Diagnosis and tuning of observation error in 1DVAR (MIRS) in all sky conditions

Yin Yang*, Wei Han#

* National Meteorological Center, CMA, Beijing, 100081

Center for Numerical Prediction, CMA, Beijing, 100081

Abstract: Retrieval of satellite microwave sounding data under clear sky is much accurate now, but the accuracy in unclear sky such as cloudy and precipitation is not so well, the main reason of such is that the observation error changes in unclear sky while which is not considered in retrieval. The observation error could be diagnosis and tuned according to the hypothesis that the relationship between background and observation error are unrelated in data assimilation. In this paper, Desrozier's approach of diagnosis and tuning to observation error is used in retrieval of temperature and water materials from AMSU sounding data in all sky conditions under the basis of MIRS. Result shows that the error tuning could get observation error of more accuracy and make retrieval improved at all weather conditions. In details the retrieved temperature profiles are close to verification data, water materials such as ice water path and rain rate are accord with CloudSate data also.

Keywords: Observation error, diagnosis and tuning, retrieval,

cloudy and precipitation

1. Introduction

Currently most operational assimilation and retrieval systems are based on the variational formalism (Courtier and Talagrand 1987). The formalism allows the use of unconventional observation such as satellite data which is not directly or linearly linked with model variables. However the variational algorithm relies on the theory of least-variance linear statistical estimation (Talagrand 1997), in this algorithm each set of information is given a weight proportional by its specified error covariance. Classically, error information is given by estimation to the observation and background of atmospheric state. Variational assimilation and retrieval method is dependent on the appropriate statistics of observation and background error. Unfortunately it is known that those statistics error are not accurate, especially in complicated weather conditions such as unclear sky and their determination remains a big challenge, which makes satellite data assimilation and retrieval behaves well in clear sky but un-ideal in unclear sky conditions such as cloudy and precipitation.

Much effects has made in determining observation error perfectly, Hollingsworth and Lonnberg (1986) had a study of getting information from the differences between observation and the background of counterpart by assuming that background errors are cross-correlations while observation errors not. Desroziers and Ivanov (2001) proposed a method to tune scaling coefficients for the observation error covariance based on an optimality criterion for the cost function at the minimum, assuming the correlations are accurately represented in the initially assumed observation error covariance. This has been applied by Chapnick *et al.* (2006) and others for diagonal observation error covariance matrices. Related to this method is the maximum-likelihood estimation (Dee and da Silva, 1999), which directly fits free parameters of covariance models to FG-departure statistics. While these methods provide useful tools, the drawback of them is that little is known about which covariance models would be appropriate for satellite radiances. Using incorrect covariance methods can lead to undesired results (Liu and Rabier, 2003; Chapnick *et al.*, 2006). G. Desroziers et al. (2005) presented an approach of diagnosis and tuning observation error based on combination of observation minus background, observation minus analysis and background minus analysis differences. Another way of tuning observation error is Hollingsworth/ Lonnberg method (Rutherford, 1972; Hollingsworth and Lonnberg, 1986), which made an assumption that FG errors are spatially correlated while observation errors are not. Niels Bormannn and Peter Bauer (2010) made a comparison between Hollingsworth/ Lonnberg method and Desroziers method and found the later method is better in tuning observation error in 4dvar assimilation. Besides, in this paper study is focus on 1dvar retrieval, which means that spatially correlation in FG errors is not exist, so the Hollingsworth/ Lonnberg method is not used in this study.

Microwave Integrated Retreieval System (MIRS) is developed by NOAA based on an assimilation-type scheme (1DVAR), with capable of optimally retrieving atmospheric and surface state parameters simultaneously. The direct outputs from MIRS include temperature, moisture and several hydrometeors atmospheric profiles, land surface temperature and emissivity at all channels. In this paper, Desrozier's approach of diagnosis and tuning to observation error is used in retrieval of temperature and water materials from AMSU sounding data in all sky conditions under the basis of MIRS.

In section 2, the approach of diagnosis and tuning observation error by G. Desroziers is introduced. The comparison of using tuned and original observation error in retrieval is in section 3. Section 4 is the part of conclusion.

