

Effects of a Cloud Filtering Method for Fengyun-3C Microwave Humidity and Temperature Sounder Measurements over Ocean on Retrievals of Temperature and Humidity postprint

Authors: HE Qiurui, WANG Zhenzhan, HE Jieying

Date: 2017-03-10T00:00:00+00:00

Abstract

For Microwave Humidity and Temperature sounder (MWHTS) measurements over ocean, a cloud filtering method is presented to filter out cloud- and precipitation-affected observations by analyzing the sensitivity of the simulated brightness temperatures of MWHTS to cloud liquid water, and using the root mean square error (RMSE) between observation and simulation in clear sky as a reference standard. The atmospheric temperature and humidity profiles are retrieved using MWHTS measurements with and without filtering by multiple linear regression (MLR), artificial neural networks (ANN) and one-dimensional variational (1DVAR) retrieval methods, respectively, and the effects of the filtering method on the retrieval accuracies are analyzed. The numerical results show that the filtering method can improve the retrieval accuracies of the MLR and the 1DVAR retrieval methods, but have little influence on that of the ANN. In addition, the dependencies of the retrieval methods upon the testing samples of brightness temperature are studied, and the results show that the 1DVAR retrieval method has great stability due to that the testing samples have great impact on the retrieval accuracies of the MLR and the ANN, but have little impact on that of the 1DVAR.

Full Text

Preamble

Effects of a Cloud Filtering Method for Fengyun-3C Microwave Humidity and Temperature Sounder Measurements over Ocean on Retrievals of Temperature and Humidity

HE Qiurui^{1,2}, WANG Zhenzhan¹, HE Jieying¹

¹. Key Laboratory of Microwave Remote Sensing, National Space Science Center, Chinese Academy of Sciences, Beijing 100190, China

². University of Chinese Academy of Sciences, Beijing 100049, China

Abstract: For Microwave Humidity and Temperature Sounder (MWHTS) measurements over ocean, we present a cloud filtering method to screen out cloud- and precipitation-affected observations by analyzing the sensitivity of simulated MWHTS brightness temperatures to cloud liquid water, using the root mean square error (RMSE) between observation and simulation in clear sky as a reference standard. Atmospheric temperature and humidity profiles are retrieved using MWHTS measurements with and without filtering by multiple linear regression (MLR), artificial neural networks (ANN), and one-dimensional variational (1DVAR) retrieval methods, respectively, and the effects of the filtering method on retrieval accuracies are analyzed. Numerical results show that the filtering method can improve the retrieval accuracies of the MLR and 1DVAR methods, but has little influence on that of the ANN. In addition, the dependencies of the retrieval methods upon the testing samples of brightness temperature are studied, and results show that the 1DVAR method has great stability because the testing samples have great impact on the retrieval accuracies of MLR and ANN, but have little impact on that of 1DVAR.

Index Terms: FY-3C/MWHTS, cloud filtering method, multiple linear regression, artificial neural networks, one-dimensional variational retrieval

1. Introduction

Space-borne radiometers represent an important data source for retrieving atmospheric parameters, measuring the thermal radiation emanating from the Earth's surface and atmosphere [?]. In addition to absorption and emission by atmospheric constituents, electromagnetic wave propagation in the atmosphere is influenced by scattering from clouds and precipitation. These scattering effects, which depend on cloud cover, cloud liquid water content, cloud ice water content, raindrop size, and other microphysical parameters, render the atmospheric radiative transfer process more complex than in clear-sky conditions and increase the nonlinearity of the radiative transfer equation, thereby complicating atmospheric parameter retrieval [?]. Because clouds do not absorb and scatter microwave radiation as strongly as visible or infrared radiation, microwave radiometers possess a robust capability for continuous all-weather detection [?]. However, a cloud filtering method capable of screening out microwave observations affected by clouds and precipitation is critical to ensuring retrieval accuracy. Although satellite-borne sounders have measured atmospheric temperature and humidity profiles for over 50 years [?], retrieval strategies can be classified into two categories: statistical methods and physical methods.

Statistical methods, including multiple linear regression and artificial neural networks, do not involve physical models but instead exploit statistical relationships

between atmospheric parameters and observations to obtain retrievals. When thick clouds and/or rain exist within the sounder's field of view, the statistical model may become more complex and inaccurate, directly affecting retrieval accuracy. For physical retrieval methods, accurate modeling of atmospheric radiative transfer is paramount. Because the electromagnetic wave scattering mechanism in clouds and rain is complicated, scattering simulations are often difficult and inaccurate, increasing uncertainty between simulations and observations during inversion of the radiative transfer equation and adversely affecting retrieval accuracy.

In summary, handling cloud- and precipitation-affected observations from satellite-borne microwave radiometers is particularly important for inversion processes.

