TELKOMNIKA, Vol. 11, No. 12, December 2013, pp. 7174~7181
e-ISSN: 2087-278X

Retrieving Atmospheric Precipitable Water Vapor using Artificial Neural Network Approach

Wang Xin*, Deng Xiaobo, Zhang Shenglan
Chengdu University of Information Technology, Chengdu, China, 610225
Key Laboratory of Atmosphere Sounding, China Meteorological Administration, Chengdu, China, 610225
*Corresponding author, e-mail: wxwell@126.com

Abstract

Discussing of water vapor and its variation is the important issue for synoptic meteorology and meteorology. In physical Atmospheric, the moisture content of the earth atmosphere is one of the most important parameters; it is hard to represent water vapor because of its space-time variation. High-spectral resolution Atmospheric Infrared Sounder (AIRS) data can be used to retrieve the small scale vertical structure of air temperature, which provided a more accurate and good initial field for the numerical forecasting and the large-scale weather analysis. This paper proposes an artificial neural network to retrieve the clear sky atmospheric radiation data from AIRS and comparing with the AIRS Level-2 standard product, and gain a good inversion results.

Keywords: *neural network, AIRS, precipitable water vapor, retrieval*

1. Introduction

Water vapor is not only the most active component and abundant greenhouse gases in atmosphere, but also a key factor to affect climatic variation. It plays an important role on maintaining ecological balance. So, water vapor is an essential factor in atmospheric greenhouse effect and the water cycle of the earth-air system, it is the central element for climate formation and climate change. Meanwhile, the water vapor content in atmosphere is one of main physical quantities to impact remote sensing application [1]. For a long time, the influence of water vapor in the climate system (such as greenhouse effect) has not been discussed further due to lacking of accurate long-term stability of the global water vapor data record. As a result, detecting global atmospheric moisture in distribution and change is of great significance to weather forecast, meteorological support work, especially water circulation and climate change research [2].

Half a century, the international conventional meteorological radiosonde network has been used in detecting water vapor. But, this kind of detection mode only provides the distribution of water vapor over the fixed points, moreover, the global radio sounding station network density far cannot satisfy the needs of business and scientific research work. Satellite is one of the best choices to obtain atmospheric information because of its unique advantage in time resolution and space resolution, it can make up for sounding data's insufficient of vast ocean, plateau, desert and polar regions [3].

At present, business processing AIRS data mainly use the feature vector statistical regression algorithm of the clear air atmosphere business retrieving International MODIS/AIRS processing software package IMAPP (International MODIS/AIRS Preprocessing Package), this statistic method is simple, so the retrieval accuracy is largely limited [4]. Artificial neural networks have been proven as a reliable technique to diagnose and have good learning capability [5], with the development of artificial neural network, neural network retrieving atmospheric components has been greatly applied as an effective method, the main benefit from it in the work is its high speed and the potential for high speed, as well as its fault tolerance ability. Neural network can directly deduce the complex and unclear relationship between input-output from the training data without doing any assumption about data distribution, so it is very suitable to be put into hyper spectral remote sensing retrieval whose data with the characteristics of high dimension, strongly correlated, and noise-sensitive[6]. In

Received June 29, 2013; Revised July 29, 2013; Accepted August 29, 2013

addition, a simple topological structure with only one hidden layer can achieve the best approximation in the range of allowable error [7]. Research shows that the method of retrieving precipitable water vapor by make use of neural network performance has higher accuracy. It will play a crucial role in storm prediction, numerical weather forecast and other extreme weather prediction by using the method of neural network retrieving precipitable water vapor.

2. The Proposed Algorithm

Known satellite instruments in different spectral band of radiation observation, then to ascertain atmospheric temperature and absorption gas (water vapor), this is the retrieval problem. The retrieval methods we commonly use are: GPS inversion, satellite hyperspectral remote sensing retrieval, satellite infrared remote sensing retrieval and MODIS near infrared retrieval. In this paper, a kind of method based on principal component analysis (PCA) neural network physical statistical retrieval is adopted.

2.1. Data Selection

AIRS is one of the six main observation instruments of NASA Aqua satellite platform, Infrared radiation material has the advantages of more channels, more information and narrower spectrum. It has 2378 infrared spectrum channels spectrum array from 3.7 μm to 15.4 μm, can provide atmospheric information from the ground to the height of 40 km. Aqua satellite goes though China about Beijing time 1:30and13:30 twice everyday [8]. We can download the required data by visiting http://disc.sci.gsfc.nasa.gov/AIRS.

