Effects of Channel Selection and Observation Error Allocation on the Assimilation of AIRS Data

Jianwei Zhang[1], Gen Wang[1], Delegerima[2], Yin Yang[3] and Jin Wang[4]

[1] *College of Mathematics and Statistics, Nanjing University of Information Science & Technology, Nanjing, China*

[2] *College of Economics and Management, Nanjing University of Information Science & Technology, Nanjing, China*

[3] *National Meteorological Centers, China Meteorological Administration, Beijing, China*

[4] *Computers and Software School, Nanjing University of Information Science & Technology, Nanjing, China*

{zhangjw, wangjin}@ nuist.edu.cn; {203wanggen,gerima_dele, mowwind}@163.com

Abstract

Hyper-spectral Atmospheric Infrared Sounder AIRS has many channels. Taking into account the correlation between channels, different detection purposes and the timeliness of assimilation, channel selection is necessary. Principal component-dual-zone-stepwise regression method is proposed in this paper and points the day and night for channel selection. The contribution rate of channel observation brightness temperature to the objective function is controlled by channel observation error, variational data assimilation system is generally given observation error and then in the process of minimization objective function remains unchanged, the paper adopt the observation error re-estimated and then provide observation error information to the variational data assimilation. Study effects of channel selection and the observational error allocation on the assimilation of AIRS data in this paper.

Keywords: *Hyper-spectral, AIRS, Error re-estimated, Channel selection*

1. Introduction

Numerical Weather Prediction (NWP) system which assimilates satellite data shows that satellite data have positive effects on weather prediction models, especially in oceans and the South and North Hemisphere where the radiosonde data are short.

Hyper-spectral Atmospheric Infrared Sounder (AIRS) which is equipped on the second satellite Aqua in the Earth Observing System (EOS). There are 2378 channels covering the infrared spectrum area and only 324 channels are used in practical [1, 2, 3]. Thus new problems are caused. One problem is too much calculation and another one is that the correlation between channels will cause a lot of redundant information, channel selection is necessary.

Variational assimilation based on linear estimation theory, the idea is that the objective functional minimization. Objective functional includes background and observation items. The contribution of observation to the objective functional is regulated by the reciprocal of the channel observation error. Variational assimilation of the specific implementation process, the observation errors are generally given and not change later. In this paper we used observation error re-estimation method to estimate the observation error, which is provided to the variational assimilation [4, 5].

Different from the current domestic and international, channel selection is a sub-step iterative method based on information entropy (written as IC) [6, 7, 8, 9]. Principal components-dual-zone stepwise regression is proposed. Principal component analysis can keep the most important information and stepwise regression can get the channels sensitive to the principal components. Partition idea is considered the global and the local optimum, to meet the information needs of the middle upper layer or middle lower layer in practical use. AIRS brightness temperature assimilation experiments adopt observation error re-estimation, which based on channel selection.

2. Channels Selection and Observation Error Allocation

Minimization objective function of variational assimilation to get analysis field $\mathbf{x}_a$

$$\mathbf{x}_a = \mathbf{x}_b + (\mathbf{H}^T \mathbf{R}_{obs}^{-1} \mathbf{H} + \mathbf{B}^{-1})^{-1} \mathbf{H}^T \mathbf{R}_{obs}^{-1} (y_{obs} - H\mathbf{x}_b)$$

where $\mathbf{x}_b$ is background field, $\mathbf{B}$ and $\mathbf{R}_{obs}$ is background and observation error covariance matrix respectively, y_{obs} is observed bright temperature. $\mathbf{H}$ is Jacobi matrix.

2.1 Principal Components-Stepwise Regression based Channel Selection

Step1: Take the temperature and humidity in $\mathbf{H}$ matrix respectively and then standardize variable of each channel in $\mathbf{H}$ to get the standardization matrix $\mathbf{R}_{PCA} = (r_{ji})_{m \times n}$ where n is layer number and m is channel number.