Many previous studies have developed cloud filtering methods for microwave observations, where thick cloud and/or rain are collectively referred to as "cloud" for handling affected observations. Karstens et al. used relative humidity profile values from meteorological observation data as thresholds to determine clear-sky cases [?]. Ishimoto et al. and Li et al. filtered cloud-affected brightness temperatures according to infrared cloud imagery or cloud products [?]. Both approaches obviously depend on third-party data sources that may introduce additional errors into the retrieval process, and cloud filtering criteria must vary depending on the characteristics of the third-party data. Compared with these methods, taking full advantage of the characteristics of the satellite data itself offers significant benefits for dealing with cloud- and precipitation-affected observations. The effects of clouds and rain on satellite microwave radiometer measurements have been studied through simulations and observations [?]. For the Advanced Microwave Sounding Unit (AMSU)-B with two window channels (at frequencies 89 and 150 GHz) and three water vapor channels (at frequencies 183 ± 1 , 183 ± 3 , and 183 ± 7 GHz), Burns et al. analyzed the correspondence between thick clouds and brightness temperature depressions and suggested a criterion based on differences between observed brightness temperatures at 183.31 ± 3 and 183.31 ± 1 GHz to screen out observations in severe convective weather. This filtering method improved agreement between observations and simulations by up to a factor of two in their case study [?]. Hong et al. developed a method to detect tropical deep convective clouds and convective overshooting using brightness temperature differences between AMSU-B water vapor channels. This method, validated by two aircraft cases and a radiative transfer model, also accounted for AMSU-B's varying viewing angle [?]. Buehler et al. developed a method combining threshold brightness temperatures from AMSU-B channel 18 suggested by Greenwald and threshold brightness temperature differences suggested by Burns to filter high and heavily ice-laden clouds in AMSU-B observations [?]. The robustness of this cloud filtering method was demonstrated through a mid-latitude winter case study, and the method was also applied to study different biases in upper tropospheric humidity climatologies. These studies show that different cloud filtering methods based on satellite microwave measurement characteristics can be developed depending on application goals. MWHTS not only

has the same water vapor channels as AMSU-B but also adds two new water vapor channels at frequencies 183.31 ± 1.8 and 183.31 ± 4.5 GHz. However, few studies have investigated cloud filtering methods for MWHTS observations prior to atmospheric parameter inversion.

In our study, we analyze the sensitivity of MWHTS water vapor channels to cloud liquid water, take the RMSE of simulations with respect to observations in clear sky as a reference value, and set three cloud filtering criteria based on mutual relationships among water vapor channel brightness temperatures to screen out cloud- and precipitation-affected observations that cannot be accurately simulated by the radiative transfer model RTTOV (Radiative Transfer for Television and Infrared Observation Satellite Operational Vertical Sounder) [?]. After determining the optimal cloud filtering criteria, MWHTS observations over ocean from 1 to 28 February 2015, before and after filtering, are used to retrieve atmospheric temperature and water vapor profiles by MLR, ANN, and 1DVAR methods, and the effects of the filtering method on retrieval accuracies are investigated. Additionally, MWHTS observations over ocean from 1 to 31 May 2015 are used to test the stability of these three retrieval methods.

2.1. FY-3C/MWHTS

The MWHTS aboard the FY-3C satellite detects the Earth-atmosphere system in a cross-track scanning manner and is a total power radiometer. MWHTS has 15 channels: eight temperature channels, five water vapor channels, and two window channels. The eight temperature channels, centered at the 118.75 GHz oxygen absorption line, are used operationally for the first time internationally to measure temperature from the surface to the upper atmosphere. The five water vapor channels, centered at the 183.35 GHz water vapor absorption line, aim to sound humidity and precipitation in the troposphere. The two window channels at 89.0 and 150.0 GHz can provide surface information. MWHTS has a swath width of 2645 km. It takes 2.667 seconds to complete one scan line including 98 fields of view (FOV), and its nominal FOV at nadir is 16 km [?]. Some MWHTS channel characteristics, including channel center frequency, polarization, bandwidth, channel sensitivity, and in-flight channel sensitivity, are listed in Table 1 [?]. The weighting functions for temperature and humidity channels, calculated from the U.S. standard atmospheric profile at nadir by RTTOV with surface emissivities set to 0.5, are shown in Figure 1 [Figure 1: see original paper].

2.2. Data and Model

The data used in our study include MWHTS L1b brightness temperature data products and European Centre for Medium-Range Weather Forecasts (ECMWF) ERA-Interim reanalysis data. MWHTS brightness temperature data products are provided by the National Satellite Meteorological Center (NSMC) (<http://www.nsmc.cma.gov.cn/NSMC/HOME/Index.html>).

ECMWF ERA-Interim reanalysis data with horizontal resolution of $1^\circ \times 1^\circ$ and temporal resolution of 6 h (available at 0000 UTC, 0600 UTC, 1200 UTC, and 1800 UTC) are provided by ECMWF [?]. The surface parameters from ERA-Interim reanalysis data used in our study include skin temperature, 2 m air temperature, 2 m air dewpoint temperature, 10 m wind speed, and surface pressure. The profile parameters include temperature, humidity, cloud liquid water, and cloud ice water across 37 pressure levels unevenly spaced from 1000 hPa to 1 hPa.

MWHTS brightness temperatures and ERA-Interim reanalysis data over ocean covering the geographic area (135°E - 165°E , 0°N - 30°N) are selected to generate two datasets. The first dataset is a statistical analysis dataset containing collocated MWHTS brightness temperatures with ECMWF reanalysis covering the period from 1 February 2014 to 31 January 2015. The second dataset is a testing dataset with the same type of collocated data but covering the period from 1 to 28 February 2014. The collocation criteria require that the time difference between brightness temperatures and reanalysis be less than 10 minutes, and the absolute distance between their positions (latitude and longitude) be less than 0.05° . Based on these criteria, the statistical analysis dataset contains 490,142 collocated samples and the testing dataset contains 37,995 collocated samples. In our study, the fast radiative transfer model RTTOV version 11.2 developed by ECMWF is used to calculate simulated MWHTS brightness temperatures and brightness temperature gradients.