Using satellite hyperspectral data retrieval atmospheric precipitation is hard to meet the training requirement if only using radial brightness from satellite actually, it is necessary to simulate AIRS radial brightness. And AIRS radial brightness data simulation is the basis on using AIRS observation of hyperspectral data retrieval. After researching and discussing, this paper chooses University of Wisconsin-Madison global clear air atmosphere profile training samples (SeeBor Version 5.0) and CRTM (Atmospheric radiative transfer model) to simulate the observation radial brightness information from atmospheric infrared detector AIRS [9].

It is necessary to establish global temperature humidity and atmospheric parameters vertical distribution database when we retrieve atmospheric parameters by using the method of statistical inversion method, statistics-physical retrieval method or neural network method. In addition, we must know the surface temperature and numerical emissivity due to the influence of the surface. In order to meet the application of remote sensing retrieval demand, University of Wisconsin-Madison meteorological satellite application research institute researchers synthesize the global clear air atmosphere profile of the training samples (CIMSS, Cooperative Institute for Meteorological Satellite Studies, University of Wisconsin-Madison) after a great deal of research analysis experimental work, which contains a total of about 15704 groups of precision of atmosphere profile lines, after processing, yields a total of 2,099,250 (15550×3×3 ×15) groups of atmosphere profile lines. At last, we choose 23,325 groups of atmosphere profile lines due to the computing needs a lot of time and some of the combination is not necessarily reasonable, then we get a total of 23,325 AIRS groups radial brightness (at-sensor radiance) by using global atmospheric model after calculating every group of atmosphere profile lines by CRTM, the selected 23,325 groups are the training sample data of neural network data set in turn.

2.2. Data Preprocessing

The solution of the problem that satellite data retrieving atmospheric parameters (mainly retrieving temperature and humidity vertical profile line) is not the only. On the one hand, the information provided by satellite data is insufficient; on the other hand, the information provided by satellite data is sometimes unusefully while retrieving. It is necessary to screen data as well as reduce the observation data dimension and eliminate redundancy, so it is very necessary to use the nonlinear data processing algorithm. In the study of neural network retrieving water vapor, the first question required to solve is data preprocessing, at present, the main method we use is PCA [10].

3. Research Method

3.1. PCA (Principal Component Analysis)

To reset the original variables to a new set of few variables that each other is independent, at the same time, according to actual needs, we can take out a few less total variables to reflect the original variables information as much as possible, this statistical approach is called principal component analysis, and it is a kind of processing dimension reduction method in mathematics.

Its basic idea is to reset many original indexes which have certain correlation (such as the number is P) to a new set of comprehensive indexes that each other is independent, then replace the original indexes by the new indexes. It is usually to put the P indexes into linear combination, then take them as new comprehensive indexes. The classical approach is to use the variance of F1 (the first linear combination, namely the first comprehensive index) to express, the bigger Var (F1) is, the more information F1contains. The F1 should be the biggest variance in all of the linear combination, so F1 is called as the first principal component. If the first principal component is not enough to represent the original information of the P indexes, then consider selecting F2 which is the second linear combination, in order to reflect the original information effectively, the existing information of F1 will not need to appear again in the F2, if it requires that Cov (F1, F2)=0, we call F2 the second principal component, By analogy, we can structure the third principal component,the forth principal component and so on.

The main steps of principal component analysis as follows:

(1)Index data standardization (SPSS software automatically);

(2) Judging the relationship between the index data;

(3) Confirming the number of principal components (m);

(4) Obtaining the expression of the principal component Fi;

(5) Naming the principal component Fi

In recent years, as technology advances and matures, PCA gets continuous development, the principal components transform (PCT) method and the projected principle components (PPC) transform method appears successively.

Principal component analysis technique is in continuous improvement and perfection with technical progress, the projection principal component (projected principle components, PPC) transform proposed by Blackwell has been proved to have better dimension reduction performance [11]. Compared with the traditional principal component, the projection of the principal component has better performance in retrieving atmospheric temperature profile line. Meanwhile, Blackwell indicates, the use of 35 PPC can make retrieval error to the minimum value in theory on water vapor retrieved by using AIRS data [12]. So, the PPC transform method is adopted in this paper. We finally select 41 water vapor channel (41 PPC) from the 2378 channels after analyzing literature and doing many experiments, then we get 24 PPC though dealing with 41 PPC by using the method of PCA, the AIRS radial brightness dimension is reduced from the original 2378 to 24. This transformation realizes the data dimension reduction as well as keeps the relative more information, consequently shorts the neural network training time and improves the efficiency of the proposed algorithm.

