https://doi.org/10.1016/j.rse.2021.112384

Water Vapor Retrieval from MERSI NIR Channels of Fengyun-3B Satellite Using Ground-based GPS Data

Jia He[a,b], Zhizhao Liu[a,b]

[a] Department of Land Surveying and Geo-Informatics, Hong Kong Polytechnic University, Hung Hum, Kowloon, Hong Kong

[b] Research Institute for Sustainable Urban Development (RISUD), Hong Kong Polytechnic University, Hung Hum, Kowloon, Hong Kong

Corresponding author: Zhizhao Liu (lszzliu@polyu.edu.hk)

Abstract

An ensemble-based empirical regression algorithm is for the first time developed to retrieve total column water vapor from the Medium Resolution Spectral Imager (MERSI) near-infrared (NIR) channels onboard the Fengyun-3B (FY-3B) satellite. This retrieval method uses precipitable water vapor (PWV) data estimated from ground-based Global Positioning System (GPS) data to build a regression model in which the reflectance ratio observed from MERSI NIR absorption channels and the corresponding GPS PWV data are the parameters. The MERSI Level 1b data, specifically the three water vapor absorption channels centered at 905 nm, 940 nm, and 980 nm are used to retrieve water vapor. PWV data observed from 256 ground-based GPS stations located in the western North America in 2016 are used as reference data for model development. Then, validation is performed with data obtained during 2017 ~ 2019 from both the western North America and Australia to assess the performance of the proposed algorithm. The results indicate that the new PWV results agree very well with ground-based PWV reference data. The mean absolute percentage error (MAPE) for ensemble median PWV is 16.72% ~ 36.74% in western North America and is 15.47% ~ 32.31% in Australia. The RMSE is 4.635 mm ~ 8.156 mm in western North America and is 5.383 mm ~ 8.900 mm in Australia. The weighted mean value using three-channel ratio transmittance has the best retrieval accuracy, with RMSE of 4.635 mm in western North America and 5.383 mm in Australia. This new PWV algorithm can retrieve PWV from FY-3B data with a higher accuracy for different regions. Different from conventional algorithms, no pre-observed information of atmospheric parameters is required in this model.

Keywords: MERSI, Near Infrared, Precipitable Water Vapor, GPS

1. Introduction

Precipitable water vapor (PWV) is one of the most important natural greenhouse gases in the atmosphere (Held and Soden, 2000). It plays a key role in the weather (Soden et al., 2002), climate (Karl and Trenberth, 2003), and environment (Elliott and Gaffen, 1991) locally or globally, and impact the hydrological cycle and energy exchange (Raval and Ramanathan, 1989; Sherwood et al., 2010). The variation of water vapor distribution is complex in space-time dimension (Elgered et al., 1997), as the water vapor field varies significantly in the time domain within a period as short as one hour (Elgered et al., 2005) and also in the space domain, ranging from about 5 cm near the equator to less than 1 mm at the poles (Mockler, 1995). Therefore, water vapor observation with high spatial-temporal resolution is critical for climate and environmental research (Belward, 2016). Water vapor measurement has benefited from the development of remote sensing techniques, such as instrument improvement with better resolution, computational advances in storage, and processing capabilities (Hanssen et al., 1999; Lindenbergh et al., 2008; Levin et al., 2014). Several operational products of water vapor were retrieved using radiative transfer models (Gao and Kaufman, 2003; Hu et al., 2011).

Validations against radiosonde show that the MOD05 data overestimated PWV with a scale factor from 1.14 to 1.20, while it overestimated by 7%–14% after comparison against GPS (Li et al., 2003). Inter-comparisons among multisource water vapor products show an underestimation of PWV by 10% to 30% for MERSI/FY-3C (Shi et al., 2018). Multi-sensory monitoring on water vapor makes it possible to continuously observe water vapor distribution with high spatial and temporal resolutions.

The Chinese meteorological satellite system is established in 1988, with the launch of its first Fengyun (FY) satellite. The FY system is composed of both sun-synchronous satellites and geostationary satellites. Specifically, the FY-3 satellite series, as the second generation of the polar-orbiting meteorological satellites, aims to provide global air temperature, humidity profiles, and meteorological parameters for scientific research in climate change, climate diagnosis, and predictions (Dong et al., 2009; Yang et al., 2012). Four satellites in this series have been successfully launched into orbit, equipped with imaging and sounding instruments. Medium Resolution Spectral Imager (MERSI) onboard the FY-3 series, is a MODIS-like sensor with 20 bands in both visible and NIR channels with resolutions from 250 to 1000 m. Similar to MODIS NIR channels, MERSI also has 5 NIR water vapor related channels, including two window channels centered at 865 nm and 1030 nm, and three water vapor absorption channels centered at 905 nm, 940 nm, and 980 nm (Wang et al., 2012). Therefore, the MERSI NIR bands are suitable for water vapor observation over land in the daytime, under cloud-free conditions.

Water vapor retrieval from NIR channels is based on the relationship between the transmittance of NIR channels and the amount of water vapor. Water vapor obtained from MERSI NIR is conventionally calculated using the simulation of radiative transfer models (Hu et al., 2011; Kaufman and Gao, 1992). Transmittance observed from MERSI NIR channels is converted into column water vapor with the pre-calculated look-up table. The look-up table is computed using atmospheric transmittance code MODTRAN. The transmittance variations in the absorption bands is affected by the efficiency of radiative transfer model. Hence large uncertainties in water vapor estimation are expected (Warner and Ellingson, 2000). Evaluations of the MERSI/FY-3A PWV product over the northwest China against water vapor derived from the Global Positioning System (GPS) show that the mean absolute percentage error (MAPE) is 22.83% (Gong et al., 2018b). Another validation on MERSI/FY-3A over the East Asia continent shows that the MAPE varies from 31.8% to 44.1% (Gong et al., 2018a). Validation analysis for MERSI/FY-3C in China suggests that the MERSI underestimates PWV by 10% to 30% when compared to ground-based observations, with a mean bias of -4.68 mm (Shi et al., 2018). In general, water vapor obtained from MERSI onboard of FY-3 satellites shows larger errors compared to other operational NIR water vapor products. For instance, our previous study showed that the RMSE of water vapor obtained from MERSI/FY-3A over the western North America region was 8.644 mm while the RMSE of water vapor product from the Moderate Resolution Imaging Spectroradiometer (MODIS) was 5.480 mm (He and Liu, 2019).

The reason to retrieve water vapor from MERSI onboard of FY-3B in this study is twofold. First, previous studies showed that operational MERSI water vapor data obtained from FY-3A and FY-3C satellite missions had relatively poor accuracy. The retrieval errors are mainly from the miscalculation of transmittance variance from the radiative transfer model (Warner and Ellingson, 2000) and the uncertainties of atmospheric condition variation. Therefore, development of a new algorithm to improve water vapor retrieval from the MERSI sensor is needed. The empirical regression algorithm developed in this work shows that it can significantly reduce the bias of water vapor products. Secondly, currently no operational water vapor product has been published by the FY-3B satellite series administrator, which greatly constrains applications in the user community. With the use of the algorithm in this work, water vapor data can be retrieved from FY-3B data published online, which will benefit the FY-3 series user community greatly.

The objectives of this study are: (1) to develop a new water vapor retrieval algorithm to retrieve water vapor from MERSI/FY-3B Level 1b data with a high accuracy. A large amount of collocated PWV data from GPS and transmittance data obtained from MERSI/FY-3B NIR water vapor absorption channels were employed to establish an ensemble-based regression model for the western North America region; (2) to validate the ensemble-based regression model using temporally and spatially independent reference PWV data. GPS PWV data observed in different time period from the western North America region and in Australia were employed to validate the performance of the new algorithm on a global scale; (3) to lay the algorithmic foundation for future work of retrieving and producing PWV data from the MERSI/FY-3B sensor for global uses.

2 Dataset and Preprocessing

2.1 Dataset

Two types of data are included in this research, namely the MERSI/FY-3B reflectance data for water vapor retrieval and the GPS PWV data as the ground truth reference.

2.1.1 MERSI L1B Data

FY-3B satellite was successfully launched into a sun-synchronous polar orbit on November 4, 2010, providing global coverage of earth surface observation every day. The MERSI onboard the FY-3B satellite has 20 channels in visible and infrared wavelength ranges. It has five NIR channels dedicated to water vapor observation, including three water vapor absorption channels centered at 905 nm, 940 nm, and 980 nm in the shortwave region and two window channels centered at 865 nm and 1030 nm. A summary of the MERSI/FY-3B channels is listed in Table 1 .

Table 1 Summary of the spectral location and band information of MERSI/FY-3B for water vapor retrieval.

Band Number	Center	FWHM	Spatial Resolution (m)	Description
16	865 nm	20 nm	1,000	Window Channel
17	905 nm	20 nm	1,000	Absorption Channel
18	940 nm	20 nm	1,000	Absorption Channel
19	980 nm	20 nm	1,000	Absorption Channel
20	1,030 nm	20 nm	1,000	Window Channel

2.1.2 GPS Data

(1) GPS PWV from SuomiNet

The hourly PWV data from 256 stations in the SuomiNet GPS network (http://www.suominet.ucar.edu/data.html) CONUS sites are employed as ground truth of water vapor for model training and testing in this research. The GPS stations are located in the western North America and they have a variety of surface types, as shown in Figure 1.

