Research on the selection method of FY-3D/MWHTS clear sky observation data based on neural network

Qiurui He1*, Xiao Guo2, Deguang Li1, Yanling Jin1, Lanjie Zhang3 and Ruanyu Zhang4

1 Luoyang Normal University, Luoyang, Henan, 471934, China.
2 Hebi Institute of Engineering and Technology, Hebi, 451191, China.
3 Beijing Information Science and Technology University, Beijing, 100101, China.
4 Shanghai Institute of Aerospace Measurement and Control Communications, Shanghai, 201109, China.
*Corresponding author’s e-mail: heqiurui@lynu.edu.cn

Abstract. The selection of clear sky data in space-borne remote sensing data is very important for its data application. For FY-3D satellite microwave humidity and temperature sounder (MWHTS), an inversion system of atmospheric cloud water content by MWHTS is established based on neural network. The cloud water content inversion value is used to select clear sky data from MWHTS observation data. The experimental results show that FY-3D/MWHTS clear sky data selection method based on neural network can effectively select MWHTS observation data, thus improving the simulation brightness temperatures accuracy of MWHTS by radiative transfer model. This method can be used to select clear sky data by using space-borne observation data itself. It is easy to operate and has important practical value for climate change research, numerical weather forecast, etc., based on space-borne observation data.

1. Introduction
At present, the simulation accuracy of brightness temperatures observed by space-borne microwave radiometer in clear air is much higher than that in cloudy and rainy atmosphere. Consequently, The calculation accuracy of space-borne microwave radiometer simulation brightness temperatures has an important impact on the data application of microwave radiometer. The extraction of clear sky brightness temperatures from space-borne microwave radiometer directly determines the retrieval accuracy of atmospheric parameters, and is closely related to the assimilation application effect of space-borne radiometer observation brightness temperatures [1].

To deal with the forward modeling of space-borne microwave radiometer, setting the threshold value of atmospheric relative humidity and using infrared cloud products are the common methods to select clear sky observed brightness temperatures from space-borne microwave radiometer observation brightness temperatures. However, the space-borne microwave radiometer observation data must be matched with the atmospheric relative humidity or infrared cloud products before selecting the clear sky data, which seriously affects the direct application of the data. Therefore, the two methods are not applicable to the inversion of microwave radiometer observation data [2-3]. In the inversion application, it is usually possible to determine the clear sky according to the statistical relationship between the brightness temperatures observed by the microwave radiometer channel. However, the
accuracy rate of the method for judging the clear sky data is low. On the other hand, the establishment of the threshold value in this method is strongly dependent on the local atmospheric state, and thus it is not suitable for large-scale clear sky observation [4].

The microwave humidity and temperature sounder (MWHTS) is an important payload on FY-3C and FY-3D satellites. It is the first microwave radiometer integrating hygrometer and thermometer in the world. MWHTS has eight temperature detection channels (channels 2-9), five humidity detection channels (channels 11-15) and two window area channels (channels 1 and 10), which can realize simultaneous detection of atmospheric temperature, water vapor parameters and surface parameters [5]. Based on the setup characteristics of radiometer channel, MWHTS has high sensitivity to water vapor parameters. Therefore, this work proposed a method of selecting MWHTS clear sky data based on neural network. Firstly, BP neural network is used to invert the atmospheric cloud water content, which is directly related to the judgment of atmospheric clear sky. Then, MWHTS clear sky brightness temperatures data can be extracted according to the inversion value of cloud water content.

2. Selection method of MWHTS clear sky observation data based on neural network

2.1. Research materials and models

MWHTS is an important payload for FY-3D. MWHTS is a full power microwave radiometer with superheterodyne receiver. Its receiver has eight temperature detection channels, five humidity detection channels and two window area channels, which can realize simultaneous detection of atmospheric and surface parameters. MWHTS adopts mechanical scanning method. The scanning width is about 2700km, and the scanning period is 2.667 seconds. Each scanning line has 98 scanning points, and each pixel corresponds to a scanning angle. The MWHTS channel parameter setting and detailed instrument characteristics can be referred to the literature [6].

The research data used in this work are FY-3D/MWHTS observation brightness temperatures climatology data set and ECMWF ERA Interim reanalysis data set. Among them, the atmospheric parameters used in reanalysis data set are temperature profile, humidity profile, cloud water profile, 2m temperature, 2m humidity, surface pressure and 10m wind speed. The atmospheric pressure stratification of the profile is 37 layers. The time range is from September 2018 to August 2019, and the geographical range is 25°N-45°N and 160°E-220°E. The data preprocessing process is as follows. According to the matching rules of time error less than 10 minutes and latitude and longitude error less than 0.1°, the brightness temperatures observed in MWHTS was matched with atmospheric parameters. Then, the matched atmospheric parameters were input into RTTOV to calculate the MWHTS simulation brightness temperatures. Finally, a matched data set (1060162 groups) including MWHTS observed brightness temperatures, MWHTS simulated brightness temperatures and atmospheric parameters was established. 80% of the matched data set was randomly selected to establish the analysis data set (848129 groups), and the remaining 20% was the validation data set (212033 groups).