3. Cloud Filtering Method

For the statistical analysis dataset containing 490,142 collocated samples described in Section 2.2, we use both the absorption-based and scattering-based models of RTTOV to simulate MWHTS brightness temperatures and evaluate simulation accuracies through RMSEs with respect to observations. To compare RTTOV simulation accuracies in clear sky versus all-weather conditions, we select 2,485 clear-sky collocated samples from the all-weather statistical analysis dataset based on zero cloud liquid water content. We simulate MWHTS observations in clear sky and all-weather conditions, obtaining simulation accuracies $\text{RMSE}_{\text{clear}}$ and RMSE_{all} , respectively, as shown in Figure 2 [Figure 2: see original paper].

In clear sky, the simulation accuracies of the two models are equivalent. The RMSEs in the two window channels are large, while RMSEs in temperature channels are within 2 K (except for channels 2, 4, and 9), and RMSEs in humidity channels are within 3 K. In all-weather conditions, channels 2, 3, and 4 are sensitive to the upper atmosphere above 100 hPa and are less affected by clouds and rain, so the absorption-based model's simulation accuracies are nearly equivalent to those in clear sky. However, for the scattering-based model, simulation accuracies are very poor, possibly due to inaccurate atmospheric parameters in the upper atmosphere, such as cloud cover and cloud ice water content profiles. The RMSEs in other channels are higher than in clear sky. Except for

channels 7, 8, 10, and 15, where the scattering-based model's simulation accuracies are slightly higher than the absorption-based model, the absorption-based model is obviously superior. Two reasons may explain this: First, ERA-Interim reanalysis lacks rain and snow parameters that cause major scattering contributions, which are set to zero in RTTOV simulations; second, the scattering-based model requires more atmospheric parameters than the absorption-based model, and inaccuracies in atmospheric parameters, especially in cloud cover profiles, can affect RTTOV simulation accuracies [?]. Based on this analysis, we choose the absorption-based RTTOV model to simulate MWHTS observations in our study. However, to achieve higher simulation accuracy, cloud- and precipitation-affected observations must be removed.

In clear sky, for MWHTS water vapor channels, brightness temperature increases with frequency distance from the 183.31 GHz absorption line center because channels farther from the line center experience less water vapor opacity and can sound warmer, lower atmospheric layers. However, when clouds and/or rain are present, brightness temperatures of channels with peak weighting function heights above the clouds and/or rain are almost unaffected. For channels with peak weighting function heights near or below clouds and/or rain, water vapor opacity increases significantly as cloud water content or rain water content increases, causing weighting functions to shift upward and the channels to sound more from higher, colder atmospheric layers. Most upwelling radiation below the cloud or rain layer is absorbed or scattered away from the satellite direction, contributing to lower brightness temperatures relative to clear-sky conditions. In other words, clouds and rain cause brightness temperature depressions [?]. Because water vapor channels have different sensitivities to clouds and rain, brightness temperature differences between water vapor channels can serve as indicators of cloud and rain presence [?].

We calculate simulated brightness temperatures from the U.S. standard atmospheric profile, including temperature and humidity parameters at nadir, using RTTOV. Based on statistical characteristics of cloud liquid water content and cloud ice water content from ERA-Interim reanalysis covering the period from 1 February 2014 to 31 January 2015, cloud liquid water content varies from 0-4 mm and cloud ice water content from 0-2 mm. Because water clouds are mainly distributed below 400 hPa in the middle and lower atmosphere, cloud liquid water content is evenly distributed below 350 hPa. Since ice clouds are mainly distributed above 600 hPa in the upper atmosphere, cloud ice water content is evenly distributed between 100 hPa and 550 hPa. If the same cloud and/or rain content exists in different FOVs at nadir, different cloud cover can cause different brightness temperatures [?]. To avoid cloud cover impacts on simulations, cloud cover is set to 1.0 when clouds or rain are present. MWHTS channel 11, which is sensitive to about 350 hPa, is scarcely influenced by clouds and rain, so channel 11 is chosen as the reference channel. $\Delta i-11$ ($i=12, 13, 14, 15$) represents brightness temperature differences between channel i and channel 11.

Because cloud ice water content has little effect on MWHTS brightness temperatures [?], only the effects of cloud liquid water content on brightness temperature differences $\Delta i-11$ are investigated here. Figure 3 [Figure 3: see original paper] shows the sensitivities of brightness temperature differences $\Delta i-11$ to cloud liquid water content.

Figure 3 shows that brightness temperature differences $\Delta i-11$ decrease as cloud liquid water content increases, with the decreasing trend gradually leveling off. Because increasing cloud liquid water content causes water vapor channel weighting function peaks to shift upward, leading to brightness temperature depressions, the magnitude of depression is larger for frequencies farther from the 183.31 GHz line center [?]. However, precipitation generally occurs near the surface when cloud liquid water content exceeds 0.5 mm [?]. MWHTS water vapor channels are less affected by rain because their peak weighting function heights are above 900 hPa, which explains why the decreasing trend drops off. As shown in Figure 3, different brightness temperature differences $\Delta i-11$ exhibit different sensitivities depending on water vapor channel weighting function distributions. $\Delta 15-11$ and $\Delta 14-11$ will be used to screen out cloud- and precipitation-affected observations in our study due to their higher sensitivity compared to $\Delta 12-11$ and $\Delta 13-11$.

For $\Delta 15-11$, we first take 11.7 K as the reference threshold, corresponding to a cloud water content of 0.5 mm. Using RMSEclear from the absorption-based model in clear sky from Figure 2 as reference values, we adjust the reference threshold with a step length of 0.2 K. Adjustment stops when differences between RMSEall from the absorption-based model in all-weather conditions shown in Figure 2 and the reference values RMSEclear are within 1 K for temperature channels and within 1.5 K for water vapor channels. After some adjustment, we determine the threshold to be 12.5 K. For $\Delta 14-11$, the method for determining the threshold is the same, but the reference threshold is set to 7.1 K, yielding a determined threshold of 8.1 K. Based on these two thresholds, we develop three criteria: the first is $\Delta 15-11 > 12.5$, the second is $\Delta 14-11 > 8.1$, and the third is $\Delta 15-11 > 12.5$ and $\Delta 14-11 > 8.1$.