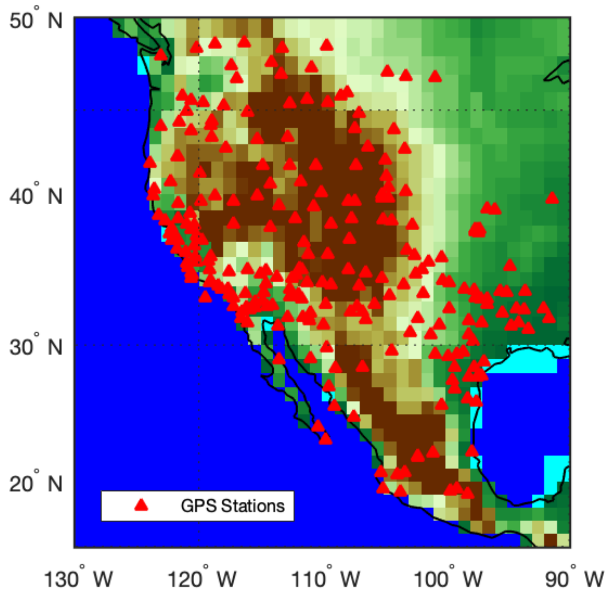


Figure 1 Distribution map of 256 GPS stations located in the western North America. They are used for FY-3B water vapor calibration and validation analysis. The color bar represents the elevation of the GPS stations, in unit of meters.

These stations are equipped with precise surveying quality, dual-frequency Trimble receivers and antennas (Ware et al., 2000), and the GPS data are processed to estimate precipitable water vapor using the BERNESE software developed at the University of Berne (Dach et al., 2015). BERNESE is a widely used software package for high-precision analysis of GPS as well as global navigation satellite system (GNSS) data. The GPS PWV data observed in 2016 are used for model development. In the model validation, GPS PWV data from 2017 to 2019 are used, in order to be independent of the data used for model development.

(2) GPS Data from Geoscience Australia

In addition, to evaluate the performance of this newly proposed retrieval algorithm outside the western North America, PWV observations obtained during 2017 to 2019 from 419 GPS stations are also used in this study for validation purpose, which are operated by the Geoscience Australia (ftp://ftp.ga.gov.au/geodesyoutgoing/gnss/products/troposphere/rapid/). The

Australian data are included in the validation process for two reasons. Firstly, Australia has a dense ground-based GPS observations. A lot of GPS-derived PWV can be used as reference. Secondly, Australia represents the south hemisphere's weather and climate. Australian climate is sensitive to El Niño – Southern Oscillation (ENSO) and La Niña events, which are strongly related to water vapor distribution (Ashcroft et al., 2016; Wang et al., 2018).

The GPS Zenith Tropospheric Delays (ZTDs) from the Australian network were also processed using the BERNESE GNSS software (Hu, 2017). These ZTD data are then converted to PWV using the surface pressure, temperature, and humidity profiles obtained from the European Centre for Medium-Range Weather Forecasts (ECMWF) ERA-5 (Hersbach et al., 2020) with millimeter accuracy (Wang et al., 2018). The distribution map of these GPS stations is presented in Figure 2.

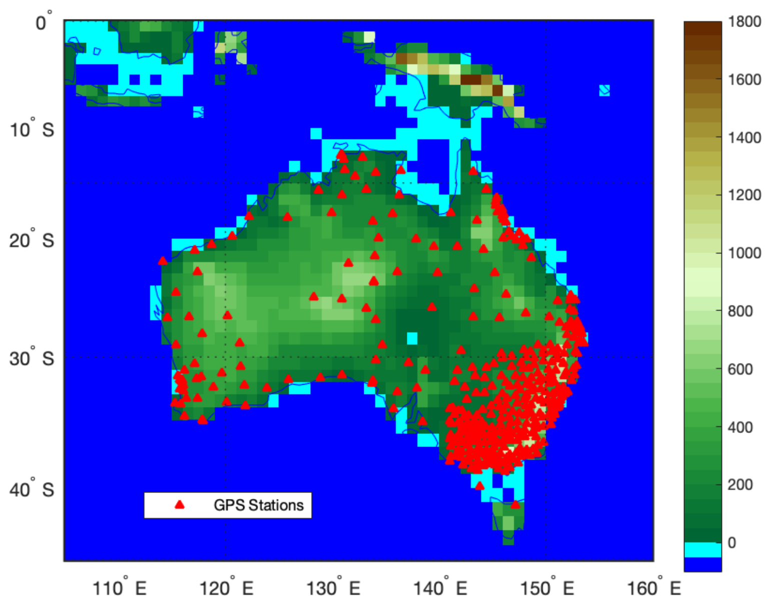


Figure 2 Distribution map of 419 GPS stations located in Australia used for FY-3B water vapor validation analysis. The color bar represents the elevation of the GPS stations, in unit of meters.

2.2 Pre-Processing

The first necessary step in the pre-processing of MERSI water vapor measurements is the screening of clouded pixels. Ideally, the cloud mask calculation algorithm should be concise and simple and provide enough information for the application effectively. A simple visible and IR window threshold approach is one of the most efficient ways to detect cloud (Ackerman et al., 1998). Because MERSI is a MODIS-like sensor, the cloud mask algorithm used for MODIS is adopted for MERSI cloud detection (Martins et al., 2002; King et al., 2003; Wind et al., 2010). The clear pixels have low radiance reflectance and high brightness temperature (Ackerman et al., 1998; Wind et al., 2010). Cloud condition is calculated at single-pixel resolution. To reduce the uncertainty caused by cloud, only confident clear pixels are further used in the following PWV retrieval training and testing procedures.

3 Theoretical Background

The measurement of the total column of atmospheric water vapor from the MERSI NIR channel is based on measuring its absorption effect on solar transmittance in the water vapor absorption channel. Based on molecular physics, the symmetric molecules, such as O_3 and other gases in the atmosphere, do not affect the transmittance in this wavelength range (Roberts et al., 1976; Fraser and Kaufman, 1985; Berk et al., 1987). As shown in Figure 3, water vapor contributes to the majority of absorption on radiance around absorption channels at the three absorption channels centered at 905 nm, 940 nm, and 980 nm. Thus the MERSI is suitable for observing total column water vapor over land in the daytime and under cloud-free conditions.

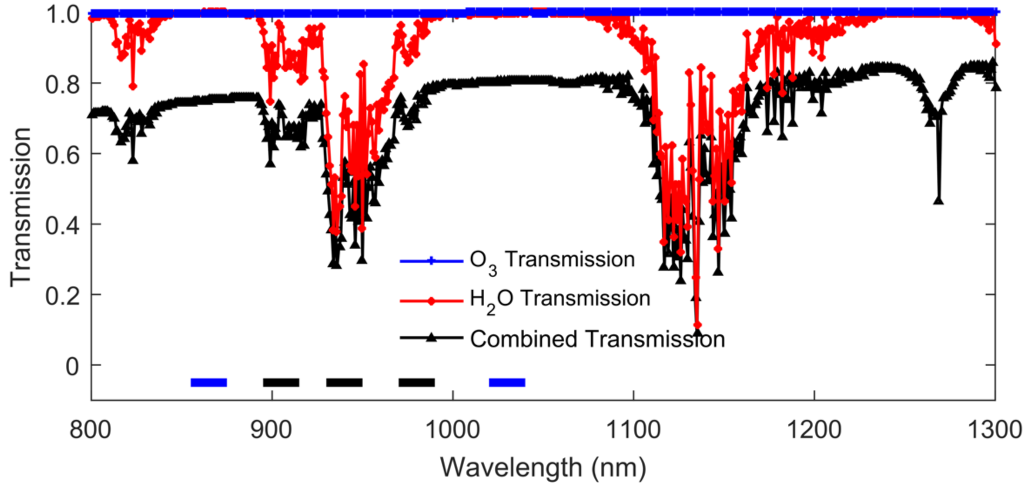


Figure 3 Spectral transmission of atmosphere contents in the presence of water vapor at 0.6 g/cm^2, considering H_2O, O_3 and the combined transmission. Computations were performed using MODTRAN 4 model. The color bars at the bottom show the location of the MERSI water vapor absorption channels (black) and the window channels (blue) used in the retrieval study.

As the transmittance cannot be measured directly, the differential absorption technique is applied while calculating transmittance from the three absorption channels (Kaufman and Gao, 1992). The technique assumes that the transmittance of the solar energy approximately equals the ratio of one absorption channel and one or two window channels (King et al., 1992), therefore it will partially eliminate the effect of surface types on the reflectance (Kaufman and Gao, 1992). The transmittance of the 2-channel ratio method is described as:

$$T_i \cong \frac{L_i}{L_{16}} \quad (1)$$

where T_i is the transmittance of band i (i=17, 18 and 19). The L_i denotes the reflectance of band i. L_{16} is reflectance in window channel 16.

For surface type with complex and mixed reflectance spectrum, additional window channel will help to eliminate the effects from surface types (Kaufman and Gao, 1992). The 3-channel ratio function, including one absorption channel and two window channels, is written as:

$$T_i \cong \frac{L_i}{(C_1 L_{16} + C_2 L_{20})} \qquad (2)$$

where the coefficients C_1 and C_2 are prescribed as 0.8 and 0.2, respectively (Seemann et al., 2006).

Then, by applying an atmospheric transmittance model, such as High-Resolution Transmission (HITRAN) (Kaufman and Gao, 1992; Rothman et al., 2009) or MODerate resolution atmospheric TRANsmission (MODTRAN) (Schläpfer et al., 1998; Berk et al., 2014), the relation between the measured radiance ratio and water vapor content could be simulated for a large variety of different atmospheric profiles (Hu et al., 2011). The relationship can be expressed by an exponential formula written as:

$$T_w = \exp(\alpha - \beta\sqrt{W^*}) \qquad (3)$$

where T_w is the transmittance of water vapor, α and β are a function of surface type, and $W^*$is the water vapor along the sun-surface-sensor path (Kaufman and Gao, 1992).

4 Methods

In this study, an ensemble-based empirical regression algorithm is proposed in order to calculate water vapor with a higher precision from MERSI NIR channels.