Step 2: Calculate the covariance matrix of $\mathbf{R}_{PCA}$ which is shown as $Cov(\mathbf{R}_{PCA})$.

Step 3: Analyze the principal components of $Cov(\mathbf{R}_{PCA})$ to take p_0 principal components, which is represented as $\mathbf{Y}_i$ ($i = 1, ..., p_0$).

Step 4: Perform stepwise regression analysis on $[\mathbf{H} \quad \mathbf{Y}_i]$ ($i = 1, ..., p_0$) where we start with $i = 1$ to perform stepwise regression analysis on $[\mathbf{H} \quad \mathbf{Y}_1]$.

Step 5: Similarly, we can repeat the operations Step 3 and 4 to obtain the channel subset which is sensitive to the previous p_0 principal components.

2.2 Re-estimated Observation Error

Assuming observation error and background error is irrelevant (both belong to the orthogonal relationship), according to the definition of the variational method, the variational method to seek the analysis field makes the goal of functional minimum, the variational assimilation analysis of the inverted the distance between the true value, relative to distance between observation and true value and between background and the true value is the shortest, so the difference of analysis and true value is orthogonal to the difference of observation and background, the relationship among three can be described with a triangular geometry [4].

According to the Pythagorean Theorem. We can get

$$(y^o - y^t)^2 = (y^o - y^a) \bullet (y^o - y^b)$$

Among them, y^a , y^b and y^t represent the information of inversion, background and the true value projected onto the observation space respectively. y^o represents the actual observed values. For observation, re-estimated the observation error is the average error of all observations.

3. Experiments and Result Analysis

3.1 The Specific Implementation of the Channel Selection

Step 1: We only consider the CO_2 channels in the long-wave region. First, we screen vapor and other channels in the pre-processing and to remove the channels affected by the sunlight. And then analyze the principal components to the temperature Jacob matrix. Several larger principal components in the front are selected and stepwise regression analysis is performed to obtain the combined channels sensitive to principle components.

Step 2: We only consider the channels near water vapor-absorbing bands and analyze the humidity Jacob matrix. The specific progress is similar to the temperature analysis above.

Step 3: We add the channels affected by the sunlight. We can further select a part of channels which are affected by non-local thermo-dynamical equilibrium.

Step4: When the total number of selected channels reaches the given number or meets the requirements of stepwise regression, whole process is terminated.

3.2 Relevant Model and Assimilation Results

One model we use in this paper is Met Office 1D-Var variational retrieval procedure from European Centre for Medium Range Forecast (ECMWF). Add to observational error re-estimated interface on the basis of this model. Another model we use is RTTOV7 model. For AIRS all the simulation observations are chosen nadir, azimuth angle zero. The remaining channels will to be selected after pretreatment.

In the course of ideal experiment, 1D-Var's own profiles are used. Select one of the profiles as real profile and corresponding simulation brightness temperature as true brightness temperature. The profiles are added Gaussian random error and perturbed for 100 times as a background profiles. True brightness temperature is added Gaussian random error and perturbed for 100 times as observed brightness temperature.