Applying these three criteria to the statistical analysis dataset to screen out cloud- and precipitation-affected observations, we obtain three additional statistical analysis datasets unaffected by clouds and rain: statistical analysis dataset 1, statistical analysis dataset 2, and statistical analysis dataset 3, containing 467,587, 470,051, and 452,683 collocated samples, respectively. A statistical analysis is performed on the RMSEs of simulations with respect to observed brightness temperatures after cloud filtering. Figure 4 [Figure 4: see original paper] shows RTTOV simulation accuracies for the filtered observations using these three criteria.

Figure 4 reveals that simulation accuracies for each channel have been obviously improved compared with RMSEall in Figure 2(b), except for channels 2, 3, and 4. MWHTS channels 15 and 14 are sensitive to levels around 900–800 hPa and 800–700 hPa, respectively. Clouds distributed in these different

atmospheric layers can be well detected by criterion 1 ($\Delta_{15-11} > 12.5$) and criterion 2 ($\Delta_{14-11} > 8.1$). However, clouds distributed around 900–800 hPa have little effect on channel 14 because its water vapor absorption is mainly located above cloud top. Clouds distributed around 800–700 hPa cause decreases in channel 15 brightness temperature due to increased absorption and scattering by ice and water droplets in clouds [?]. In other words, criterion 1 can screen out more cloud- and precipitation-affected observations than criterion 2. Another point to note is that criterion 2 has better sensitivity to high-level clouds than criterion 1, especially in atmospheric layers around 800–700 hPa. Based on this analysis, for water vapor channels, the simulation accuracy of channel 15 for criterion 1 is higher than that for criterion 2, but criterion 2 yields better simulation results for channels 13 and 14. For temperature channels 6, 7, 8, 9 and window channels 1 and 10, criterion 1 can improve simulation accuracies to varying degrees compared with criterion 2, with the largest improvement of 0.4 K in channel 9. This occurs because criterion 1 better identifies low-level clouds in the near-surface atmospheric layer where ice and water droplets can introduce scattering to brightness temperatures of channels whose peak weighting function heights are closer to the surface. Criterion 3 combines criteria 1 and 2, enabling it to detect clouds distributed between approximately 900 and 700 hPa, screen out more cloud- and precipitation-affected observations, and obtain higher simulation accuracies than the first two criteria, as shown in Figure 4. Through this analysis, criterion 3 is adopted as the cloud filtering method in this paper. Compared with AMSU-B, MWHTS provides greater sensitivity to clouds in different tropospheric layers due to the new water vapor channel at 183.31 ± 4.5 GHz. Consequently, cloud filtering criterion 3 demonstrates better performance in filtering out cloud- and precipitation-affected observations and effectively improving RTTOV simulation accuracy. To further evaluate the performance of the cloud filtering method in atmospheric parameter inversion and provide reference values for different retrieval techniques, retrieval experiments are conducted.

4.1. MLR Retrieval Algorithm

In essence, the MLR retrieval algorithm converts observations to atmospheric parameters through a linear regression model representing the linear relationship between radiometer observations and atmospheric parameters including temperature, humidity, and cloud water parameters. This model is given by [?]:

where x is the atmospheric state vector, p is pressure level, \bar{y} represents the statistical mean value, y is the observed radiance vector, i is the radiometer channel index, and D is the retrieval operator. In our study, collocated samples including atmospheric temperature profiles, humidity profiles, and MWHTS brightness temperatures from the statistical analysis dataset are used to calculate the retrieval operator D . We then retrieve MWHTS brightness temperatures from the testing dataset to obtain the retrievals.

4.2. ANN Retrieval Algorithm

ANN is a statistical regression model that, unlike MLR, can describe not only linear relationships between input and output data but also any nonlinear relationships in theory. Following the majority of publications on applying ANN to atmospheric parameter inversion, we choose a three-layer backpropagation ANN with one hidden layer for the retrieval problem. Figure 5 [Figure 5: see original paper] shows a schematic diagram of the selected network. Additional details on initialization, training, optimization, and other advanced topics can be found in Polyakov et al. and Yao et al. [?, ?].

In our study, MWHTS brightness temperatures from the statistical analysis dataset are taken as the input vector ($i=1\sim 15$ is the MWHTS channel index). Profile parameters including temperature and humidity, and surface parameters including 2 m air temperature, 2 m air dewpoint temperature, and 10 m wind speed from the statistical analysis dataset are taken as the output vector ($j=1\sim 37$ and $j=38\sim 74$ represent pressure levels from upper atmosphere to surface for temperature and humidity, respectively; $j=75, 76, 77, 78$ represent 2 m air temperature, 2 m air dewpoint temperature, 10 m U wind component, and 10 m V wind component). The input and output vectors comprise the training samples, and the steepest descent method is selected for the training phase. Based on extensive testing, a hidden layer with 16 hidden nodes was found to be optimal in our study. Weights and biases are determined by training on 90% of the training samples, while the remaining 10% are used for validation to determine when to stop training.