The retrieval scheme is illustrated in Figure 4. Firstly, the clouded pixels are defined as null (NAN) value, as the algorithm retrieves water vapor under clear conditions only. Secondly, the differential absorption method is utilized to calculate the transmittance of NIR channels for clear pixels. Subsequently, iterative optimization is performed using the bootstrap method to resample the training dataset for water vapor retrieval. The determination of the relationship between transmittance and water vapor content in three absorption channels from the MERSI band centered at 905 nm, 940 nm, and 980 nm is the most critical step. The least-squares fitting method is used to establish the relationship between PWV and transmittance in the MERSI three water vapor absorption channels.

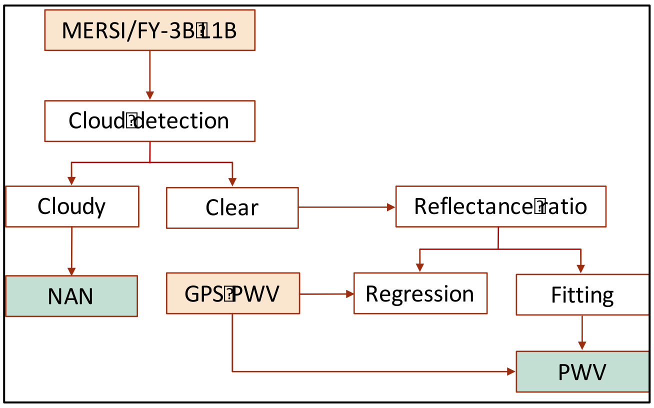


Figure 4 Flow chart of the newly proposed method to retrieve PWV from NIR channels of the MERSI sensor onboard the FY-3B satellite. PWV estimated from ground-based GPS observations are used as reference values. The light brown boxes denote the input data, and the green boxes are the output results.

4.1 Ensemble Analysis

After getting the transmittance for clear pixels, a total of 6,036 pairs of valid points collected in 2016 over the western North America under cloud-free conditions are used for model development. Because regression modeling is highly data-dependent, the training dataset with uneven distribution in time might affect the performance of the retrieval results. As shown in Figure 5, summertime has more valid data observed than wintertime. Therefore, an ensemble-based bootstrap resampling algorithm is introduced to balance the class distribution and reduce the potential effects of the training dataset (Batista et al., 2004). The approach is to divide the training dataset into several groups with slightly overlapping chunks and conduct regression fitting for all subsets concurrently (Efron, 1979; Wu, 1986). The multiple classifiers would have a better description of the relationship than a single one and would reduce bias and retrieval errors (Batista et al., 2004). In this procedure, the training data are resampled into 10 independent subsets with 4,200 data points (around 70% of the total dataset used for model development). By applying this bootstrap method, the errors introduced from random sampling and the uncertainties caused by the possible channel drifting in the channel position are expected to be reduced (He and Liu, 2020).

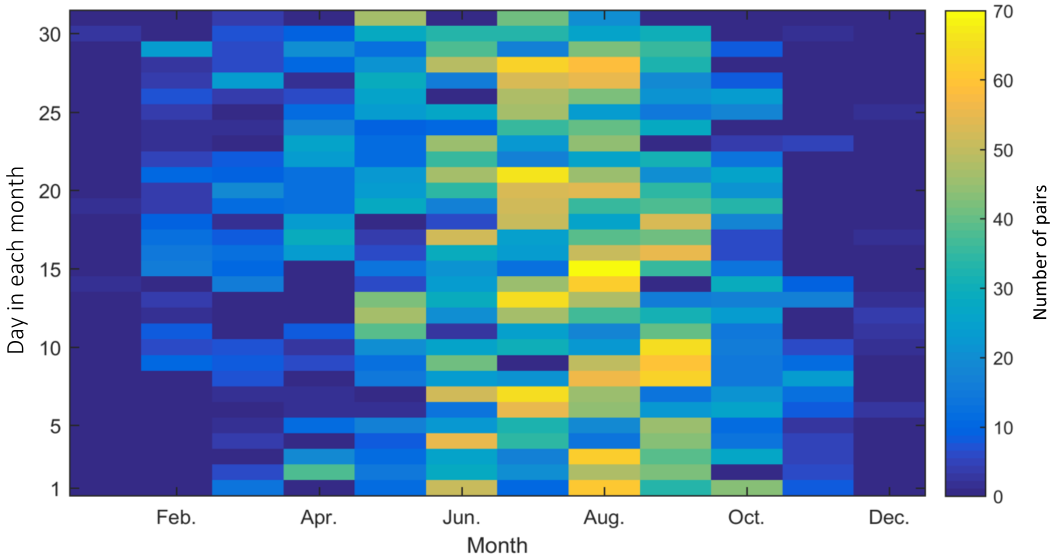


Figure 5 The number of data pairs of the collocated GPS and MERSI/FY-3B L1 NIR channel reflectance recorded in each day in each month of 2016 over the western North America under cloud free conditions. They are used for model development. The color bar denotes the number of data pairs.

4.2 Training of Regression Algorithm

A scatterplot of the relation between GPS observed water vapor and the transmittances of the three absorption channels of MERSI/FY-3B are shown in Figure 6. Band 18 centered at 940 nm is the strongest absorption channel in the NIR wavelength range, with the maximum variation in transmittance. In the remaining two bands, band 17 (centered at 905 nm) is more sensitive to water vapor variation than band 19 (centered at 980 nm).

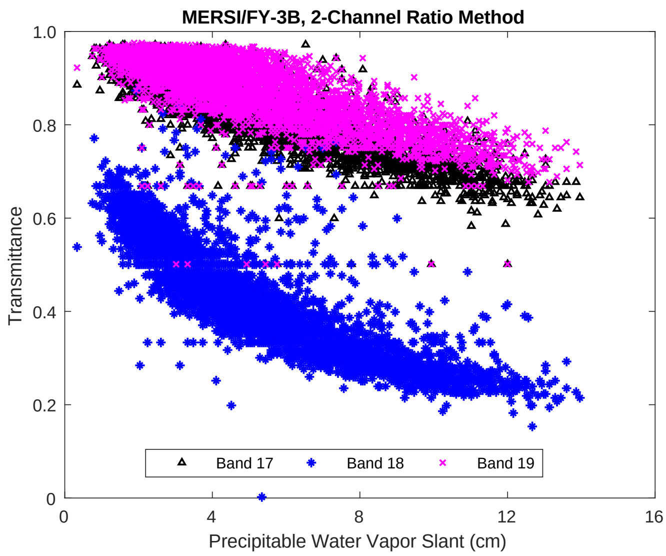


Figure 6 Example of scatterplot of the relationship between optical path (slant) column water vapor observed from GPS and the transmittance from three absorption channels of MERSI/FY-3B with 2-channel ratio method.

The selection of the exponential function is based on the examination of the numerical relationship between GPS observed PWV and the transmittance in the three absorption channels of MERSI. The least-squares fitting method is employed for each subset of data to model the relationship. After analyzing the characteristics of many types of functions, the exponential function description given below is selected to characterize the relationship:

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

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

where T_i denotes the transmittance of MERSI NIR absorption channel i (i=17, 18 and 19); W_i^* represents the water vapor content observed from channel i in the slant viewing angle; the coefficients a, b, and c are to be determined by the regression fitting; W_i is the total column water vapor in the vertical view; θ and θ_0 represent the sensor zenith angle and solar zenith angle, respectively (Gao and Kaufman, 2003). It should be noted that the development of equation (4) is stimulated by the physical model in equation (3). The two equations have a kind of similarity but they are different as shown by our results below. It is also worth mentioning that data points with a distance larger than three standard deviations of the fitting function are considered as outliers and are excluded from the training dataset.

4.3 Optimization of Channel Selection

Water vapor ensemble members can be retrieved from each absorption channel of the MERSI using the regression function. On the other hand, these absorption channels have different sensitivities under different levels of water vapor concentration. Band 18 (centered at 940 nm) is the strong absorption band, having the largest decrease in transmittance with the increase of water vapor. Band 19 (centered at 980 nm) is the weakest absorption band, having the least decrease in transmittance. Band 17 (centered at 905 nm) has a moderate sensitivity to water vapor variation. As a result, the transmittance in the three absorption channels can represent the magnitude of radiance attenuation resulting from water vapor. The weighted mean value of the three absorption channels, which is calculated based on the sensitivity of the transmittance, could obtain a more accurate retrieval of water vapor (Gao and Kaufman, 2003):

$$W = f_{17}W_{17} + f_{18}W_{18} + f_{19}W_{19} \tag{6}$$

where W_{17}, W_{18} and W_{19} are water vapor calculated from MERSI water vapor absorption bands 17, 18, and 19, respectively; the f_{17}, f_{18} and f_{19} are normalized weighting parameters corresponding to each band and it is calculated as (Gao and Kaufman, 2003):

$$f_i = \frac{\eta_i}{\eta_1 + \eta_2 + \eta_3} \tag{7}$$

The weighting factor η_i is subject to the sensitivity of transmittance in the absorption band and η_i is estimated from (Gao and Kaufman, 2003):

$$\eta_i = \left|\frac{dT_{W_i}}{dW_i}\right| \tag{8}$$

where dT_{W_i} is the transmittance variation in one unit length; dW_i is the water vapor variation in one unit length. It is computed numerically from simulated curves of transmittance versus precipitable water vapor.

One set of the ensemble members of the regression functions derived for water vapor retrieval in each absorption band and the corresponding weighting factor (f_{i}) of the three channels based on their sensitivity to transmittance are shown in Figure 7. The results confirm that band 19 is sensitive to atmospheric conditions with low water vapor content. It contributes the most to the weighted mean value when the PWV in the optical path is less than 8 cm. In the contrast, band 18 is the most sensitive to the atmospheric conditions of high water vapor concentration (higher than 8 cm).