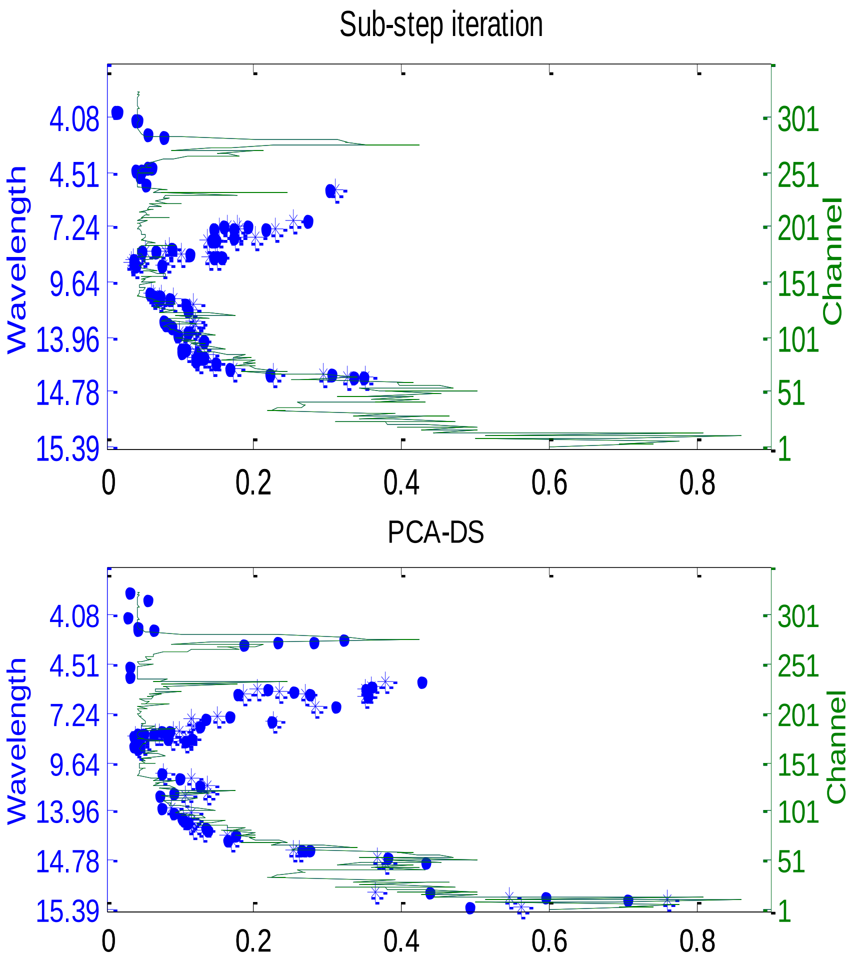

Figure 1. Observation Error of Channels which Get Based on IC and PCA-DS

Figures 1 show the original observation error of 324 channels (the green line) and re-estimated observation error, which adopt information entropy sub-step iteration (IC, above figure) and principal components-dual-zone stepwise regression(PCA-DS, below figure)respectively.“*”and “.” represent 42 and 62 channels respectively.

Seen from Figure 1, channels of water vapor absorption band (the center near 6.3 μm) re-estimate the error value is larger, indicating that the contribution rate of water vapor channel brightness temperature to objective functional is reduce.

The following adopt IC method and PCA-DS variational assimilation AIRS bright temperature by re-estimating observational error. In the statistical analysis of observations and models of space error (error is defined as the final value of the inversion and the real value of the difference) RMS value. In Figure 2, IC+C means using the classical variational assimilation method assimilation the channel combination of AIRS, which adopt IC method. Here given 100 profiles RMS statistics of inverted bright temperature (written as: BT), temperature and humidity error. The rest of figure notations are done similarly: (1) IC+Re, adopt observation error re-estimation, channels get by information entropy sub-step iteration. (2) PCA-DS+C, adopt classical variational assimilation, channels get by principal components-dual-zone stepwise regression. (3) PCA-DS+Re, adopt observation error re-estimation, channels get by principal components-dual-zone stepwise regression. Here only given 62 channels combination of nighttime.