2.2. Establishment of MWHTS clear sky threshold based on cloud water content

Taking the cloud water content of 0 as the strict clear sky threshold and selecting the strict clear sky data in the analysis data set, 5830 sets of matching data can be obtained. By calculating the root mean square error between MWHTS observed brightness temperatures and MWHTS simulated brightness temperatures in strict clear sky data, strict clear sky accuracy can be obtained. All weather accuracy can be obtained by calculating and analyzing the root mean square error between all MWHTS observed brightness temperatures and MWHTS simulated brightness temperatures in the data set. The calculation accuracy of MWHTS simulated brightness temperatures under strict clear sky conditions and all weather conditions (i.e. strict clear sky accuracy and all weather accuracy) is shown in figure 1. It can be seen from figure 1 that the strict clear sky accuracy is higher than the all weather accuracy in the 15 channels of MWHTS, especially in channel 1 and channel 10 in window area, as well as channels 7-9 and 13-15, which are close to the surface. Due to the detection path is not affected by cloud and rain, the simulation accuracy is greatly improved.
Figure 1. Calculation accuracy of MWHTS simulated brightness temperatures under the strict clear sky and all weather conditions

Figure 2. The difference between the adjusted clear sky accuracy and the strict clear sky accuracy

The strict clear sky threshold is gradually increased with step length of 0.01mm to form the adjusted clear sky threshold. In the analysis data set, the matching data less than the adjusted clear sky threshold is selected. Also, the root mean square error between the MWHTS observed brightness temperatures and the MWHTS simulated brightness temperatures is calculated, which is recorded as the adjustment of clear sky accuracy. When the difference between the adjusted clear sky precision and the strict clear sky precision of any channel in MWHTS temperature detection channel (channels 2-9) and water vapor detection channel (channels 11-15) increases to 1K, the adjusted clear sky threshold is the clear sky threshold. When the clear sky threshold is adjusted to 0.11mm, the difference between the adjusted clear sky accuracy and the strict clear sky accuracy is shown in figure 2. The accuracy difference in channel 8 is 1K, and thus the clear sky threshold is 0.11mm.

2.3. Selection method of MWHTS clear sky data based on neural network

In recent years, data processing algorithms based on neural network have developed rapidly in the field of remote sensing. Among them, BP neural network based on error back propagation algorithm is
widely used in atmospheric parameter inversion because of its strong nonlinear mapping capability. According to the research requirements of this work, the three-layer BP neural network is selected to carry out the inversion research of atmospheric temperature profile. The schematic diagram of three-layer BP neural network is shown in figure 3.

In figure 3, there are $L$ nodes in the input layer, representing the length of the input vector $X$. The input layer does not carry out any calculation, and each node is connected with $M$ nodes of the hidden layer. Each node of the hidden layer performs nonlinear operations and outputs the calculation results to the output layer. The output layer contains $N$ nodes, and the output vector $Z$ is obtained by summing the weights of the output results in the hidden layer. The initialization, training and optimization of BP neural network are detailed in reference.

In this work, the training set of neural network is generated by using the data in the analysis data set. Among them, the observed brightness temperatures is taken as the input vector $X_L$. The cloud water content is taken as the output vector $Z_N$. The MWHTS brightness temperatures of 848129 sets in the analysis data set is used as the input of the algorithm model, and the corresponding cloud water content is taken as the output of the algorithm model. The number of hidden layer neurons changed from 5 to 50, and the neural network was trained respectively. When the number of neurons in the hidden layer is 43, the mean square difference of cloud water content predicted by neural network model is the smallest compared with the 46 trained neural network models, which is 0.0091. Then, when the number of hidden layer neurons is 43, the neural network model is denoted as the best model for retrieving cloud water content at brightness temperatures observed by MWHTS.

3. Experimental results and analysis

The 212033 group of MWHTS brightness temperatures in the validation data set was input into the best inversion model of cloud water content inversion by MWHTS observation brightness temperatures, and the cloud water content inversion value of 212033 group was obtained. Taking the cloud water content in the validation data set as the real value, the average deviation between the true cloud water content and the cloud water content inversion value is 0.0022 mm.