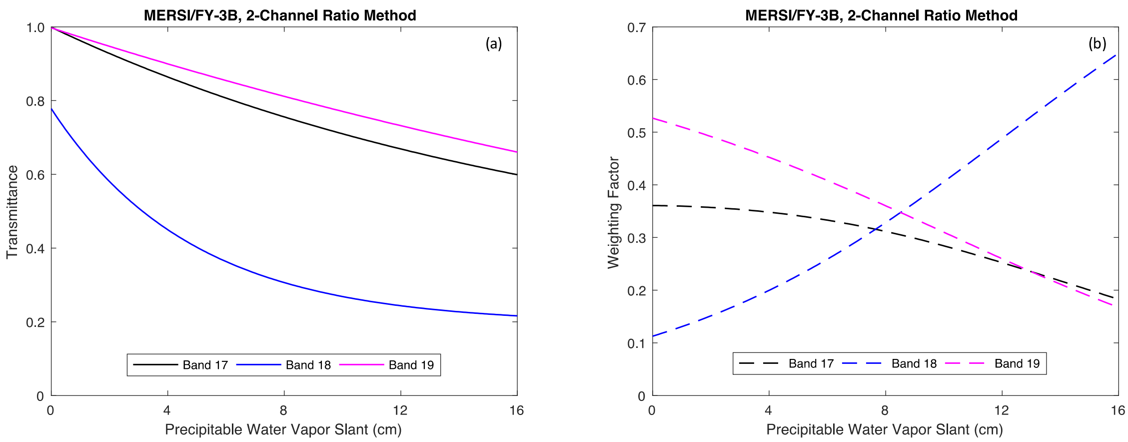


Figure 7 (a) Example of regression functions from MERSI/FY-3B using 2-channel ratio method; (b) the corresponding normalized weighting factors of the three absorption channels based on their sensitivity to transmittance.

5 Results

With the above procedures discussed, new sets of ensemble members of water vapor can be estimated from MERSI/FY-3B using the regression functions for each absorption channel. The weighted mean values could be further calculated from three channels based on their sensitivity to water vapor variation. The least-squares estimated parameters of the ensemble members are list in Table 1.

Table 1 Least-squares estimated parameters of regression function $T_i = a\,exp\,(b\,W_i^{*}) + c$ for ensemble members of MERSI/FY-3B water vapor absorption channels. The reference PWV data are from ground-based GPS observations over the western North America from January 1, 2016 to December 31, 2016.

Training Dataset	**Band**	**2-Channel Ratio Method**			**3-Channel Ratio Method**		
		a	**b**	**c**	**a**	**b**	**c**
	17	0.684	-0.055	0.315	0.618	-0.063	0.387
1	18	0.583	-0.207	0.195	0.585	-0.211	0.199
	19	0.997	-0.026	0.000	1.000	-0.025	0.000
	17	0.701	-0.053	0.296	0.624	-0.062	0.381
2	18	0.582	-0.205	0.194	0.584	-0.209	0.198
	19	0.997	-0.026	0.000	1.000	-0.025	0.000
	17	0.724	-0.050	0.272	0.642	-0.060	0.362
3	18	0.583	-0.201	0.189	0.584	-0.205	0.194
	19	0.997	-0.026	0.000	1.000	-0.025	0.000
	17	0.688	-0.053	0.308	0.618	-0.062	0.384
4	18	0.581	-0.207	0.195	0.583	-0.210	0.200
	19	0.995	-0.025	0.000	0.998	-0.025	0.000
	17	0.732	-0.050	0.265	0.643	-0.059	0.361
5	18	0.583	-0.203	0.192	0.585	-0.207	0.197
	19	0.997	-0.026	0.000	1.000	-0.025	0.000
	17	0.667	-0.056	0.331	0.606	-0.064	0.399
6	18	0.581	-0.208	0.196	0.583	-0.211	0.201
	19	0.996	-0.026	0.000	0.999	-0.025	0.000
	17	0.713	-0.052	0.284	0.643	-0.059	0.360
7	18	0.584	-0.206	0.194	0.585	-0.210	0.199
	19	0.997	-0.026	0.000	1.000	-0.025	0.000
	17	0.779	-0.046	0.216	0.689	-0.054	0.313
8	18	0.581	-0.203	0.192	0.583	-0.206	0.196
	19	0.996	-0.026	0.000	1.000	-0.025	0.000

9	17	0.734	-0.050	0.262	0.647	-0.059	0.356
	18	0.583	-0.208	0.195	0.585	-0.212	0.200
	19	0.997	-0.026	0.000	1.000	-0.025	0.000
10	17	0.761	-0.047	0.233	0.679	-0.055	0.323
	18	0.579	-0.202	0.192	0.580	-0.205	0.196
	19	0.996	-0.026	0.000	1.000	-0.025	0.000

To evaluate the performance of the retrieval model that is developed based on the data of 2016, water vapor data obtained during 2017 to 2019 over the western North America and the Australia, representing two regions with different weather and climate conditions, are used in the validation process. Four statistic metrics are employed to evaluate the validation results. They are the mean absolute percentage error (MAPE), root mean squares error (RMSE), mean bias (MB), and coefficient of determination (R^2). The MAPE is used to measure the retrieval accuracy, defined as:

$$\mathbf{MAPE} = \frac{\mathbf{1}}{\mathbf{n}}\sum_{\mathbf{i=1}}^{\mathbf{n}}\left|\frac{\mathbf{PWV_{RS}-PWV_{GPS}}}{\mathbf{PWV_{GPS}}}\right| * \mathbf{100\%} \qquad (9)$$

where n denotes the number of data pairs; PWV_{RS} is the PWV column obtained from remote sensor i.e. the MERSI/FY-3B; PWV_{GPS} is the PWV observed from GPS data.

The RMSE is used to quantify the PWV differences between remote sensing PWV and reference data. It is defined as:

$$\mathbf{RMSE} = \sqrt{\frac{\mathbf{1}}{\mathbf{n}}\sum_{\mathbf{i=1}}^{\mathbf{N}}(\mathbf{PWV_{RS}} - \mathbf{PWV_{GPS}})^{\mathbf{2}}} \qquad (10)$$

The MB is used to estimate the mean bias between the two sets of PWV data. It is defined as:

$$\mathbf{MB} = \frac{\mathbf{1}}{\mathbf{n}}\sum_{\mathbf{i=1}}^{\mathbf{n}}(\mathbf{PWV_{RS}} - \mathbf{PWV_{GPS}}) \qquad (11)$$

The R^2 provides strength information between MERSI/FY-3B PWV and GPS PWV data. It is calculated as:

$$\mathbf{R^2} = \left[\frac{\sum_{i=1}^{n}(\mathbf{PWV_{GPS}}-\overline{\mathbf{PWV}}_{\mathbf{GPS}})(\mathbf{PWV_{RS}}-\overline{\mathbf{PWV}}_{\mathbf{RS}})}{\sqrt{\sum_{i=1}^{n}(\mathbf{PWV_{GPS}}-\overline{\mathbf{PWV}}_{\mathbf{GPS}})^2(\mathbf{PWV_{RS}}-\overline{\mathbf{PWV}}_{\mathbf{RS}})^2}}\right]^2 \quad (12)$$

where the $\overline{PWV}_{RS}$ and $\overline{PWV}_{GPS}$ denote the average values of remote sensing PWV and GPS PWV, respectively.

5.1 Validation in western North America

The first validation of the model is conducted using the water vapor data retrieved from the western North America. A total of 10,566 pairs of collocated PWV data observed during 2017 to 2019 are employed. The performance of PWV retrieval results from both single absorption channel and the weighted mean value is discussed in details.

5.1.1 Single Channel PWV Retrieval

The scatterplot displayed in Figure 8 reveals that the MERSI water vapor estimated from each absorption channel has a high accuracy in all the comparison studies using transmittance calculated with both 2-channel ratio method and 3-channel ratio method. The MAPE is in the range of 16.50% to 36.74% and RMSE is in the range of 4.964 mm to 8.156 mm. For band 17, the PWV retrieved using 2-channel ratio transmittance performs slightly better than the data calculated using 3-channel ratio transmittance. In contrast, for both band 18 and band 19, the retrieval results have a better retrieval accuracy using 3-channel ratio transmittance. The improvement of the retrieval accuracy may be due to the addition of another window channel. The use of another window channel in transmittance calculation might reduce the spectroscopic uncertainties. Previous studies showed that additional window channel could mitigate the impact on the water vapor continuum caused by surface types (Gao and Kaufman, 2003).