4.3. 1DVAR Retrieval Algorithm

The 1DVAR method is a typical representative of physical retrieval approaches that obtains atmospheric parameters by directly solving the radiative transfer equation. The algorithm consists primarily of two components: the radiative transfer model used to calculate simulated brightness temperatures and brightness temperature gradients, and a scheme for minimizing a cost function that weights the relative contributions of satellite observations and background information. If errors in background information and satellite observations are unbiased, uncorrelated, and have Gaussian distributions, the atmospheric state vector can be solved by minimizing the following cost function:

where B is the background covariance matrix, x_b is the background state vector, R is the sum of covariance error in simulated brightness temperature and instrument channel noise, T represents matrix transpose, H is the forward operator that simulates satellite observations at atmospheric state vector x , and y is the observed brightness temperature vector. Setting the cost function gradient to zero yields the optimal estimate:

where n is the iteration index, the starting point x_0 is the first-guess profile, and H' is the tangent linear function of H at point x_n . Equation (3) shows that retrieval algorithm parameters including background covariance matrix B ,

background state vector x_b , error covariance matrix R , first-guess profile, and system bias must be determined before retrieval.

A priori information including background covariance matrix and background state vector directly affects retrieval accuracy in the physical retrieval process, aiming to constrain retrievals within physically realistic solutions during iteration in equation (3). In our study, atmospheric temperature and humidity profiles from ERA-Interim reanalysis covering the period from 1 February 2014 to 31 January 2015 are used to compute background covariance matrix B , following Boukabara et al. for detailed computation methods [?]. The means of atmospheric parameters used to compute the background covariance matrix are taken as the background state vector x_b . Outputs from the BP ANN retrieval method in Section 4.2 are taken as first-guess profiles in our study.

For biases between simulations and observations, statistical regression correction is given by [?]:

where I^* is the corrected MWHTS brightness temperature, I is the uncorrected MWHTS brightness temperature, a is the slope, b is the intercept, $i=1\sim 15$ is the MWHTS channel index, and $j=1\sim 98$ is the MWHTS scan position index. We determine correction coefficients through statistical analysis of simulations and observations from the statistical analysis dataset, then correct observed brightness temperatures from the testing dataset.

For error covariance matrix R , assuming measurements in one channel are unrelated to those in others, only diagonal elements of R are used. Differences between simulations and observations from the statistical analysis dataset and in-flight sensitivities from Table 1 are utilized.

The convergence criterion adopted in our retrieval system stops iteration in equation (3) when the relative difference in cost function between two iterations is less than 0.01, with a maximum of 10 iterations. If iterations reach 10, retrieval is set to the first guess. The retrieval quality control criterion discards observations if residuals between corrected brightness temperatures and RTTOV-simulated values using the first guess exceed 20 K in any channel.

5. Validation and Analysis of Retrieval Results

MWHTS channel weighting function distributions shown in Figure 1 indicate that temperature channels are not sensitive to the top of the atmosphere and water vapor channels are not sensitive to levels above 300 hPa. Therefore, we verify retrievals at levels from 1000 to 50 hPa for temperature and 1000 to 250 hPa for relative humidity. MWHTS profile retrieval results are evaluated by mean error (ME) and RMSE with respect to ECMWF ERA-Interim reanalysis taken as truth. ME and RMSE are defined as follows [?]:

where x_{MWHTS} is the MWHTS retrieval, x_{ECMWF} is the ECMWF reanalysis, and N is the total number of comparisons.

Applying the three cloud filtering criteria suggested in our study to the testing dataset in Section 2.2 to screen out cloud- and precipitation-affected observations, we obtain three additional testing datasets unaffected by clouds and rain: testing dataset 1, testing dataset 2, and testing dataset 3, containing 35,211, 35,517, and 34,021 collocated samples, respectively. RTTOV simulation calculations using atmospheric parameters from these four testing datasets show that simulation accuracies for each channel have been obviously improved after filtering cloud- and precipitation-affected observations in the testing dataset, with criterion 3 achieving higher simulation accuracies than criteria 1 and 2, consistent with conclusions from the statistical analysis datasets in Section 3. To further evaluate cloud filtering method performance in atmospheric parameter inversion, observed brightness temperatures from these four testing datasets are used to retrieve atmospheric temperature and humidity profiles.

5.1. Retrieval Results and Analysis of MLR Method

The statistical analysis dataset in Section 2.2 and the three statistical analysis datasets in Section 3 are used to separately calculate retrieval operator D, with retrieval results shown in Figure 7 [Figure 7: see original paper]. For temperature, retrieval MEs of brightness temperatures in the testing dataset before and after filtering are nearly equal, reaching maximum values of about 2.8 K at 70 hPa and 850 hPa, and minimum values of about 0 K between 200 hPa and 300 hPa, with mean errors within 1.5 K at other levels. However, retrieval RMSEs with filtering are smaller than those without filtering between 400 hPa and 1000 hPa, and criterion 3, which performs better in filtering cloud- and precipitation-affected observations, yields the highest retrieval accuracy. RMSEs are almost the same between 50 hPa and 400 hPa because temperature channels sensitive to the upper atmosphere are nearly unaffected by clouds and rain. Retrieval RMSEs reach a minimum of 1.08 K near 200 hPa and a maximum of about 3.4 K near 850 hPa. For humidity, retrieval MEs of brightness temperatures before and after filtering are nearly equal (similar to temperature retrieval results), at about 9% between 500 hPa and 600 hPa, and within 5% at other levels. Retrieval accuracies with filtering are higher than without filtering between 300 hPa and 800 hPa because the peak weighting function heights of water vapor sounding channels are mainly distributed in this pressure range. Similar to temperature retrieval accuracy, criterion 3 achieves the highest humidity retrieval accuracies, with RMSEs reaching a maximum of 19.5% near 800 hPa.