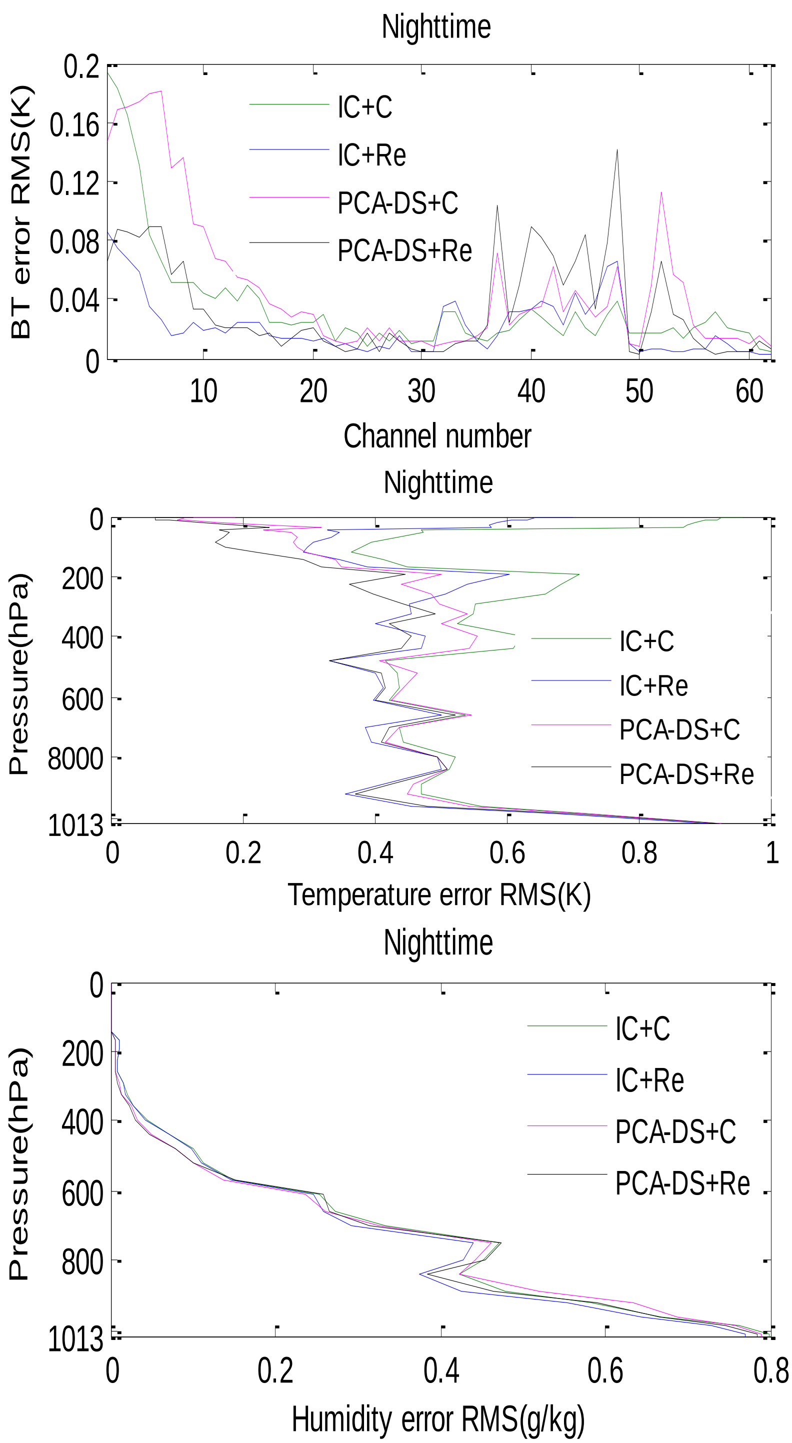

Figure 2. Brightness Temperature (BT), Temperature and Humidity Error RMS Statistics

Seen from Figure 2, the ultimate effect of different channels combination is different, assimilation effect based on the observation error re-estimated is better than classical variational assimilation, which adopt original observation error.

4. Conclusions

Channel selection is necessary due to the factors of correlation between channels of hyper-spectral sounder, and the timeliness of assimilation. Channel brightness temperature of contribution rate to objective functional by observation error control. Effects of different channel selection method and observation errors are re-estimated on the assimilation of AIRS data in this paper.