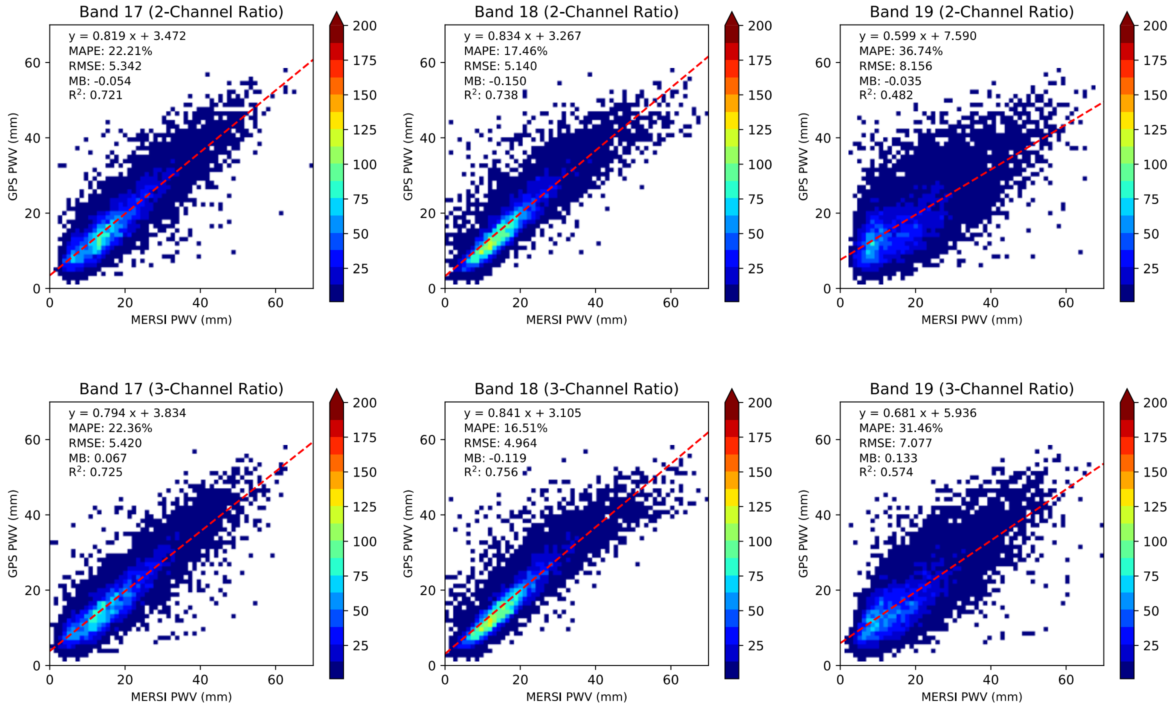


Figure 8 Normalized frequencies of the ground-based GPS PWV data and water vapor products retrieved from MERSI/FY-3B over the western North America using 2-channel ratio transmittance (upper panel, a total of 10,566 data points) and 3-channel ratio transmittance (lower panel, a total of 10,566 data points) over the period January 1, 2017 to December 31, 2019. The ensemble median is considered as the value of water vapor content calculated for corresponding pixel. The color bar represents the sample size.

Among the three channels, band 18 performs the best with the smallest RMSE of 4.964 mm and the MAPE of 16.51% while using 3-channel ratio transmittance. Band 19 has the lowest accuracy with RMSE of 8.156 mm and MAPE of 36.74%.

5.1.2 Weighted Mean PWV of Three Channels

The weighted mean PWV validation results calculated from the three absorption channels with both 2-channel ratio transmittance and 3-channel ratio transmittance are presented in Figure 9. Both results have a better accuracy than single-channel PWV retrieval, with RMSE of 5.157 mm using 2-channel ratio transmittance, and RMSE of 4.635 mm using 3-channel ratio transmittance. The results using 3-channel ratio transmittance have a MAPE of 16.72%, smaller than those calculated with 2-channel ratio transmittance.

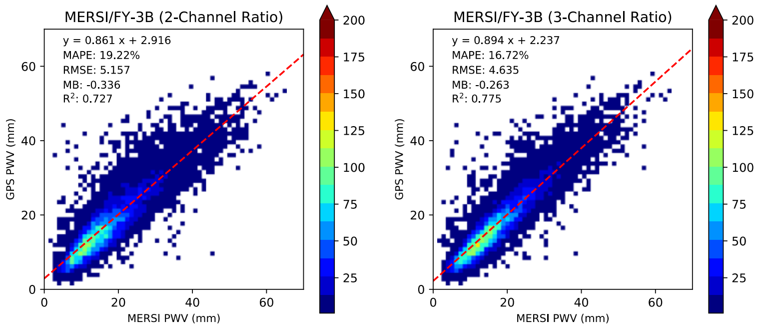
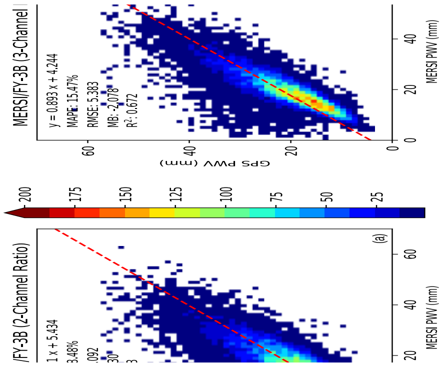


Figure 9 Normalized frequencies of the ground-based GPS PWV data and weighted mean PWV retrieved from MERSI/FY-3B over the western North America using 2-channel ratio transmittance and 3-channel ratio transmittance (a total of 10,566 data points) over the period January 1, 2017 to December 31, 2019. The ensemble median is considered as the value of water vapor content calculated for corresponding pixel. The color bar represents the sample size.

5.2 Validation in Australia

To further investigate the applicability of the algorithm in other regions of the world, PWV are retrieved from MERSI/FY-3B observed over Australia during 2017 to 2019 using the same set of coefficients shown in Table 2, which are developed using the 2016 data collected in the western North America. The retrieved PWV are then validated against collocated ground-based GPS PWV observations. A total of 15,600 collocated data pairs under clear conditions are observed over 419 GPS stations during the period of 2017 to 2019. The ensemble median is used to represent the corresponding retrieval value for each pixel.

Validation results in



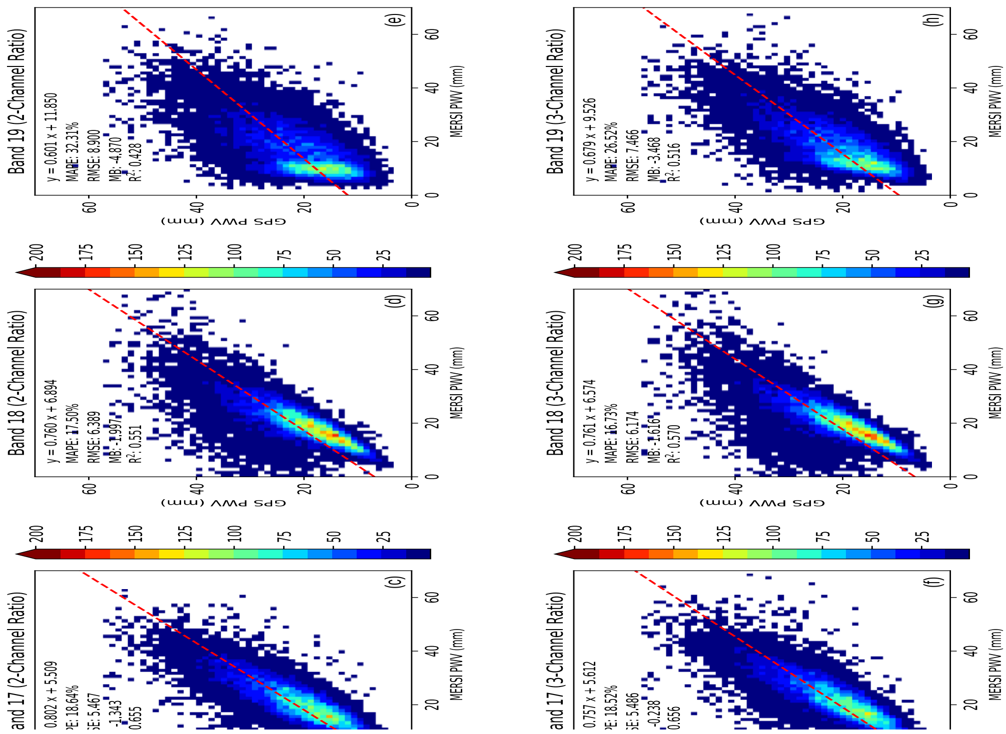
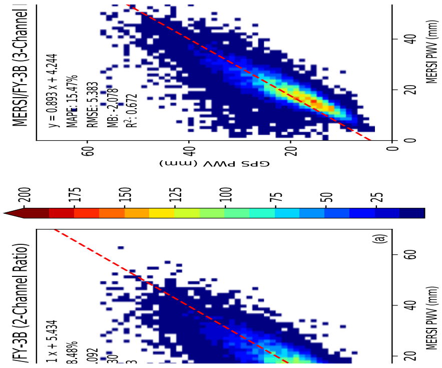


Figure *10* show that all PWV results retrieved from MERSI/FY-3B agree well with GPS PWV. The MAPE is in the range of 15.47% ~ 32.31%, while the RMSE is in the range of 5.383 mm ~ 8.900 mm. The weighted mean PWV have a better agreement with GPS than those retrieved from a single absorption channel. Moreover, the data retrieved using 3-channel ratio transmittance performs better than those calculated using 2-channel ratio transmittance. As shown in



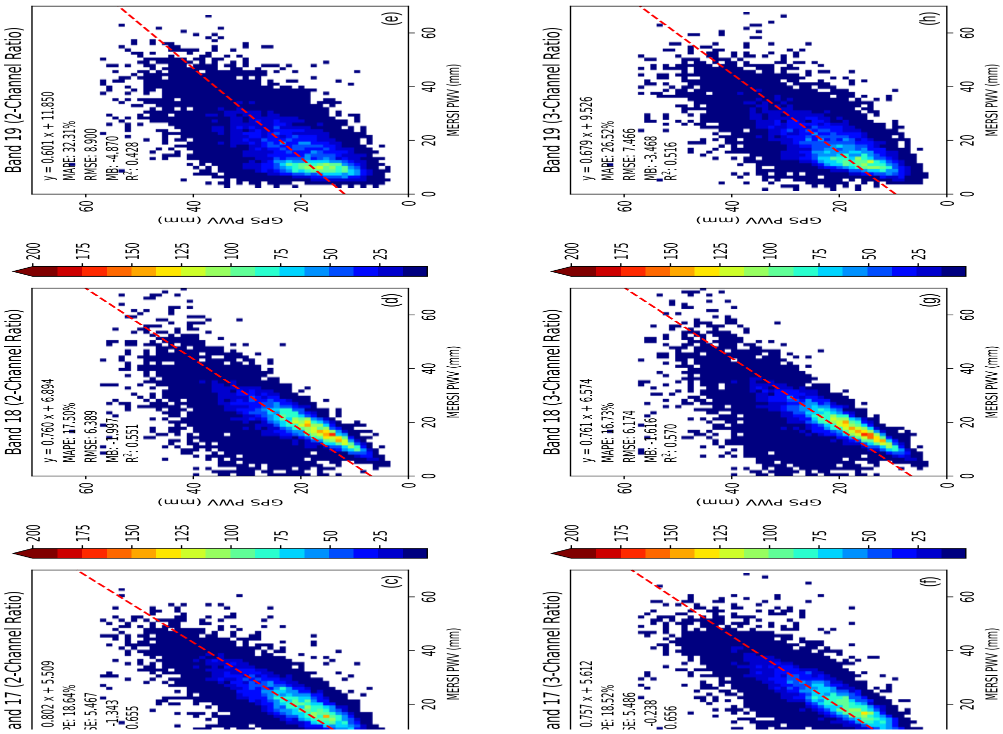


Figure *10*, the weighted mean PWV retrieved using 3-channel ratio transmittance has the best accuracy, with MAPE of 15.47% and RMSE of 5.383 mm. The observation from band 19 using 2-channel ratio transmittance has the worst retrieval accuracy, with MAPE of 32.31% and RMSE of 8.900 mm.