For the cloud filtering criteria developed in this paper, criterion 3 makes greater strides in improving retrieval accuracy of multiple linear regression, which uses linear relationships between atmospheric parameters and satellite observations, though criteria 1 and 2 also improve retrieval accuracy to some extent. The filtering methods screen out cloud- and precipitation-affected brightness temperatures, further reducing nonlinearity caused by clouds and rain and conforming to the linear model used by the retrieval approach. However, it is worth noting that relationships between atmospheric parameters and satellite observations

are nonlinear, which limits the retrieval accuracy of multiple linear regression.

5.2. Retrieval Results and Analysis of ANN Method

The statistical analysis dataset in Section 2.2 and the three statistical analysis datasets in Section 3 are used to separately train the ANN and determine weights and biases, with retrieval results shown in Figure 8 [Figure 8: see original paper]. For temperature, retrieval MEs of brightness temperatures before and after filtering with the three criteria are nearly equal, both within 1 K. Retrieval RMSEs with filtering are smaller than without filtering between 300 hPa and 200 hPa, but higher between 900 hPa and 800 hPa, and nearly equal at other levels. Retrieval accuracies reach a maximum of about 2.5 K at 100 hPa and are within 1.8 K at other levels. For humidity, retrieval MEs of brightness temperatures before and after filtering with the three criteria are all within about 2%, and their RMSEs are nearly equal at all pressure levels, reaching a maximum of about 15% at 800 hPa.

Analysis of ANN retrieval results shows that the cloud filtering method using the three criteria cannot improve retrieval accuracy and even worsens it at some levels. This occurs because ANN's strong nonlinear mapping capability can describe the nonlinear relationships between satellite observations and atmospheric parameters caused by clouds and rain. However, the number of samples trained in the statistical analysis dataset differs with and without filtering, and ANN algorithm performance is affected by the number of training samples, which may lead to worse retrieval accuracies.

5.3. Retrieval Results and Analysis of 1DVAR Method

The statistical analysis dataset in Section 2.2 and the three statistical analysis datasets in Section 3 are used to separately determine correction coefficients for biases between observations and simulations. Observed brightness temperatures in the four testing datasets are corrected and retrieved, with results shown in Figure 9 [Figure 9: see original paper]. For temperature, retrieval MEs of brightness temperatures before and after filtering are nearly equal above 400 hPa in the upper atmosphere, but at other levels, retrieval MEs with filtering are obviously smaller than without filtering, all within 0.4 K. Retrieval MEs with the three criteria are nearly equal. Retrieval accuracies with filtering are higher than without filtering between 400 hPa and 1000 hPa where temperature channels 7, 8, and 9 are sensitive, and criterion 3 achieves the highest accuracy of 1.7 K between 150 hPa and 1000 hPa. For humidity, retrieval MEs and RMSEs with filtering are significantly smaller than without filtering, especially between 200 hPa and 800 hPa where water vapor channels are sensitive, with MEs close to zero. Criterion 3 improves humidity retrieval accuracies by up to 11.2% and achieves the highest retrieval accuracy of about 19%.

From a physical perspective, the 1DVAR method retrieves atmospheric parameters by directly solving the radiative transfer equation. As seen in equation (3),

in addition to prior information, differences between observations and simulations directly affect retrieval accuracies. Filtering out cloud- and precipitation-affected observations can improve retrieval accuracies of atmospheric temperature and humidity profiles. Compared with criteria 1 and 2, criterion 3 can effectively screen out more cloud- and precipitation-affected observations and better improve retrieval accuracies.

5.4. Comparison Results of Three Retrieval Methods

Based on the above analysis, for temperature retrievals, MLR method shows the largest MEs and RMSEs, while 1DVAR retrieval accuracies are slightly better than ANN. For humidity retrievals, both ANN and 1DVAR methods have MEs nearly zero at all levels. ANN retrieval accuracies are the best, while other methods have similar retrieval accuracies. However, we should note that 1DVAR retrieval accuracies depend on many factors evident from equation (3). Therefore, several approaches can further improve retrieval accuracy, such as optimizing bias correction methods, improving first-guess profiles, and further controlling satellite observation quality.

From the perspective of algorithm complexity, statistical inversion methods including multiple linear regression and ANN are simple and straightforward. Generally, a good statistical model corresponds to high retrieval accuracy, as exemplified by the ANN trained in our study. However, it is worth noting that retrieval algorithm stability is very important for operational applications of atmospheric parameter retrieval.

In our study, we change testing samples to examine the stability of these three retrieval methods. MWHTS brightness temperatures covering the period from 1 to 31 May 2015 are selected, and collocation criteria with ERA-Interim reanalysis are the same as in Section 2, yielding 43,137 collocated samples in all-weather conditions and 38,034 samples after filtering by criterion 3. These 38,034 brightness temperature samples are used to retrieve atmospheric temperature and humidity profiles using the three retrieval methods, with results shown in Figure 10 [Figure 10: see original paper]. For the MLR method, compared with retrieval results using criterion 3 in Figure 7, different retrieval results are obtained due to different brightness temperature samples retrieved. Temperature retrieval MEs and RMSEs are both significantly smaller than in Figure 7 between 400 hPa and 1000 hPa, with temperature retrieval accuracies improved by about 1.0 K. Humidity retrieval accuracies decrease by 5% near 500 hPa. For the ANN method, compared with retrieval results using criterion 3 in Figure 8, temperature retrieval MEs and RMSEs both increase significantly, with temperature retrieval accuracies declining about 0.4 K between 400 hPa and 800 hPa. Humidity retrieval accuracies decline 5% near 800 hPa. However, for the 1DVAR method, compared with retrieval results using criterion 3 in Figure 9, retrieval accuracies remain nearly unchanged. These comparison analyses demonstrate that MLR and ANN retrieval methods have great dependence on retrieved brightness temperatures, which have very little impact on 1DVAR

retrieval accuracies.