Known satellite instruments in different spectral band of radiation observation, then to ascertain atmospheric temperature and absorption gas (water vapor), this is the retrieval problem. The retrieval methods we commonly use are: GPS inversion, satellite hyperspectral remote sensing retrieval, satellite infrared remote sensing retrieval and MODIS near infrared retrieval. In this paper, a kind of method based on principal component analysis (PCA) neural network physical statistical retrieval is adopted.

3.2. Neural Network

BP neural network is the most widely used kind of neural network in all artificial neural network, it also called error back propagation neural network, it is a feed forward network composed by the nonlinear transformation unit [13].But this paper ultimately chooses multilayer feed forward neural network radial basis function neural network RBF (Radical Basis Function) which has the best approximation performance after many test and analysis, the RBF network avoids trivial lengthy computation of back propagation between the input layer and hidden layer, it makes learning $10^3 \sim 10^4$ times faster than usual BP neural network [14]. Figure 1 is the neural network topology. As shown in figure, the input layer has 24 neurons, output layer has one neuron, two hidden layer containing respectively (18, 9) neurons.

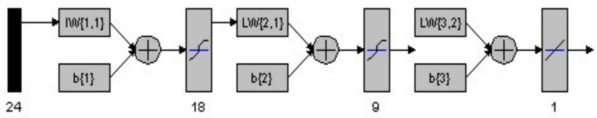

Figure 1. The Neural Network Training Topology

This paper takes AIRS radial brightness PPC data simulated by CRTM as the input of the neural network; correspondingly takes the atmospheric precipitation data from University of Wisconsin-Madison clear sky profile lines data as the output of the neural network. At first, sending the input-output data set to multilayer feed forward neural network training, choosing a optimized neural network as a practical application algorithm; Then the optimization neural network whose simulation error is small enough is applied to actual AIRS detection data, at last, we verify the precision and stability of the algorithm.

After analysis and processing, the 24 PPC handled by principal component analysis technology transformation are treated as the input of the neural network, PWV as the only output of the neural network. The atmospheric precipitation data which correspond to the input of the neural network from University of Wisconsin-Madison clear sky profile of the training samples (SeeBor Version 5.0) is treated as the target output of the neural network, we put the AIRS hyper spectral data simulated by the neural network into CRTM, and consider the calculated results as actual output, we calculate the root mean square error between the actual output and target output calculated is 0.063g/ cm2, while using statistical approach, as shown in Figure 2.

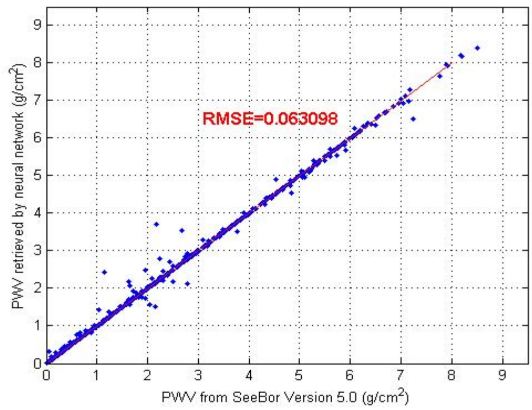

Figure 2. Comparison between the PWV Retrieved from Neural Network based on Algorithm and PWV from University of Wisconsin-Madison.

4. Results and Analysis

This paper select typical region of low and middle latitude area as the research object, middle latitude area's surface type is complex, include: mountain, steppe, desert, rivers and glacier, etc. surface temperature emissivity and atmospheric precipitation's spatial distribution has a large change; the low-latitude area, mainly surrounded by ocean and island, its temperature is high, the water vapor content is rich, and atmospheric precipitation's distribution changes very largely in time and space. We can effectively verity the applicability of the algorithms while taking the two areas as an example. In order to validate the multilayer feed forward neural network inversion algorithm's accuracy and feasibility(which based on principal component analysis), we select a group of AIRS L1B infrared radial brightness data, only part of the measured data are listed to test the result owing to thesis length.