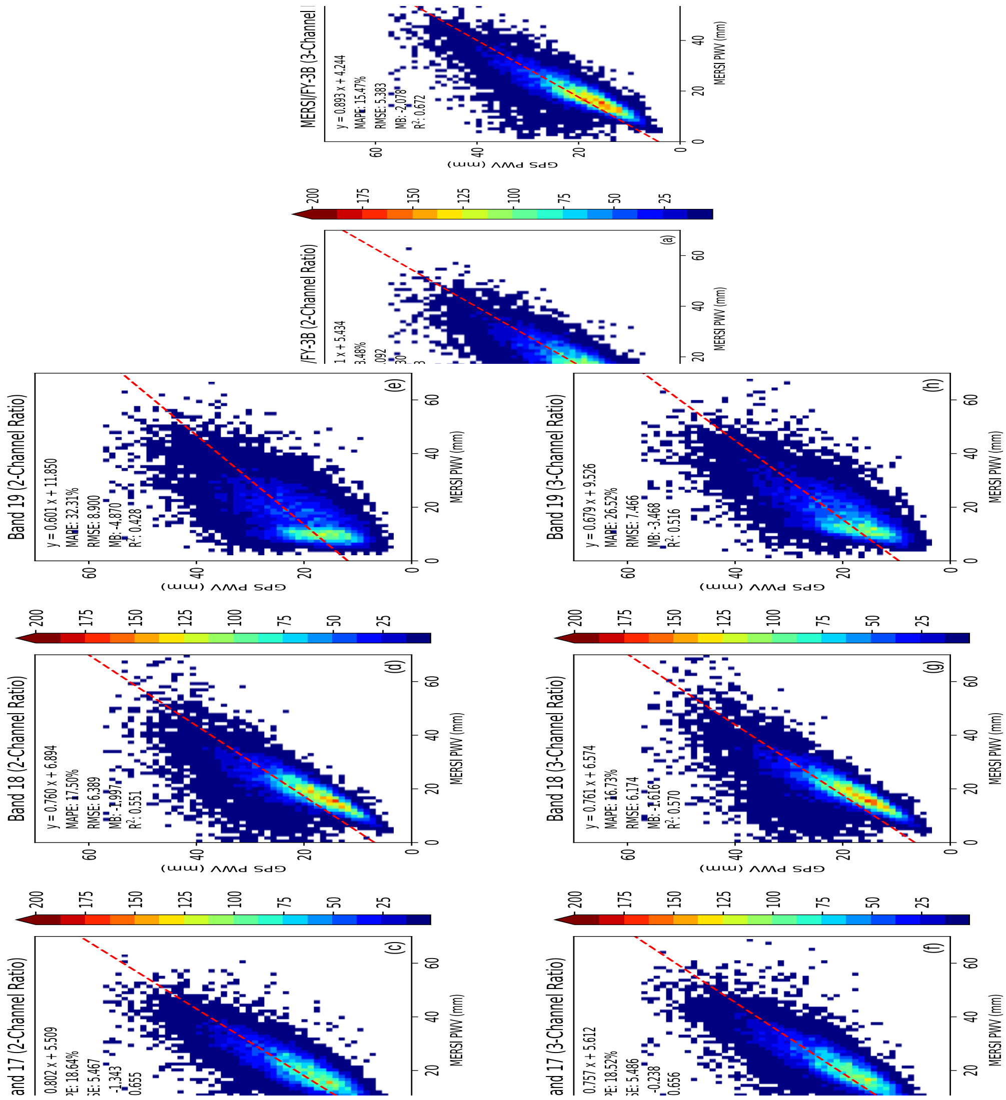


Figure 10 Normalized frequencies of the ground-based GPS PWV data and weighted mean PWV retrieved from MERSI/FY-3B over Australia over the period January 1, 2017 to December 31, 2019 (a total of 15,600 data points) (a) weighted mean PWV estimated using 2-channel ratio transmittance; (b) weighted mean PWV estimated using 3-channel ratio transmittance; (c) PWV retrieved from band 17 using 2-channel ratio transmittance; (d) PWV retrieved from band 18 using 2-channel ratio transmittance; (e) PWV retrieved from band 19 using 2-channel ratio transmittance; (f) PWV retrieved from band 17 using 3-channel ratio transmittance; (g) PWV retrieved from band 18 using 3-channel ratio transmittance; (h) PWV retrieved from band 19 using 3-channel ratio transmittance. The ensemble median is considered as the value of water vapor content calculated for corresponding pixel. The color bar represents the sample size.

6 Discussion

6.1 Stability of the algorithm

An ensemble-based empirical regression algorithm for water vapor retrieval from MERSI/FY-3B NIR channels has been developed with the observation data recorded in 2016 in the western North America. To assess the model's performance, we used temporally and spatially independent datasets in the model validation. The validation data were collected in the period 2017-2019 from both western North America and Australia. The water vapor distribution and variation of the Australia in the south hemispehre are very different from those in the western North America in the north hemisphere. The model test period 2017-2019 is also different from the model building period 2016. Therefore such an assessment is expected to objectively have a full evaluation the model's performance in both time and space domains.

A total of 10,566 pairs of collocated data points from the western North America and 15,600 pairs of data points from Australia are employed for validation analysis. The results show that all data records retrieved from MERSI/FY-3B have good agreements with reference GPS PWV reference data, indicating that the algorithm provides an effective and accurate way for water vapor retrieval. Moreover, no obvious difference in MAPE or RMSE has been found in the results , showing that the model is spatial-independent and temporal-independent and that the coefficients derived from the model are applicable on a global scale.

To study the temporal independency of the algorithm, the annual validation results for the period of 2017 ~ 2019 are summarized in Table 3 Annual validation results of the weighted mean PWV retrieved for both western North America and Australia region. For retrieval over western North America, the MAPE is in the range of 15.98% to 17.16%, and the RMSE is in the range of 3.764 mm to 6.493 mm. For retrieval over Australia, the MAPE is between 13.77% and

18.09%, and the RMSE is between 4.580 mm and 6.459 mm. Although the results in both sites show relatively good agreement with ground truth, an increase of retrieval error was revealed, with RMSE increase year by year. This is possibly caused by channel drifting.

Table 3 Annual validation results of the weighted mean PWV retrieved using 3-channel ratio transmittance compared against GPS observed PWV. The validation data were obtained from western North America and Australia for the period 2017 to 2019.

Region	Data Points	Year	3-Channel Ratio					
			Slope	Off-set	MAPE	RMSE (mm)	MB (mm)	R^2
Western North America	5964	2017	0.866	2.091	16.86%	3.764	0.429	0.865
	2187	2018	0.983	0.643	17.16%	4.311	-0.318	0.808
	2415	2019	0.934	3.099	15.98%	6.493	-1.923	0.555
Australia	5843	2017	0.938	2.230	13.77%	4.580	-1.005	0.770
	5420	2018	0.835	5.128	15.21%	5.236	-1.661	0.627
	4337	2019	0.893	6.183	18.09%	6.459	-4.046	0.614

In general, the water vapor retrieved using this newly developed ensemble-based regression algorithm is spatial and temporal independent. The algorithm coefficients are derived from a series of generally independent subsets of training data with slight overlapping. The multiple subsets have balanced the class distribution and reduced the potential sampling errors. Therefore, the ensemble median is bias-robust.

6.2 Comparison with previous studies

In previous studies, only a limited volume of MERSI/FY-3 series data have been used to retrieve water vapor and their results have shown large retrieval errors compared with ground-based PWV observation. For instance, recent work by Gong et al. (2018a) showed that the MAPE of MERSI varied in the range of 31.8–44.1% in the East Asian. In the work by Gong et al. (2018b), the MAPE of MERSI compared with GPS water vapor is 22.83% when using North

China's data. In general, the accuracy of water vapor retrieved from MERSI is worse than other water vapor products with NIR sensors, such as MODIS and MERIS (He and Liu, 2019). Using this newly proposed ensemble-based empirical regression algorithm, our study shows that the water vapor data retrieved from MERSI have a good agreement with the reference GPS data, with MAPE of 16.72% and RMSE of 4.635 mm in western North America, and MAPE of 15.47% and RMSE of 6.092 mm in Australia.