6. Summary and Discussion

MWHTS water vapor channels have high sensitivity to clouds and rain and can detect different tropospheric cloud layers well due to new channels at 183.31 ± 4.5 GHz. We have derived a cloud filtering method to screen out cloud- and precipitation-affected observations using brightness temperature differences between water vapor channels. Based on sensitivities of MWHTS water vapor channels 14 and 15 to cloud water content, and taking RMSEs between observation and simulation by RTTOV in clear sky as reference values, we developed cloud filtering criteria: $\Delta_{15-11} > 12.5$ and $\Delta_{14-11} > 8.1$. The proposed cloud filtering method can improve agreement between simulated and observed brightness temperatures. To evaluate the cloud filtering method in the retrieval process, brightness temperatures before and after filtering are used to retrieve atmospheric temperature and humidity profiles through different retrieval methods. Retrieval results show that the cloud filtering method can effectively improve retrieval accuracies of MLR and 1DVAR, which depend on the nonlinearity of the radiative transfer model, but has little effect on ANN because ANN can represent the nonlinear relationship between atmospheric parameters and satellite observations. Comparison of retrieval results using the three methods shows that MLR accuracies, which only describe linear relationships between atmospheric parameters and satellite observations, are worst. ANN and 1DVAR can describe nonlinear relationships and have similar temperature retrieval accuracies, but 1DVAR humidity retrieval accuracy is lower than ANN. From an operational application perspective, retrieval algorithm stability is tested, and results show that 1DVAR is the best choice.

This paper focuses on improving differences between simulated and observed brightness temperatures. Generally, differences between simulations and observations are caused by: satellite instrument issues (e.g., poor calibration or adverse environmental effects), the radiative transfer model linking atmospheric parameters to sounder-measured radiation (e.g., spectroscopy errors, simplistic modeling of sounder viewing geometry, or inaccurate modeled scattering effects), and errors in atmospheric parameters input to the radiative transfer model. However, filtering out cloud- and precipitation-affected observations as a pre-processing step is very important for subsequent bias correction that can further improve retrieval accuracy.

It is worth noting that the proposed cloud filtering method does not account for limb darkening for MWHTS, which is a cross-track scanning instrument. However, from a statistical computation perspective in this study, effects of limb darkening on simulation accuracies may be partially filtered out using the cloud filtering method. The cloud filtering method may be optimized by deriving cloud filtering criteria based on statistical analysis for different MWHTS viewing geometries, which is our next investigation. Additionally, we only study MWHTS measurements over a portion of ocean in this paper; therefore, devel-

oping a cloud filtering method for MWHTS observations over the global range is also a future research direction.

Acknowledgments: The authors thank ECMWF for providing reanalysis data and NSMC for providing MWHTS brightness temperature data. This work was supported by the Key Fostering Project of National Space Science Center, Chinese Academy of Sciences (No. Y62112f37s) and the National 863 Project of China (No. 2015AA8126027).

References

- [1] MORADI I, FERRARO R R, SODEN B J, et al. Retrieving layer-averaged tropospheric humidity from advanced technology microwave sounder water vapor channels [J]. *IEEE Trans on Geosci Remote Sens*, 2015, 53(12): 6675-6688.
- [2] NAVAS-GUZMAN F, STAHLI O, KAMPFER N. An integrated approach toward the incorporation of clouds in the temperature retrievals from microwave measurements [J]. *Atmos Meas Tech*, 2014, 7(6): 1619-1628.
- [3] PERRO C, LESINS G, DUCK T, et al. A microwave satellite water vapor column retrieval for polar winter conditions [J]. *Atmos Meas Tech*, 2016, 9(8): 2241-2252.
- [4] POLYAKOV A, TIMOFEYEV Y M, VIROLAINEN Y. Comparison of different techniques in atmospheric temperature-humidity sensing from space [J]. *Int J Remote Sensing*, 2014, 35(15): 5899-5912.
- [5] TAN Hao-bo, MAO Jie-tai, CHEN Huan-huan, et al. A study of a retrieval method for temperature and humidity profiles from microwave radiometer observations based on principal component analysis and stepwise regression [J]. *J Atmos Ocean Tech*, 2011, 28(3): 378-389.
- [6] MATZLER C, MORLAND J. Refined physical of integrated water vapor and cloud liquid for microwave radiometer data [J]. *IEEE Trans on Geosci Remote Sens*, 2009, 47(6): 1585-1594.
- [7] ISHIMOTO H, OKAMOTO K, OKAMOTO H, et al. One-dimensional variational (1D-Var) retrieval of middle to upper tropospheric humidity using AIRS radiance data [J]. *J Geophys Res*, 2014, 119(12): 7633-7654.
- [8] LI Juan, ZOU Xiao-lei. A quality control procedure for FY-3A MWTS measurements with emphasis on cloud detecting using VIRR cloud fraction [J]. *J Atmos Ocean Tech*, 2013, 30(8): 1704-1715.
- [9] MULLER B M, FUELBERG H E, XIANG Xu-wu. Simulations of the effects of water vapor, cloud liquid water, and ice on AMSU moisture channel brightness temperatures [J]. *J Appl Meteorol*, 1994, 33(10): 1133-1154.
- [10] BURNS B A, WU Xiang-qian, DIAK G R. Effects of precipitation and cloud ice on brightness temperatures in AMSU moisture channels [J]. *IEEE Trans on Geosci Remote Sens*, 1997, 35(6): 1429-1437.