2. Desrozier's approach of diagnosis and tuning observation error

In this section, Desrozier's approach of diagnosis and tuning observation error is introduced briefly. In variation algorithm the analysis is usually archived by deriving the minimization of cost function (Lorenc 1986). The cost function J(x) is

$$J(x)=[\tfrac{1}{2}(x-x_b)^T\times B^{-1}\times(x-x_b)]+[\tfrac{1}{2}(y_o-H(x))^T\times R^{-1}\times(y_o-H(x))] \quad (1)$$

Cost function J(x) has minimum value when its' derivative equal to zero in mathematics. And the analysis result is

$$x_a = x_b + \{(B^{-1} + \mathrm{H}^T R^{-1} \mathrm{H}\ )^{-1} \mathrm{H}^T R^{-1}\}[y_o - H(x_b)] \tag{2}$$

Another form of x_a is that:

$$x_a = x_b + \sigma x_a = x_b + K d_b^0 \tag{3}$$

Where x_b is the background information and σx_a is the analysis increment. K is the gain matrix.

$$K = BH^T (HBH^T + R)^{-1} \tag{4}$$

Where B and R is the background and observation error covariance matrix independently, *H* is the nonlinear observation operator when background and observation are different. H is the matrix corresponding to linearized *H.* d_b^0 is the innovation vector, which means the difference between observation and background in counterpart.

$$d_b^o = y^o - H(x_b) = y_o - H(x_t) + H(x_t) - H(x_b) \approx \varepsilon_o - H\varepsilon_b \tag{5}$$

In variational algorithm it is always assumed that the relationship of errors between background and observation are unrelated. Considering the linearity of the statistical expectation operator E, it is clear that.

$$E[d_b^o (d_b^o)^T] = R + \mathrm{HBH}^T \tag{6}$$

Also the difference between analysis and background can be defined as $d_b^a = H(x_a) - H(x_b) \approx \mathrm{H}\sigma x_a = \mathrm{HK} d_\mathrm{b}^o$ (7)

So the multiply of d_b^a and d_b^o is $d_b^a (d_b^o)^T = \mathrm{HK} d_b^o (d_b^a)^T$

The statistics expression of above equation is

$$E[d_b^a(d_b^o)^T] = \mathrm{HBH}^T(\mathrm{HBH}^T + R)^{-1}E[d_b^o(d_b^o)^T] \tag{8}$$

Considering gain matrix K, equation (8) can be expressed in another form: $E[d_b^a(d_b^o)^T] = \mathrm{HBH}^T$ (9)

Similarly the difference between observation and analysis is

$$d_a^o = y_o - H(x^b + \sigma x_a) \approx y_o - H(x_b) - \mathrm{HK}d_b^o \tag{10}$$

The statistical expectation of cross-product between d_b^o and d_a^o is

$$E[d_a^o d_b^{oT}] = R(\mathrm{HBH}^T + R)^{-1}E[d_b^o d_b^{oT}] \tag{11}$$

Making a combination of equation (6) and (10) it is clear that

$$E[d_a^o d_b^{oT}] = R \tag{12}$$

According to equation (12), in statistical observation error could be diagnosis and tuned if observation, background and analysis are known.

3. Application of observation error tuning in retrieval

Desrozier's approach of diagnosis and tuning observation error is introduced in section two. In this section that algorithm is used in retrieval of (Advanced microwave sounding unit) AMSU and (microwave humidity sounder) MHS observation of NOAA-19 on the bias of MIRS.

3.1 Data choosing

Considering of the land emissivity simulating on complex surface does not accurate enough and observation should be verified by other data, the typhoon ‘FAPABI’ is selected as observation target with the channels sensitive to land surface are not used, for example AMSU-A’s channel 1-4,15 and MHS’s channel 1 and 2. Besides, channel 12 to 14 of AMSU-A is not used also. The 11^{th} typhoon ‘FAPABI’ of year 2010 is generated on 15^{th} Sep 2010 and perished on 21th, which caused severe property damage and casualty to China. NOAA-19’s observation to ‘FAPABI’ is selected from 00 to 06UTC on 17^{th} Sep 2010 for the reason of satellite CloudSate has an observation to that typhoon also during this period. And the observation area is selected on open sea with latitude between 0^{o}N and 30^{o}N, longitude between 100^{o}E and 240^{o}E. Figure 1 is the observation orbit of NOAA-19 to ‘FAPABI’.

