighted water vapor data from three absorption channels have been discussed for MODIS both Terra and Aqua satellites using the new algorithm. For water vapor retrieved from single absorption channel, the data calculated using three-channel ratio transmittance method performs better than using two-channel ratio transmittance in most scenarios. For the MODIS/Terra satellite, the channel 18 performs the best among three absorption channels. Its RMSE is 2.262 mm and 2.231 mm using two-channel and three-channel ratio transmittance method, respectively. For the MODIS/Aqua satellite, the channel 19 performs the best with RMSE of 2.455 mm and 2.438 mm using the two-channel and three-channel ratio transmittance method, respectively. On the other hand, channel 17 has the worst accuracy among the three channels for both satellites.

The weighted water vapor, which is a weighted combination of the water vapor calculated from three absorption channels, further improves the retrieval accuracy. For MODIS/Terra satellite, the RMSE has been reduced by 50.78% to 2.229 mm using two-channel ratio transmittance method and been reduced by 53.06% to 2.126 mm using three-channel ratio transmittance method. For MODIS/Aqua satellite, the RMSE has been reduced by 45.54% to 2.423 mm using two-channel ratio transmittance method and been reduced by 45.34% to 2.432 mm using three-channel ratio transmittance method.

In terms of land surface type, the water vapor data retrieved from shrublands have the highest accuracy among all surface types using the new algorithm. For MODIS/Terra, the RMSE is 1.971 mm and 1.935 mm for data using two-channel and three-channel ratio transmittance method, respectively. For MODIS/Aqua, the RMSE is 2.197 mm and 2.243 mm for data using two-channel and three-channel ratio transmittance method, respectively.

In summary, the new algorithm proposed in this work has significantly improved the water vapor retrieval accuracy of MODIS NIR channels under cloud-free conditions. It reduces the wet bias for most occasions. However, water vapor retrieval over water bodies remains problematic because of the larger uncertainties of atmospheric scattering. Further investigation for this surface type is needed.

Acknowledgment

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, water vapor data and land cover product. 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.

References

[1] 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, doi: 10.1007/s00382-005-0017-4.

[2] A. Raval and V. Ramanathan, “Observational determination of the greenhouse effect,” *Nature*, vol. 342, no. 6251, pp. 758–761, Dec. 1989, doi: 10.1038/342758a0.

[3] 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, doi: 10.1126/science.296.5568.727.

[4] T. R. Karl and K. E. Trenberth, “Modern Global Climate Change,” *Science*, vol. 302, no. 5651, pp. 1719–1723, Dec. 2003, doi: 10.1126/science.1090228.

[5] N. Kämpfer, *Monitoring Atmospheric Water Vapour: Ground-Based Remote Sensing and In-situ Methods*. Springer Science & Business Media, 2012.

[6] S. Lee, T. Gong, S. B. Feldstein, J. A. Screen, and I. Simmonds, “Revisiting the Cause of the 1989–2009 Arctic Surface Warming Using the Surface Energy Budget: Downward Infrared Radiation Dominates the Surface Fluxes,” *Geophys. Res. Lett.*, vol. 44, no. 20, p. 10,654-10,661, 2017, doi: 10.1002/2017GL075375.

[7] B. Luo, D. Luo, L. Wu, L. Zhong, and I. Simmonds, “Atmospheric circulation patterns which promote winter Arctic sea ice decline,” *Environ. Res. Lett.*, vol. 12, no. 5, p. 054017, May 2017, doi: 10.1088/1748-9326/aa69d0.

[8] Y. Yao, D. Luo, A. Dai, and I. Simmonds, “Increased Quasi Stationarity and Persistence of Winter Ural Blocking and Eurasian Extreme Cold Events in Response to Arctic Warming. Part I: Insights from Observational Analyses,” *J. Clim.*, vol. 30, no. 10, pp. 3549–3568, Jan. 2017, doi: 10.1175/JCLI-D-16-0261.1.

[9] J. A. Screen, T. J. Bracegirdle, and I. Simmonds, “Polar Climate Change as Manifest in Atmospheric Circulation,” *Curr. Clim. Change Rep.*, vol. 4, no. 4, pp. 383–395, Dec. 2018, doi: 10.1007/s40641-018-0111-4.

[10] Z.-L. Li *et al.*, “Satellite-derived land surface temperature: Current status and perspectives,” *Remote Sens. Environ.*, vol. 131, pp. 14–37, Apr. 2013, doi: 10.1016/j.rse.2012.12.008.

[11] M. Bevis, S. Businger, T. A. Herring, C. Rocken, R. A. Anthes, and R. H. Ware, “GPS meteorology: Remote sensing of atmospheric water vapor using the global positioning system,” *J. Geophys. Res. Atmospheres*, vol. 97, no. D14, pp. 15787–15801, Oct. 1992, doi: 10.1029/92JD01517.

[12] 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, doi: 10.1126/science.283.5406.1297.

[13] A. E. Niell *et al.*, “Comparison of Measurements of Atmospheric Wet Delay by Radiosonde, Water Vapor Radiometer, GPS, and VLBI,” *J. Atmospheric Ocean. Technol.*, vol. 18, no. 6, pp. 830–850, Jun. 2001, doi: 10.1175/1520-0426(2001)018<0830:COMOAW>2.0.CO;2.

[14] 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, doi: 10.1146/annurev.energy.25.1.441.