4.1. Low-latitude Region Water Vapor Retrieval Results and Analysis

One AIRS level-1B image covering (0N-22N, 60E-80E) on May 9th, 2012 at 08:47:24-08:53:24 (universal time) is selected for studying low-latitude region.

In order to discuss the influence of the clouds, we don't eliminate pollution clouds pixel in the inversion process, instead retrieve the whole scene images. Because of obtaining a wide range of atmospheric precipitation measured data is difficult, we analyze and compare the algorithm mentioned in this paper with AIRS Level-2 standard products while doing reality testing.

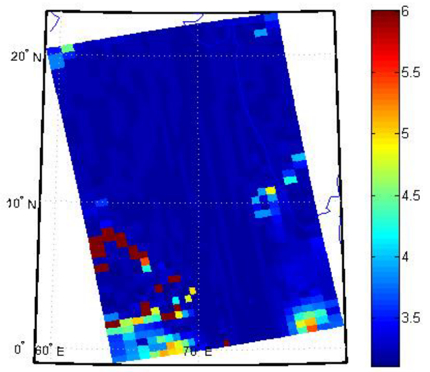

Figure 3. The Spatial Distribution of PWV Retrieved from Neural Network

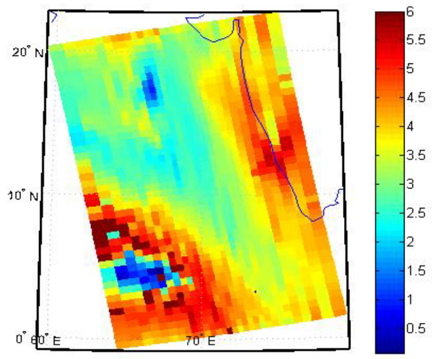

Figure 4. The Spatial Distribution of PWV Retrieved from AIRS Level-2 Standard Product

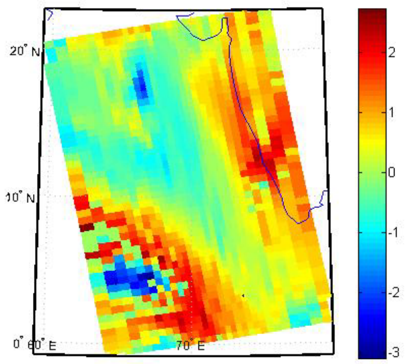

Figure 5. The Absolute Error of PWVs Retrieved by Neural Network and AIRS Product

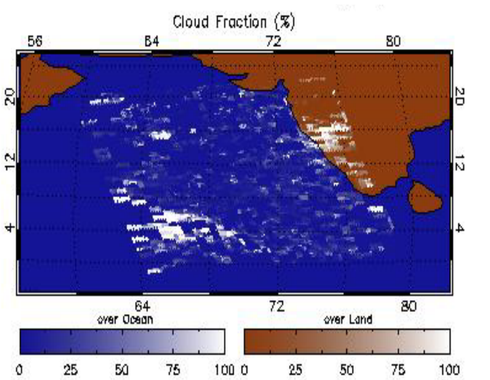

Figure 6. The Cloud Distribution from AIRS Level-2 Standard Product

The spatial distribution of PWV retrieved by neural network and AIRS PWV product are shown by Figure 3 and Figure 4. Compared to figure 6, it can be seen that blue area for little cloud area, and red yellow area for cloudy area. Due to the selected area is larger, marine climate type is specially, the maximum and minimum water vapor content difference is big at the same time. Figure 5, interpolated through AIRS standard products, can analyze their difference intuitively, the root mean square error is 0.7552g/cm^2, it can intuitively find out the absolute error distribution between PWV retrieved by neural network and AIRS PWV product. As is shown in Figure 5, RMSE is smaller in little cloud area; similarly, RMSE is bigger in cloudy area, the absolute error distribution matches well with the AIRS Level-2 standard products.

4.2. Middle-latitude Region Water Vapor Retrieval Results and Analysis

One AIRS level-1B image covering (32N-54N, 66E-97E) on May 2th, 2012 at 20:47:24-20:53:24 (universal time) is selected for studying middle-latitude region.