Fig 1. The observation orbit of NOAA-19 to ‘FAPABI’

3.2 Approach to distinguish weather condition

It is well known that the retrieval accuracy changes in different weather conditions such as clear sky, cloudy and precipitation. The mainly reason for which is the simulation error which contains in observation error has non-linearly variety in different weather conditions. So dividing observation into different types according to weather is the precondition in studying of retrieval in all weather conditions.

Clear sky observation is used only in the past for the reason of simulation has great error in un-clear sky. There are many researches in division clear and un-clear observation. Bennartz advanced an index, Bennartz index $I_S = (BT_1 - BT_2) - (-39.2010 + 0.1140\Theta)$ to weigh whether an observation is affected by cloud and rain, where BT_1 and BT_2 are bright temperature (BT) observations of channel 1 and 2 on MHS or AMSU-B, Θ is the local zenith angle. The absolute value of difference between observation and simulated BT $|BT_{ob} - BT_{fg}|$ of some channel (for example channel 4 of AMSU-A and channel 2 of MHS or AMSU-B) could express the influence of unclear sky to observation, where the BT_{ob} means the observation and BT_{fg} means the simulated background BT. It always believed that an observation is affected (or contained) by cloud and (or) precipitation when the absolute value of the difference exceeds a given threshold. The precipitation probability index could represent the probability of an observation been contained by precipitation. The precipitation probability index are: $p = 1/(1 + e^{-f}) \times 100$, $f = 10.5 + 0.184 \times BT_1 - 0.221 \times BT_{15}$, where BT_1 and BT_{15} represents channel 1 and 15 of AMSU-A independently.

Those approaches mentioned above are used in dividing observation into clear sky, cloudy and precipitation, and observation to ‘FAPABI’ in visible light is used as the verification. Result shows that the approach of difference between observation and simulated BT in channel 2 of MHS could distinguish the clear and un-clear sky with the threshold of 5 (fig.2), and the precipitation probability index exceed 60 means precipitation exists (fig.3).

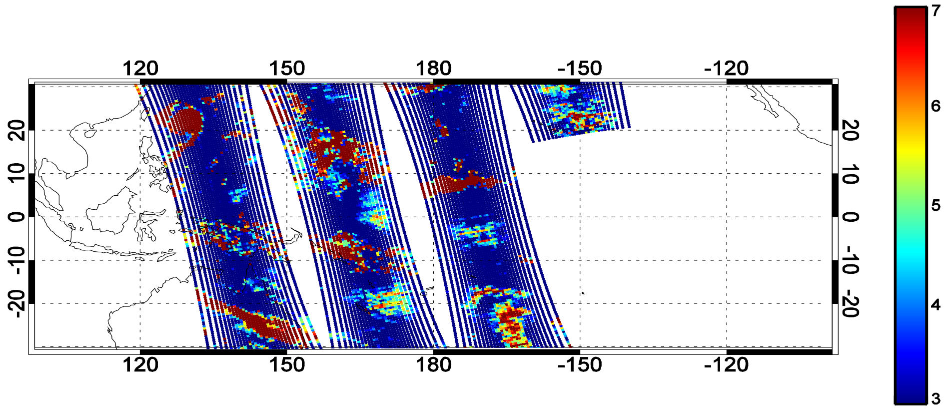

Fig.2. $|BT_{ob} - BT_{fg}|$ in channel 2 of MHS to ‘FAPABI’, bigger than 5 means cloudy

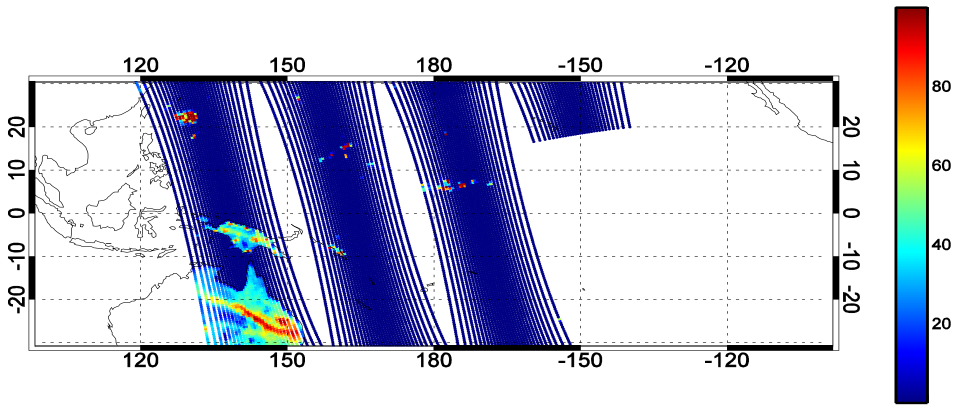

Fig.3. Precipitation probability index to ‘FAPABI’, value bigger than 60 means precipitation