- [11] SKOFRONICK-JACKSON G M, GASIEWSKI A J, WANG J R. Influence of microphysical cloud parameterizations on microwave brightness temperatures [J]. *IEEE Trans on Geosci Remote Sens*, 2002, 40(1): 187-196.
- [12] GREENWALD T J, CHRISTOPHER S A. Effect of cold clouds on satellite measurements near 183 GHz [J]. *J Geophys Res*, 2002, 107(D13): 4170.
- [13] BENNARTZ R, BAUER P. Sensitivity of microwave radiances at 85-183 GHz to precipitating ice particles [J]. *Radio Sci*, 2003, 38(4): 8075-8090.
- [14] HONG Gang, HEYGSTER G, MIAO Jun-gang, et al. Detection of tropical deep convective clouds from AMSU-B water channels measurements [J]. *J Geophys Res*, 2005, 110(5): D05 205-1-D05 205-15.
- [15] BUEHLER S A, KUVATOV M, SREEREKHA T R, et al. A cloud filtering method for microwave upper tropospheric humidity measurements [J]. *Atmos Chem Phys*, 2007, 7(21): 5531-5542.
- [16] HOCKING J, RAYER P, RUNDLE D, et al. RTTOV v11 users guide. NWP-SAF, Exeter, UK, 2014.
- [17] GUO Yang, LU Nai-meng, QI Cheng-li, et al. Calibration and validation of microwave humidity and temperature sounder onboard FY-3C satellite [J]. *Chinese J Geophys-Ch*, 2015, 58(1): 20-31.
- [18] BAO Jing-hua. The principle of FY-3C satellite microwave humidity and temperature sounder and preliminary analysis on its in-orbit performance [D]. Beijing: Chinese Academy of Science, 2014.
- [19] DEED P, UPPALA S M, SIMMONS A J, et al. The ERA-Interim reanalysis: configuration and performance of the data assimilation system [J]. *Q J Roy Meteor Soc*, 2011, 137(656): 553-597.
- [20] GEER A J, BAUER P, DELL C W. A revised cloud overlap scheme for fast microwave radiative transfer in rain and cloud [J]. *J Appl Meteorol Clim*, 2009, 48(11): 2257-2270.
- [21] HONG Gang, HEYGSTER G, MIAO Jun-gang, et al. Sensivity of microwave brightness temperatures to hydrometeors in a tropical deep convective cloud system at 89-190 GHz [J]. *Radio Sci*, 2005, 40(4): RS4003.
- [22] GUO Yang, LU Nai-meng, GU Song-yan. Simulation of the radiometric characteristics of 118 GHz and 183 GHz channels for FY-3C new microwave radiometer sounder [J]. *J Infrared Millim W*, 2014, 33(5):481-491.
- [23] KARSTENS U, SIMMER C, RUPRECHT E. Remote sensing of cloud liquid water [J]. *Meteorol Atmos Phys*, 1994, 54: 159-168.
- [24] CHEN Hao, JIN Ya-qiu. Data validation of Chinese microwave FY-3A for retrieval of atmospheric temperature and humidity profiles during Phoenix typhoon [J], *Int J Remote Sens*, 2008, 32(23): 8541-8554.

- [25] YAO Zhi-gang, CHEN Hong-bin, LIN Long-fu. Retrieving atmospheric temperature profiles from AMSU-A data with neural networks [J]. Adv Atmos Sci, 2005, 22(4): 606-616.
- [26] LIU Quan-hua, WENG Fu-zhong. One-dimensional variational retrieval algorithm of temperature, water vapor, and cloud water profiles from advanced microwave sounding unit(AMSU) [J]. IEEE Trans on Geosci Remote Sens, 2005, 43(5): 1087-1095.
- [27] BOUKABARA S A, GARRETT K, CHEN Wan-chun, et al. MIRS: An all-weather 1DVAR satellite data assimilation and retrieval system [J]. IEEE Trans on Geosci Remote Sens, 2011, 49(9): 3249-3272.
- [28] LI Jun, WOLF W W, MENZEL W P, et al. Global soundings of the atmosphere from ATOVS measurements: the algorithm and validation [J]. J Appl Meteorol, 2000, 39(8): 1248-1268.
- [29] MATHUR A K, GANGWAR R K, GOHIL B S, et al. Humidity profile retrieval from SAPHIR on-board the Megha-Tropiques [J]. Curr Sci India, 2013, 104(12): 1650-1655.

Author Biographies:

HE Qiurui (1986-), male, from Huaxian, Henan Province, is a Ph.D. candidate at the National Space Science Center, Chinese Academy of Sciences. His research focuses on satellite data processing and microwave radiometer atmospheric parameter retrieval applications. Phone: 18811793515. Email: heqiurui@126.com

WANG Zhenzhan (1969-), male, from Qinglong, Hebei Province, is a professor. His research focuses on microwave remote sensing technology and applications, including: fully polarized microwave radiometry technology, calibration and application fundamentals, microwave radiometer ocean-atmospheric parameter retrieval applications, microwave remote sensor calibration/validation research, and terahertz atmospheric detection mechanisms. Email: wangzhenzhan@mirslab.cn

HE Jieying (1984-), female, from Tianjin, is an associate professor. Her research focuses on ground-based and satellite-borne microwave radiometer data processing and applications. Email: hejieying@mirslab.cn

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv – Machine translation. Verify with original.