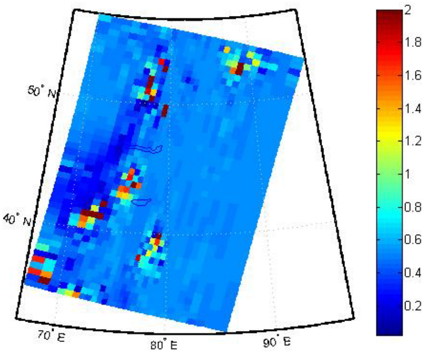

Figure 7. The Spatial Distribution of PWV Retrieved from Neural Network

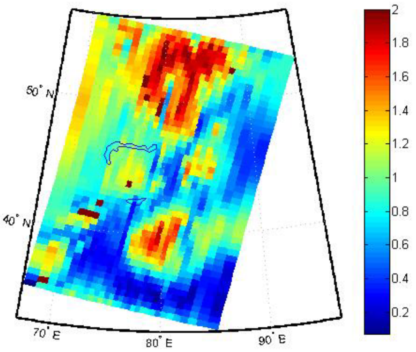

Figure 8. The Spatial Distribution of PWV Retrieved from AIRS Level-2 Standard Product

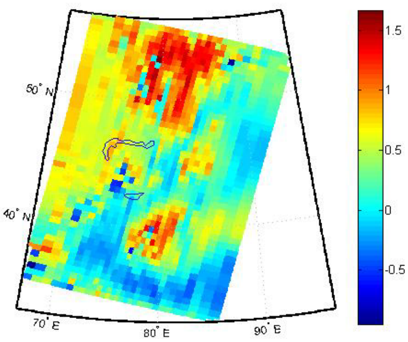

Figure 9. The Absolute Error of PWVs Retrieved by Neural Network and AIRS Product

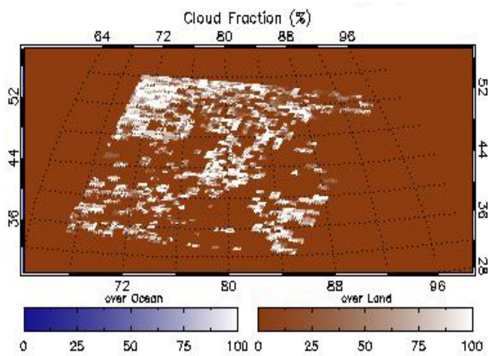

Figure 10. The Cloud Distribution from AIRS Level-2 Standard Product

The surface climate type selected in this paper is rather complex, include mountain land, lakes, rivers, etc. The water vapor changes largely in time and space. The spatial distribution of PWV retrieved by neural network and AIRS PWV product are shown by Figure 7 and Figure 8. Figure 9, interpolated through AIRS standard products, can analyze their difference intuitively, the root mean square error is 0.3359g/cm^2, which more close to and more fit the AIRS standard products. As shown in Figure 9 and Figure 10, we can find the absolute error distribution matches well with the cloud distribution of AIRS Level-2 standard inversion products.

This paper mainly research the atmospheric precipitation water vapor in clear air condition, but from the AIRS Level-2 standard product we can see whether low latitude region or middle latitude region has a certain amount of cloud, no doubt which will cause certain influence to the inversion result. According to the analysis and comparison, we can see the absolute error of the low-latitude region is bigger than the middle latitude region's. The main reason is that the temperature is high in low latitude region, meanwhile, the water vapor content is rich and it changes very largely in time and space, so the absolute error is slightly bigger.

In addition, this study has certain significance for popularization, although we select low and middle regions as subject to study, it also applies to discuss high latitude area.

5. Conclusion

Information resources integration in collaborative logistics network makes use of superior resources in network, achieves the purpose of sharing network resources by technical means. It mainly virtual integrates the resources which have competitive advantages and geospatial discrete distribution in network, strengthens cohesive relations, improves resource response speed and reduces resource redundant waste [15].

Neural network nonlinear retrieval algorithm not only shows the stability and effectiveness of the statistical regression method, but also shows the accuracy of the physical inversion method. The algorithm is especially suitable for atmospheric remote sensing retrieval problem based on the characteristics of the neural network and the nonlinear, non-Gaussian relationship of atmospheric remote sensing data. With the rapid development of network, open and reconfigurable network facility can well achieve the goal for multi-network integration, so that the facility for NGN (next generation network) will be more and more popular [16].

AIRS infrared detection data can display the subtle atmospheric structure, it contains more information several hundred times than previous detection instrument, the method of neural network is used in this paper to handle AIRS data, then invert atmospheric precipitable water vapor, finally we compare the PWVs retrieved by neural network and AIRS Level-2 standard product, the result shows that Neural network retrieval algorithm is simple and feasible, what's more, its error is rather small.