Conventional water vapor retrieval methods use a pre-calculated look-up table generated from radiative transfer models (Hu et al., 2011). In some cases, such as the MERSI/FY-3B sensor, the conventional methods can show significant bias. In this situation the empirical algorithm of this paper can be used operationally to estimate water vapor directly with a significantly reduced bias. It can therefore contribute to greater accuracy in global water vapor estimation.

7 Conclusion

Despite many studies on MERSI onboard of FY-3 series, the accuracy of water vapor retrieval methods is still low. In this study, our conclusions are:

(1) an ensemble-based empirical regression retrieval algorithm is for the first time proposed to retrieve PWV from the NIR channels of MERSI sensor onboard the Chinese FY-3B satellite. This algorithm uses real-world data to mathematically establish the relationship between the transmittance and PWV. Ensemble analysis with the bootstrap method is employed to resample the model training dataset into 10 independent training subsets. The ensemble members are expected to have biases randomly so that the ensemble median could be bias-free.

(2) the new method has been validated using independent GPS PWV data collected during 2017 to 2019 in both western North America and Australia. Water vapor data from MERSI/FY-3B

satellite are calculated from both each of single absorption channels and the weighted mean of three channels. Validation results show that the PWV retrieved from both western North America and Australia agree well with ground-based GPS PWV observations.

(3) the weighted mean PWV calculated using 3-channel ratio transmittance performs the best, with MAPE of 16.72% and RMSE of 4.635 mm in the western North America, and MAPE of 15.47% and RMSE of 5.383 mm in Australia.

(4) The differences of the two sets of in MAPE and RMSE for western North America and Australia are small. It is reasonable to state that our proposed model is spatially independent. It has the potential to be applied to other global regions as well.

(5) the annual accuracy of the validated PWV is reasonably small over the period 2017-2019 though the model's coefficients are estimated based on dataset of 2016. It is reasonable to state that this algorithm is temporally independent. However the RMSE of the validation results also shows a slight increase trend over the years at both western North America and Australia. This is probably because of the channel drifting of MERSI sensor. The channel drifting will result in a change of transmittance observed in the water vapor absorption channels and then affect the retrieval accuracy. Therefore, it is suggested to calibrate the MERSI sensor regularly in order to maintain the PWV retrieval accuracy.

Acknowledgment

The support from the Key Program of the National Natural Science Foundation of China (project No. 41730109) is acknowledged. The support by the Hong Kong Research Grants Council (RGC) (project No. B-Q52W PolyU 152149/16E) is acknowledged. The support by the Emerging Frontier Area (EFA) Scheme of Research Institute for Sustainable Urban Development (RISUD) of the Hong Kong Polytechnic University (project No. 1-BBWJ) is appreciated. The authors would like to thank the four anonymous reviewers and the RSE editorial team for their invaluable comments that improve the quality of this manuscript. They thank the research staffs in China Meteorological Administration /National Satellite Meteorological Center (CMA/NSMC) (http://satellite.nsmc.org.cn) for providing MERSI/FY3B level 1 reflectance 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 Geoscience Australia (ftp://ftp.ga.gov.au/geodesyoutgoing/gnss/products/troposphere/rapid/) is acknowledged for providng the PWV data for the period 2017-2019.

References:

Ackerman, S.A., Strabala, K.I., Menzel, W.P., Frey, R.A., Moeller, C.C., Gumley, L.E., 1998. Discriminating clear sky from clouds with MODIS. J. Geophys. Res. 103, 32. https://doi.org/10.1029/1998JD200032

Ashcroft, L., Gergis, J., Karoly, D.J., 2016. Long-term stationarity of El Niño–Southern Oscillation teleconnections in southeastern Australia. Clim. Dyn. 46, 2991–3006. https://doi.org/10.1007/s00382-015-2746-3

Batista, G.E.A.P.A., Prati, R.C., Monard, M.C., 2004. A Study of the Behavior of Several Methods for Balancing Machine Learning Training Data. SIGKDD Explor Newsl 6, 20–29. https://doi.org/10.1145/1007730.1007735

Belward, A., 2016. The Global Observing System for Climate: Implementation Needs. Ref. Number GCOS-200 315, 349.

Berk, A., Bernstein, L.S., Robertson, D.C., 1987. MODTRAN: A Moderate Resolution Model for LOWTRAN. Spectral Sciences. Inc., Burlington, MA.

Berk, A., Conforti, P., Kennett, R., Perkins, T., Hawes, F., Bosch, J. van den, 2014. MODTRAN6: a major upgrade of the MODTRAN radiative transfer code, in: Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XX. Presented at the Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XX, International Society for Optics and Photonics, p. 90880H. https://doi.org/10.1117/12.2050433

Dach, R., Lutz, S., Walser, P., Fridez, P., 2015. Bernese GNSS Software Version 5.2, Dach, Rolf; Lutz, Simon; Walser, Peter; Fridez, Pierre (eds.) (2015). Bernese GNSS Software Version 5.2. Bern: University of Bern, Bern Open Publishing. University of Bern, Bern Open Publishing, Bern. https://doi.org/10.7892/boris.72297

Dong, C., Yang, J., Yang, Z., Lu, N., Shi, J., Zhang, P., Liu, Y., Cai, B., Zhang, W., 2009. An Overview of a New Chinese Weather Satellite FY-3A. Bull. Am. Meteorol. Soc. 90, 1531–1544. https://doi.org/10.1175/2009BAMS2798.1

Efron, B., 1979. Bootstrap Methods: Another Look at the Jackknife. Ann. Stat. 7, 1–26. https://doi.org/10.1214/aos/1176344552

Elgered, G., Johansson, J.M., Rönnäng, B.O., Davis, J.L., 1997. Measuring regional atmospheric water vapor using the Swedish Permanent GPS Network. Geophys. Res. Lett. 24, 2663–2666. https://doi.org/10.1029/97GL02798

Elgered, G., Plag, H.P., Marel, H. van der, Barlag, S., Nash, J. (Eds.), 2005. Exploitation of ground-based GPS for operational numerical weather prediction and climate applications Final Report, COST Action 716 EUR 21639 edition. ed. COST Office, Luxembourg.

Elliott, W.P., Gaffen, D.J., 1991. On the Utility of Radiosonde Humidity Archives for Climate Studies. Bull. Am. Meteorol. Soc. 72, 1507–1520. https://doi.org/10.1175/1520-0477(1991)072<1507:OTUORH>2.0.CO;2

Fraser, R.S., Kaufman, Y.J., 1985. The Relative Importance of Aerosol Scattering and Absorption in Remote Sensing. IEEE Trans. Geosci. Remote Sens. GE-23, 625–633. https://doi.org/10.1109/TGRS.1985.289380

Gao, B.-C., Kaufman, Y.J., 2003. Water vapor retrievals using Moderate Resolution Imaging Spectroradiometer (MODIS) near-infrared channels. J. Geophys. Res. Atmospheres 108. https://doi.org/10.1029/2002JD003023

Gong, S., Hagan, D.F.T., Lu, J., Wang, G., 2018a. Validation on MERSI/FY-3A precipitable water vapor product. Adv. Space Res. 61, 413–425. https://doi.org/10.1016/j.asr.2017.10.005

Gong, S., Hagan, D.F.T., Wu, X., Wang, G., 2018b. Spatio-temporal analysis of precipitable water vapour over northwest china utilizing MERSI/FY-3A products. Int. J. Remote Sens. 39, 3094–3110. https://doi.org/10.1080/01431161.2018.1437298

Hanssen, R.F., Weckwerth, T.M., Zebker, H.A., Klees, R., 1999. High-Resolution Water Vapor Mapping from Interferometric Radar Measurements. Science 283, 1297–1299. https://doi.org/10.1126/science.283.5406.1297

He, J., Liu, Z., 2020. Water Vapor Retrieval From MODIS NIR Channels Using Ground-Based GPS Data. IEEE Trans. Geosci. Remote Sens. 1–12. https://doi.org/10.1109/TGRS.2019.2962057

He, J., Liu, Z., 2019. Comparison of Satellite-Derived Precipitable Water Vapor Through Near-Infrared Remote Sensing Channels. IEEE Trans. Geosci. Remote Sens. 57, 10252–10262. https://doi.org/10.1109/TGRS.2019.2932847

Held, I.M., Soden, B.J., 2000. Water Vapor Feedback and Global Warming. Annu. Rev. Energy Environ. 25, 441–475. https://doi.org/10.1146/annurev.energy.25.1.441

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., Chiara, G.D., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R.J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., Rosnay, P. de, Rozum, I., Vamborg, F., Villaume, S., Thépaut, J.-N., 2020. The ERA5 global reanalysis. Q. J. R. Meteorol. Soc. 146, 1999–2049. https://doi.org/10.1002/qj.3803

Hu, G., 2017. Report on the analysis of the Asia Pacific Regional Geodetic Project (APRGP) GPS campaign 2016. Geoscience Australia. https://doi.org/10.11636/Record.2017.017

Hu, X.-Q., Huang, Y.-B., Lu, Q.-F., Zheng, J., 2011. Retrieving Precipitable Water Vapor Based on the Near-infrared Data of FY-3A Satellite. J. Appl. Meteorol. Sci.

Karl, T.R., Trenberth, K.E., 2003. Modern Global Climate Change. Science 302, 1719–1723. https://doi.org/10.1126/science.1090228

Kaufman, Y.J., Gao, B.C., 1992. Remote sensing of water vapor in the near IR from EOS/MODIS. IEEE Trans. Geosci. Remote Sens. 30, 871–884. https://doi.org/10.1109/36.175321

King, M.D., Kaufman, Y.J., Menzel, W.P., Tanre, D., 1992. Remote sensing of cloud, aerosol, and water vapor properties from the Moderate Resolution Imaging Spectrometer (MODIS). IEEE Trans. Geosci. Remote Sens. 30, 2–27.