In this paper the approaches selected to divide observation into clear sky, cloudy and precipitation is as follow:

1. Observation is in clear sky if the absolute value of difference between observation and simulated BT is smaller than 5 on channel 2 of

MHS and the precipitation probability index is smaller than 60.

2. Observation is in cloud if the absolute value of difference between observation and simulated BT is bigger than 5 on channel 2 of MHS but the precipitation probability index is smaller than 60.

3. Observation is in precipitation if the absolute value of difference between observation and simulated BT is bigger than 5 on channel 2 of MHS and the precipitation probability index is bigger than 60.

3.3. Comparison of retrieval result

In this section, the tuned and original observation error and the retrieval result of counterpart are analyzed and compared.

3.3.1 Comparison of new and old temperature retrieval

Make MIRS run once and the original retrieval result are got, then the tuned observation error is got using the Desrozier's approach also. Figure 4 is comparison of original and tuned observation error on selected channels.

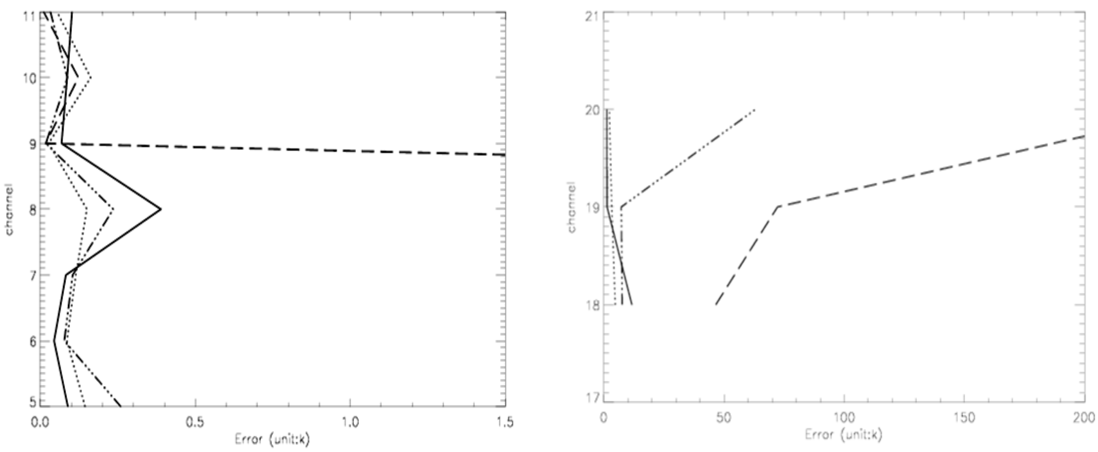

Fig. 4. The comparison of original and tuned observation error on used channels of AMSU-A (left) and MHS (right). Original is in solid, tuned in clear sky (dashed), cloudy (dashed-dotted) and precipitation (datted).

Instead of repesent all weather conditions in unification (such as the original observation error) , the tuned observation error repesents weather conditions in details, which could make retrieval accurate improved comparing with the original one in theory. To verify that, the FNL data is used as the truth to checking which retrieval is more accuracy between the oringinal and tuned observation error. Figure 5 to 7 are the comparison of retrieved temperature in RMS between tuned and original. It is obviously that in clear sky at all levels the RMS of retrieved temperature by tuned observation error is smaller than the original in figure 6. From figure 7 and 8, it is clear also that the RMS of retrieved temperature by tuned observation error is smaller than the original in unclear weather condition of cloudy and precipitation.