In addition, it is worth our attention is that the retrieval spatial resolution of neural network algorithm is 13.5km [17], the spatial resolution of AIRS standard product is 45km, by comparison, this algorithm improves the spatial resolution 2 times higher than AIRS standard product, Therefore, the neural network algorithm products can better reflect the actual distribution of water vapor, compared with AIRS standard products, the finer levels have been obviously improved. It makes great sense to exactly descript and analyze the exchange and transport of water vapor, and discuss small range of atmospheric state activities.

Acknowledgement

The data used in this study are acquired from University of Wisconsin-Madision, and the authors thank the Key Laboratory of Atmosphere Sounding, China Meteorological Administration for this work support.

References

[1] HL Senior. Climate and Water. England: *Nature*. 2002; 419 (6903).
[2] E Smith, G Asrar, Y Furuhama et al. International Global Precipitation Measurement (GPM) Program and Mission: An Overview. In Measuring Precipitation From Space. *USA: Springer*. 2007; 611-653.
[3] Wang Ying, Huang Yong, Siyuan Huang. A Preliminary Study of the Retrieval Methods for Atmosphere and Humidity Profiles. *Remote Sensing for Land & Resources.* 2008; 3(1).
[4] Xuehui Zhang, Guan Li, Zhenhui Wang, Han Jing. Retrieving Atmospheric Temperature Profiles Using Artificial Neural Network Approach. 2009; 11(35).
[5] Li Guoping, Zhang Qingwei, Ma Xiao. Combination of Fault Tree and Neural Networks in Excavator Diagnosis. *TELKOMNIKA*. 2013; 11(4): 1787-1796.
[6] Hornik K, Stinchcombe M, White H. Multilayer feedforward networks are universal approximators. *Neural Networks*. 1989; 2(5): 359-366.
[7] A Gallant, H White. There exists a neural network that does not make avoidable mistakes. *In Proceedings of the International Joint Conference on Neural Networks*. 1988; 657–664.
[8] Wenhua Gao，Fengsheng Zhao，Changsong Gai. Validation of Airs Retrieval Temperature and Moisture Products and their Application in numerical Models. *Acta Meteorologica Sinica.* 2006; 64 (3): 2712280.
[9] Shenglan Zhang. A Study of Precipitable Water Vapor Retrieval Using Infrared Hyperspectral Soundings (degree paper) Chengdu; Chengdu University of Information Technology. 2011.
[10] JA Sobrino, ZL Li, MP Stoll, et al. Multi-channel and multi-angle algorithms for estimating sea and land surface temperature with ATSK data. *International Journal of Remote Sensing*. 1996; 17: 2089-2114.

[11] WJ Blackwell. A Neural-Network Technique for the Retrieval of Atmospheric Temperature and Moisture Profiles from High Spectral Resolution Sounding Data. *IEEE Trans. Geosci. Remote Sensing*. 2005; 43(11): 2535-2546.
[12] William J Blackwell, Frederick W Chen. Neural Networks in Atmospheric Remote Sensing. MIT. 2009.
[13] Tie Wang, Chao Wang. Bora Engine CH Emissions Diagnosis Based on Neural Network. *TELKOMNIKA*. 2012; 10(8): 2343-2350.
[14] Liangjun Zhang, Cao Jing, Shizhong Jiang. Neural Network Practical Tutorial. China Machine Press. 2008; 2.
[15] Xiaofeng Xu, Jinlou Zhao, Yirui Deng. Integrated Optimization and Deployment Mechanism of Information Resources in Complex Manufacturing Collaborative Logistics Network. *Journal of Theoretical and Applied Information Technology*. 2012; 44(1): 153 – 160.
[16] Zhou Minhui, Wang Weiming, Zhou Jingjing. Routing Optimiza for Forces Based on Traffic Matrix. *Journal of Theoretical and Applied Information Technology*. 2012; 44(1); 007 - 011
[17] Shenglan Zhang, Lisheng Xu, Jilie Ding, Hailei Liu, Xiaobo Deng. Precipitable Water Vapor Retrieval Using Neural Network from Infrared Hyperspectral Soundings. Satellite Remote Sensing Lab College of Atmospheric Sounding, Chengdu University of Information Technology. 2010.