In the cloud water content inversion value, the clear sky threshold value of 0.11mm was compared with the cloud water content inversion value. 76332 sets of MWHTS brightness temperatures corresponding to the cloud water content inversion value less than the clear sky threshold are selected, which are recorded as the selection data of MWHTS clear sky observation brightness temperatures. In the validation of the cloud water content real value in the analysis data set, the clear sky threshold of 0.11mm was compared with the true cloud water content value. 76332 sets of MWHTS brightness temperatures corresponding to the cloud water content real value less than the clear sky threshold are selected, which are recorded as the real data of MWHTS clear sky observation brightness temperatures. Among them, the selection data of brightness temperatures in MWHTS clear sky observation and the real data of brightness temperatures in MWHTS clear sky observation are 73590 groups. In other words, the method based on cloud water content inversion can select 73590 sets of real MWHTS clear sky brightness temperatures data (i.e., 92.84% of MWHTS clear sky brightness temperatures can be selected by this method).
Figure 4. The difference of simulated brightness temperatures accuracy between MWHTS selected clear sky and real clear sky

Based on the neural network, the difference between the simulated brightness temperatures accuracy of the clear sky observation data selected by the MWHTS clear sky observation data selection method and the simulated brightness temperatures accuracy of the MWHTS real clear sky observation data is shown in figure 4.

It can be seen from figure 4 that the calculation accuracy of MWHTS simulated brightness temperatures in selected clear sky data is close to that in real clear sky data in temperature detection channels 2-9 and water vapor detection channels 11-15, and the difference is kept within 0.3K. Although the selection method of MWHTS clear sky observation data based on neural network is processed, there is still a big difference between channel 1 and channel 10. The comparison between the calculation accuracy of MWHTS simulated brightness temperatures in clear sky data and that of MWHTS simulated brightness temperatures under all weather conditions is shown in figure 5.

Figure 5. The calculation accuracy comparison of MWHTS simulated brightness temperatures under the selected clear sky and all weather conditions

It can be seen from figure 5 that the simulation brightness temperatures accuracy of the clear sky data selected by the neural network MWHTS observation data selection method is higher than the simulation brightness temperatures accuracy under all weather conditions. Especially in channel 1 and channel 10, the accuracy of simulated brightness temperatures is improved significantly. In the
temperature detection channels 2-5, the radiation contributions of these channels are all from the upper atmosphere, and the clouds and rain are usually distributed in the middle and low altitude atmosphere. Therefore, although the MWHTS clear sky observation data selection method based on neural network can effectively filter out the cloud and rain observation data, the calculation accuracy of simulated brightness temperatures for temperature detection channels 2-5 has not been improved. However, the radiation contribution of temperature detection channels 6-9 and humidity detection channels 11-15 come from the middle and low altitude atmosphere. Therefore, filtering out the cloud and rain influence observation data can effectively improve the simulation brightness temperatures accuracy of these detection channels.

4. Conclusion
The selection method of FY-3C/MWHTS clear sky observation data based on neural network firstly uses the information of MWHTS observation data to invert atmospheric cloud water content and obtain the inversion value of cloud water content. Then, a clear sky threshold based on cloud water content is established. Finally, the clear sky data of MWHTS observation data is selected based on the clear sky threshold value and cloud water content inversion value, and the accuracy rate is 92.84%. Compared with the simulated brightness temperatures under all weather conditions, the accuracy of simulated brightness temperatures in clear sky data is significantly improved. However, the selection effect of this method for MWHTS clear sky observation data depends on the accuracy of MWHTS inversion cloud water content. The next step is to use MWHTS data to invert cloud water content with high accuracy and improve the accuracy of MWHTS clear sky data selection.

Acknowledgments
This work was supported by The National Natural Science Foundation (NO. 41901297), Science and Technology Plan of Henan Province (NO. 202102310017, NO. 192102210259, NO. 202102210370) and Key Research Projects for the Universities of Henan Province (NO. 20A170013).

References
[1] Sahoo, S., Bosch-Liuis, X., Reising, S. C., Vivekanandan, J. (2015) Radiometric information content for water vapor and temperature profiling in clear skies between 10 and 200 GHz. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 8: 859-871.
[2] Karstens, U., Simmer, C., Ruprecht, E. (1994) Remote sensing of cloud liquid water. Meteorology and Atmospheric Physics, 54: 157-171.
[3] Ishimoto, H., Okamoto, K., Okamoto, H., Sato, K. (2014) One-dimensional variational (1D-Var) retrieval of middle to upper tropospheric humidity using AIRS radiance data. Journal of Geophysical Research: Atmospheres, 119: 7633-7654.
[4] Buehler, S. A., Kuvatov, M., Srereekha, T. R., John, V. O., Rddb, B., Eriksson, P., Notholt, J. (2007) A cloud filtering method for microwave upper tropospheric humidity measurements. Atmospheric Chemistry and Physics, 7: 5531-5542.
[5] He, Q. R., Wang, Z. Z., He, J. Y. (2016) Bias Correction for Retrieval of Atmospheric Parameters from the Microwave Humidity and Temperature Sounder Onboard the Fengyun-3C Satellite. Atmosphere, 7: 1-20.
[6] He, Q. R., Wang, Z. Z., He, J. Y. (2017) Retrieval of clear sky temperature and humidity profiles over land using measurements of FY-3C/MWHTS. Journal of Remote Sensing, 21: 27-39.