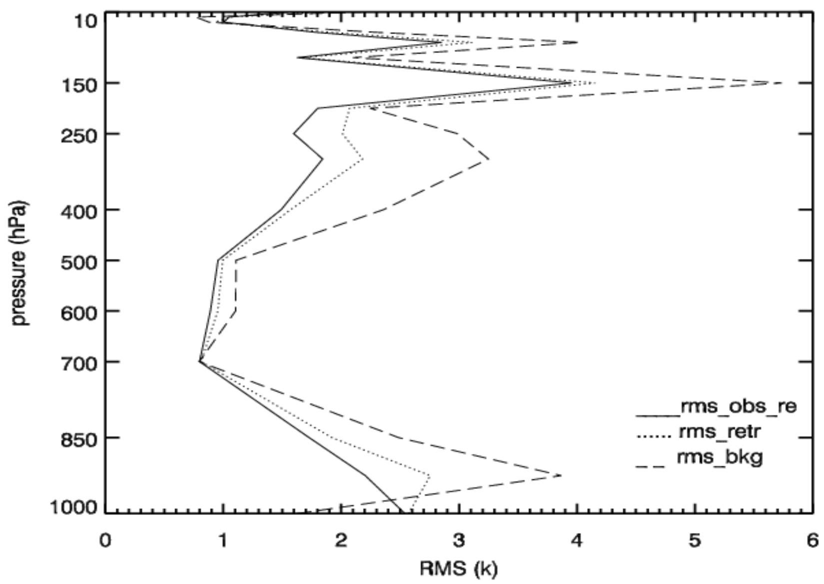

Fig.5. The comparison of RMS in retrieved temperature of background (dashed), original (dotted) and tuned (solid) in clear sky

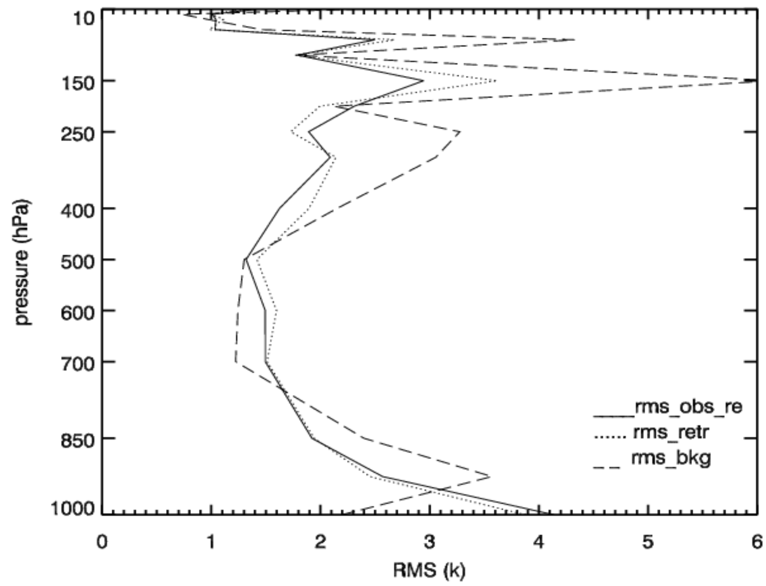

Fig.6. The comparison of RMS in retrieved temperature of background (dashed), original (dotted) and tuned (solid) in cloudy

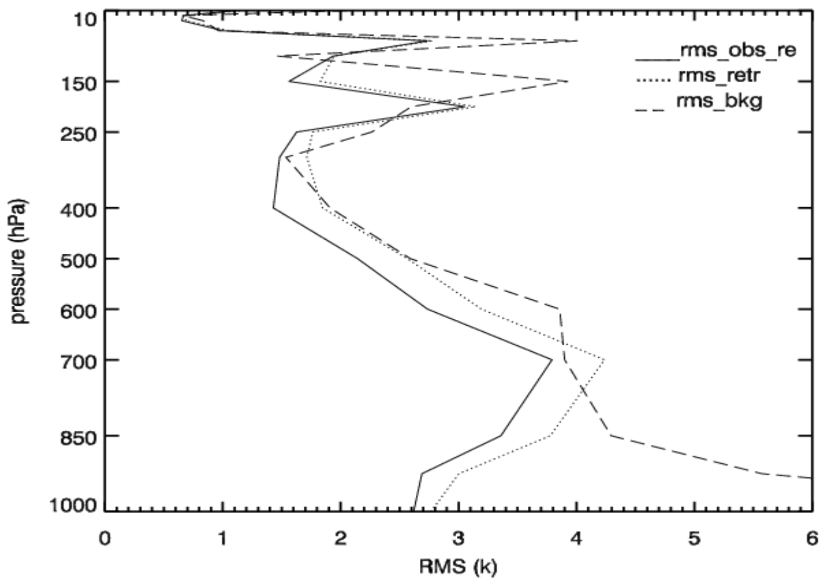

Fig.7. The comparison of RMS in retrieved temperature of background (dashed), original (dotted) and tuned (solid) in precipitation