King, M.D., Menzel, W.P., Kaufman, Y.J., Tanre, D., Gao, B.-C., Platnick, S., Ackerman, S.A., Remer, L.A., Pincus, R., Hubanks, P.A., 2003. Cloud and aerosol properties, precipitable water, and profiles of temperature and water vapor from MODIS. IEEE Trans. Geosci. Remote Sens. 41, 442–458. https://doi.org/10.1109/TGRS.2002.808226

Levin, N., Johansen, K., Hacker, J.M., Phinn, S., 2014. A new source for high spatial resolution night time images — The EROS-B commercial satellite. Remote Sens. Environ. 149, 1–12. https://doi.org/10.1016/j.rse.2014.03.019

Li, Z., Muller, J.-P., Cross, P., 2003. Comparison of precipitable water vapor derived from radiosonde, GPS, and Moderate-Resolution Imaging Spectroradiometer measurements. J. Geophys. Res. Atmospheres 108, 4651. https://doi.org/10.1029/2003JD003372

Lindenbergh, R., Keshin, M., Marel, H. van der, Hanssen, R., 2008. High resolution spatio-temporal water vapour mapping using GPS and MERIS observations. Int. J. Remote Sens. 29, 2393–2409. https://doi.org/10.1080/01431160701436825

Martins, J.V., Tanré, D., Remer, L., Kaufman, Y.J., Mattoo, S., Levy, 2002. MODIS Cloud screening for remote sensing of aerosols over oceans using spatial variability. Geophys. Res. Lett. 29. https://doi.org/10.1029/2001GL013252

Mockler, S.B., 1995. Water Vapor in the Climate System. American Geophysical Union, Washington, D.C.

Raval, A., Ramanathan, V., 1989. Observational determination of the greenhouse effect. Nature 342, 758–761. https://doi.org/10.1038/342758a0

Roberts, R.E., Selby, J.E.A., Biberman, L.M., 1976. Infrared continuum absorption by atmospheric water vapor in the 8–12-μm window. Appl. Opt. 15, 2085–2090. https://doi.org/10.1364/AO.15.002085

Rothman, L.S., Gordon, I.E., Barbe, A., Benner, D.C., Bernath, P.F., Birk, M., Boudon, V., Brown, L.R., Campargue, A., Champion, J.-P., Chance, K., Coudert, L.H., Dana, V., Devi, V.M., Fally, S., Flaud, J.-M., Gamache, R.R., Goldman, A., Jacquemart, D., Kleiner, I., Lacome, N., Lafferty, W.J., Mandin, J.-Y., Massie, S.T., Mikhailenko, S.N., Miller, C.E., Moazzen-Ahmadi, N., Naumenko, O.V., Nikitin, A.V., Orphal, J., Perevalov, V.I., Perrin, A., Predoi-Cross, A., Rinsland, C.P., Rotger, M., Šimečková, M., Smith, M.A.H., Sung, K., Tashkun, S.A., Tennyson, J., Toth, R.A., Vandaele, A.C., Vander Auwera, J., 2009. The HITRAN 2008 molecular spectroscopic database. J. Quant. Spectrosc. Radiat. Transf., HITRAN 110, 533–572. https://doi.org/10.1016/j.jqsrt.2009.02.013

Schläpfer, D., Borel, C.C., Keller, J., Itten, K.I., 1998. Atmospheric Precorrected Differential Absorption Technique to Retrieve Columnar Water Vapor. Remote Sens. Environ. 65, 353–366. https://doi.org/10.1016/S0034-4257(98)00044-3

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

Sherwood, S.C., Roca, R., Weckwerth, T.M., Andronova, N.G., 2010. Tropospheric water vapor, convection, and climate. Rev. Geophys. 48, RG2001. https://doi.org/10.1029/2009RG000301

Shi, F., Xin, J., Yang, L., Cong, Z., Liu, R., Ma, Y., Wang, Y., Lu, X., Zhao, L., 2018. The first validation of the precipitable water vapor of multisensor satellites over the typical regions in China. Remote Sens. Environ. 206, 107–122. https://doi.org/10.1016/j.rse.2017.12.022

Soden, B.J., Wetherald, R.T., Stenchikov, G.L., Robock, A., 2002. Global Cooling After the Eruption of Mount Pinatubo: A Test of Climate Feedback by Water Vapor. Science 296, 727–730. https://doi.org/10.1126/science.296.5568.727

Wang, X., Zhang, K., Wu, S., Li, Z., Cheng, Y., Li, L., Yuan, H., 2018. The correlation between GNSS-derived precipitable water vapor and sea surface temperature and its responses to El Niño–Southern Oscillation. Remote Sens. Environ. 216, 1–12. https://doi.org/10.1016/j.rse.2018.06.029

Wang, X., Zhao, D.-Z., Su, X., Yang, J.-H., Ma, Y.-J., 2012. Retrieving precipitable water vapor based on FY-3A near-IR data. J Infrared Millim Waves 31, 550–555. https://doi.org/10.3724/SP.J.1010.2012.00550

Ware, R.H., Fulker, D.W., Stein, S.A., Anderson, D.N., Avery, S.K., Clark, R.D., Droegemeier, K.K., Kuettner, J.P., Minster, J.B., Sorooshian, S., 2000. SuomiNet: A Real-Time National GPS Network for Atmospheric Research and Education. Bull. Am. Meteorol. Soc. 81, 677–694. https://doi.org/10.1175/1520-0477(2000)081<0677:SARNGN>2.3.CO;2

Warner, J.X., Ellingson, R.G., 2000. A New Narrowband Radiation Model for Water Vapor Absorption. J. Atmospheric Sci. 57, 1481–1496. https://doi.org/10.1175/1520-0469(2000)057<1481:ANNRMF>2.0.CO;2

Wind, G., Platnick, S., King, M.D., Hubanks, P.A., Pavolonis, M.J., Heidinger, A.K., Yang, P., Baum, B.A., 2010. Multilayer Cloud Detection with the MODIS Near-Infrared Water Vapor Absorption Band. J. Appl. Meteorol. Climatol. 49, 2315–2333. https://doi.org/10.1175/2010JAMC2364.1

Wu, C.F.J., 1986. Jackknife, Bootstrap and Other Resampling Methods in Regression Analysis. Ann. Stat. 14, 1261–1295. https://doi.org/10.1214/aos/1176350142

Yang, J., Zhang, P., Lu, N., Yang, Z., Shi, J., Dong, C., 2012. Improvements on global meteorological observations from the current Fengyun 3 satellites and beyond. Int. J. Digit. Earth 5, 251–265. https://doi.org/10.1080/17538947.2012.658666

List of Figure Captions

Figure 1 Distribution map of 256 GPS stations located in the western North America. They are used for FY-3B water vapor calibration and validation analysis. The color bar represents the elevation of the GPS stations, in unit of meters.

Figure 2 Distribution map of 419 GPS stations located in Australia used for FY-3B water vapor validation analysis. The color bar represents the elevation of the GPS stations, in unit of meters.

Figure 3 Spectral transmission of atmosphere contents in the presence of water vapor at 0.6 g/cm^2, considering H_2O, O_3 and the combined transmission. Computations were performed using MODTRAN 4 model. The color bars at the bottom show the location of the MERSI water vapor absorption channels (black) and the window channels (blue) used in the retrieval study.

Figure 4 Flow chart of the newly proposed method to retrieve PWV from NIR channels of the MERSI sensor onboard the FY-3B satellite. PWV estimated from ground-based GPS observations are used as reference values. The light brown boxes denote the input data, and the green boxes are the output results.

Figure 5 The number of data pairs of the collocated GPS and MERSI/FY-3B L1 NIR channel reflectance recorded in each day in each month of 2016 over the western North America under cloud free conditions. They are used for model development. The color bar denotes the number of data pairs.

Figure 6 Example of scatterplot of the relationship between optical path (slant) column water vapor observed from GPS and the transmittance from three absorption channels of MERSI/FY-3B with 2-channel ratio method.

Figure 7 (a) Example of regression functions from MERSI/FY-3B using 2-channel ratio method; (b) the corresponding normalized weighting factors of the three absorption channels based on their sensitivity to transmittance.

Figure 8 Normalized frequencies of the ground-based GPS PWV data and water vapor products retrieved from MERSI/FY-3B over the western North America using 2-channel ratio transmittance (upper panel, a total of 10,566 data points) and 3-channel ratio transmittance (lower panel, a total of 10,566 data points) over the period January 1, 2017 to December 31, 2019. The ensemble median is considered as the value of water vapor content calculated for corresponding pixel. The color bar represents the sample size.

Figure 9 Normalized frequencies of the ground-based GPS PWV data and weighted mean PWV retrieved from MERSI/FY-3B over the western North America using 2-channel ratio transmittance and 3-channel ratio transmittance (a total of 10,566 data points) over the period January 1, 2017 to December 31, 2019. The ensemble median is considered as the value of water vapor content calculated for corresponding pixel. The color bar represents the sample size.

Figure 10 Normalized frequencies of the ground-based GPS PWV data and weighted mean PWV retrieved from MERSI/FY-3B over Australia over the period January 1, 2017 to December 31, 2019 (a total of 15,600 data points) (a) weighted mean PWV estimated using 2-channel ratio transmittance; (b) weighted mean PWV estimated using 3-channel ratio transmittance; (c) PWV retrieved from band 17 using 2-channel ratio transmittance; (d) PWV retrieved from band 18 using 2-channel ratio transmittance; (e) PWV retrieved from band 19 using 2-channel ratio transmittance; (f) PWV retrieved from band 17 using 3-channel ratio transmittance; (g) PWV retrieved from band 18 using 3-channel ratio transmittance; (h) PWV retrieved from band 19 using 3-channel ratio transmittance. The ensemble median is considered as the value of water vapor content calculated for corresponding pixel. The color bar represents the sample size.