Acknowledgements

This work is supported by the National Natural Science Foundation of China (61173072), Chinese Ministry of Finance project (GYHY200906006) and College Graduate Student Research and Innovation Program of Jiangsu province (CXZZ12_0511). It was also supported by the Natural Science Foundation of Jiangsu Province (No. BK2012461). Prof. Jianwei Zhang is the corresponding author.

References

[1] H. H. Aumann, et al., “AIRS/AMSU/HSB on the Aqua mission: design, science objective, data products, and processing systems”, IEEE Trans on Geosci. And Remote Sensing, vol. 4, no. 2, **(2003)**.

[2] A. D. Collard, R. Saunders, J. Cameron, B. Harris, Y. Takeuchi and L. Horrocks, “Assimilation of data from AIRS for improved numerical weather prediction”, Proc. Thirteenth International TOVS Study Conference **(2003)** October 28- November 4, Sainte-Adèle, Canada.

[3] M. D. Goldberg, Y. Qu, L. M. McMillin, W. Wolf, L. H. Zhou and M. Divakarla, “AIRS near-real-time products and algorithms in support of operational numerical weather prediction”, IEEE Trans. Geosci. Remote Sens., vol. 41, no. 2, **(2003)**.

[4] G. Desroziers, L.Berre, B. Chapnik and P.Poli, “Diagnosis of observation, background and analysis-error statistics in observation space”, Q. J. R. Meteorol.Soc., vol. 131, **(2005)**, pp. 613.

[5] N. Bormann, A. J. Geer and P. Bauer, “Estimates of observation-error chaaracteristics in clear and cloudy regions for microwave imager radiances from numberical weather prediction”, Q. J. R. Meteorol. Soc., vol. 137, **(2011)**, pp. 611.

[6] C. D. Rodgers, “Information content and optimisation of high spectral resolution measure-ments”, Advances in Space Research, vol. 21, no. 3, **(1998)**.

[7] F. Rabier, N. Fourrie, D. Chafai and P. Prunet, “Channel selection methods for Infrared atmospheric sounding interferometer radiances”, Q. J. R. Meteorol. Soc., vol. 128, **(2002)**, pp. 581.

[8] N. Fourrie and J. N. Thepaut, “Validation of the NESDIS near real time AIRS channel selection”, ECMWF Tech.Memo.390, European Center for Medium Range, http://www.ecmwf.int/publications/.

[9] N. Fourrie and J. N.Thepaut, “Evaluation of the AIRS near-real-time channel selection for application to numerical weather prediction”, Q. J. R. Meteorol. Soc., vol. 129, **(2003)**, pp. 592.

Author

Jianwei Zhang

Dr. Jianwei Zhang received his master degree in 1998 from Wuhan University and his Ph. D degree in 2006 from Nanjing University of Science & Technology. He is a professor in Nanjing University of Information Science & Technology. His main research interests include image processing, pattern recognition and numerical analysis.

Gen Wang

He received master degree in 2011 from Nanjing University of Information Science & Technology and now as a Ph.D candidate at Nanjing University of Information Science & Technology. His main research interests include meteorological satellite data assimilation and image processing.

Delegerima

She received master degree in 2009 from Inner Mongolia University of Technology and now as a Ph.D candidate at Nanjing University of Information Science & Technology. Her main research interests include short-term weather and climate analysis and meteorological disaster risk assessment. .

Yin Yang

He graduated from Chinese Academy of Meteorological Sciences, with a master's degree in 2011, and now works at National Meteorological Centers of China Meteorological Administration. His main research interest is meteorological satellite data assimilation.

Jin Wang

Dr. Jin Wang received the B.S. and M.S. degree in the Electronical Engineering from Nanjing University of Posts and Telecomm., China in 2002 and 2005, respectively. He received Ph.D. degree in the Ubiquitous Computing laboratory in Computer Engineering Dept. of Kyung Hee University Korea in 2010. Now, he is a professor in Nanjing University of Information Science & technology. His research interests include routing protocol and algorithm design, analysis and optimization, and performance evaluation for wireless ad hoc and sensor networks.