However, in figure 7 both of the tuned and original observation error makes retrieval RMS worse than the background at middle level from 700 to 500hPa. Here is a possible interpretation for that, in cloudy the exist of cloud particle makes simulation error bigger than that in clear sky, and the effect of observation error tuning is not magntiude enough to make retrieved temperature more accurate than background at middle level. The jacobian matrix of temperature on used channel of AMSU-A may offer some help in explaining that (Fig 8). In figure 8, the jacobian in cloudy (dashed) at channel 5 and 6 of AMSU-A are not so smooth as that in clear sky (solid). It is known that channel 5 and 6 of AMSU-A are sensitivity to levels about 700 to 400hPa. The abnormity of jacobian on channel 5 and 6 of AMSU-A in cloudy makes the temperature retrieval un-ideal in middle levels. Manly because of background error is too big under the condition of precipitation, the retrieval is better both than background in figure 7.

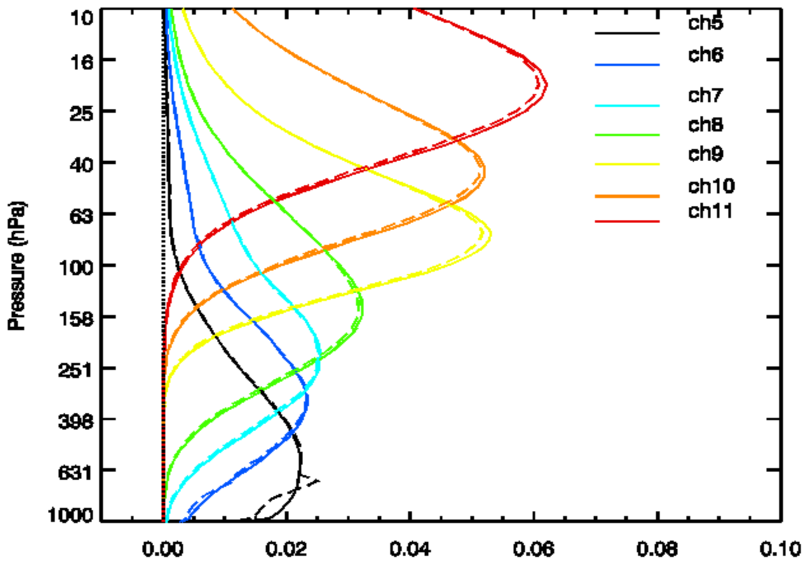

Fig.8. The temperatrue jacobian of channels on AMSU-A in clear sky (solid) and cloudy (dotted)

3.3.2 Verification of retrieved water materials

MIRS could retrieve water materials such as rain rate from satellite observation, figure 9 is the retrieved rain rate and ice water path (IWP) of typhoon ‘FAPABI’ by tuned observation error on the basis of MIRS .

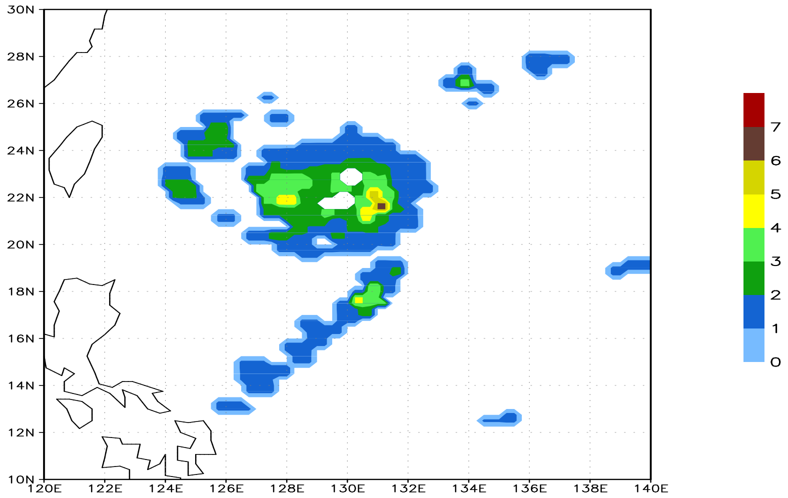

Fig.9a. Retrieved rain rate of ‘FAPABI’ (unit: mm/hr)