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

[16] 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, doi: 10.1029/2003JD003372.
[17] 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, doi: 10.1002/joc.3412.
[18] 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, doi: 10.1016/j.rse.2017.09.028.
[19] I. Durre, R. S. Vose, and D. B. Wuertz, “Overview of the Integrated Global Radiosonde Archive,” *J. Clim.*, vol. 19, no. 1, pp. 53–68, Jan. 2006, doi: 10.1175/JCLI3594.1.
[20] A. Rinke *et al.*, “Trends of Vertically Integrated Water Vapor over the Arctic during 1979–2016: Consistent Moistening All Over?,” *J. Clim.*, vol. 32, no. 18, pp. 6097–6116, Jul. 2019, doi: 10.1175/JCLI-D-19-0092.1.
[21] 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, Apr. 2000, doi: 10.1175/1520-0477(2000)081<0677:SARNGN>2.3.CO;2.
[22] 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, doi: 10.1016/j.atmosres.2017.07.021.
[23] 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, doi: 10.1080/01431161.2014.999884.
[24] 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, doi: 10.1016/j.rse.2015.10.037.
[25] J. Du, J. S. Kimball, and L. A. Jones, “Satellite Microwave Retrieval of Total Precipitable Water Vapor and Surface Air Temperature Over Land From AMSR2,” *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 5, pp. 2520–2531, May 2015, doi: 10.1109/TGRS.2014.2361344.
[26] 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, doi: 10.1109/TGRS.2011.2160184.
[27] 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, doi: 10.1002/2016JD024806.
[28] 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.
[29] J. He and Z. Liu, “Comparison of Satellite-Derived Precipitable Water Vapor Through Near-Infrared Remote Sensing Channels,” *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 12, pp. 10252-10262, Dec. 2019.
[30] J. Fischer and R. Bennartz, “Retrieval of Total Water Vapor Content from MERIS Measurements,” *Algorithm Theor. Basis Doc. PO-NT-MEL-GS-005*, 1997, Accessed: Jun. 15, 2017. [Online]. Available: https://earth.esa.int/documents/700255/2042855/MERIS_ATBD_2.4_v4.2+-+2011.pdf.
[31] 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.
[32] 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, doi: 10.5194/amt-8-823-2015.
[33] R. Bennartz and J. Fischer, “Retrieval of columnar water vapour over land from backscattered solar radiation using the Medium Resolution Imaging Spectrometer,” *Remote Sens. Environ.*, vol. 78, no. 3, pp. 274–283, Dec. 2001, doi: 10.1016/S0034-4257(01)00218-8.
[34] 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, doi: 10.1175/JTECH1708.1.
[35] R. Lindstrot *et al.*, “A global climatology of total columnar water vapour from SSM/I and MERIS,” *Earth Syst. Sci. Data*, vol. 6, no. 1, pp. 221–233, Jun. 2014, doi: https://doi.org/10.5194/essd-6-221-2014.
[36] K. Mao, J. Shi, Z.-L. Li, and H. Tang, “An RM-NN algorithm for retrieving land surface temperature and emissivity from EOS/MODIS data,” *J. Geophys. Res. Atmospheres*, vol. 112, no. D21, p. D21102, Nov. 2007, doi: 10.1029/2007JD008428.
[37] 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, doi: 10.3390/atmos8080139.
[38] 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, doi: 10.1109/TGRS.2017.2716414.

[39] M. A. Friedl *et al.*, “Global land cover mapping from MODIS: algorithms and early results,” *Remote Sens. Environ.*, vol. 83, no. 1, pp. 287–302, Nov. 2002, doi: 10.1016/S0034-4257(02)00078-0.

[40] M. A. Friedl *et al.*, “MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets,” *Remote Sens. Environ.*, vol. 114, no. 1, pp. 168–182, Jan. 2010, doi: 10.1016/j.rse.2009.08.016.

[41] 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, doi: 10.1080/0143116031000102502.

[42] 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. Accessed: Nov. 24, 2016. [Online]. Available: http://eospso.nasa.gov/sites/default/files/atbd/atbd_mod07.pdf.

[43] B. Efron, “Bootstrap Methods: Another Look at the Jackknife,” *Ann. Stat.*, vol. 7, no. 1, pp. 1–26, Jan. 1979, doi: 10.1214/aos/1176344552.

[44] 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, doi: 10.1364/AO.15.002085.

[45] 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, doi: 10.1109/TGRS.1985.289380.

[46] A. Berk, L. S. Bernstein, and D. C. Robertson, “MODTRAN: A Moderate Resolution Model for LOWTRAN,” Spectral Sciences. Inc., Burlington, MA, Jul. 1987.

[47] 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*, Jun. 2014, vol. 9088, p. 90880H, doi: 10.1117/12.2050433.

[48] 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, p. 4389, 2003, doi: 10.1029/2002JD003023.

[49] M. D. King *et al.*, “Cloud and aerosol properties, precipitable water, and profiles of temperature and water vapor from MODIS,” *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 2, pp. 442–458, Feb. 2003, doi: 10.1109/TGRS.2002.808226.

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 Wuhan University, Wuhan, China, in 1997, and the Ph.D. degree in geomatics engineering from the University of Calgary, Calgary, AB, 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. He is a guest Professor with Liaoning Technical University, Huludao, China, and Wuhan University. 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.