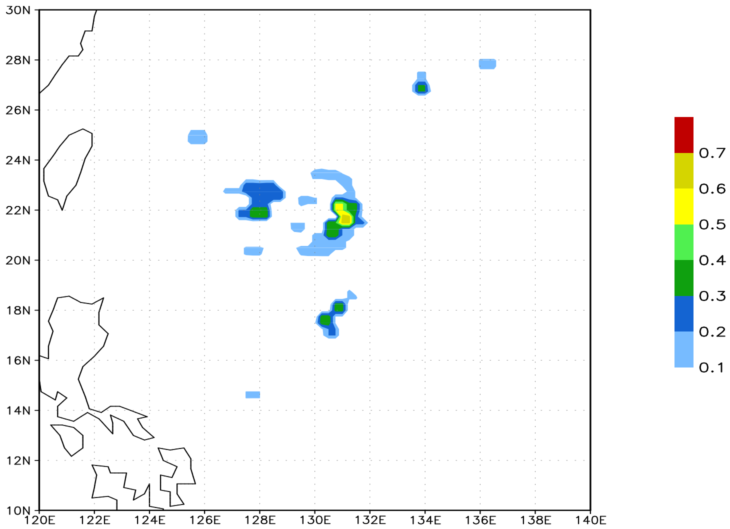

Fig.9b. Retrieved ice water path of ‘FAPABI’ (unit: g/kg)

The satellite CloudSate has an observation to typhoon ‘FAPABI’ also during the period of NOAA-19’s scanning to ‘fAPABI’. CloudSate data is used here to verify the retrieved rain rate and IWP generated by tuned observation error. In figure 10 the red line is the scan orbit of CloudSate to typhoon ‘FAPABI’.

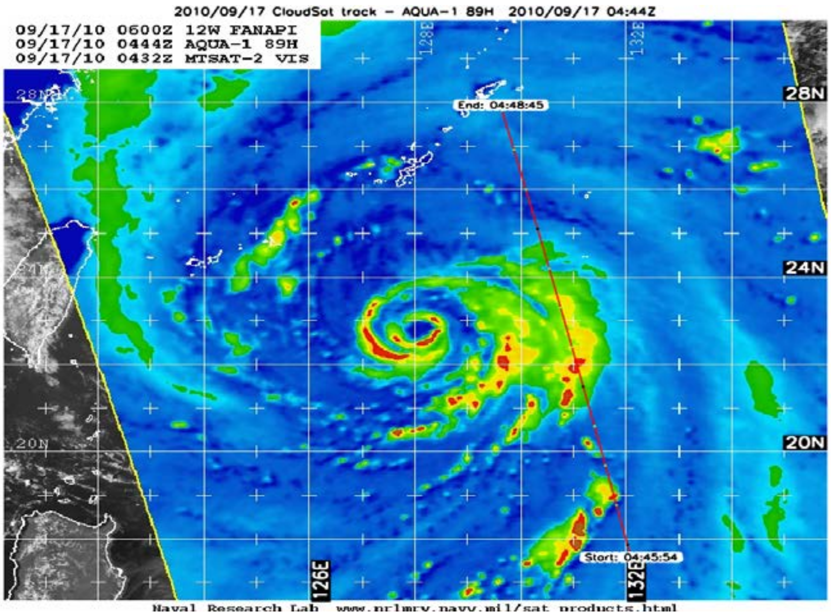

Fig.10. The scan orbit of CloudSate to ‘FAPABI’ (red line)

Retrieved rain rate and IWP of typhoon ‘FAPABI’ are interpolated onto CloudSate’s orbit by bilinear interpolation to make comparison. Figure 11 are the comparison of retrieved water materials to CloudSate. From figure 11 it is clear that the retrieved rain rate and ice water path has a similar distribution to CloudSate. To the magnitude, the retrieved water materials are basically the same as CloudSate considering the resolution of two observation data is different (AMSU-A is 45KM while CloudSate is about 2km), which means that the big value in high resolution data (CloudSate) may be averaged down by low resolution data (AMSU).

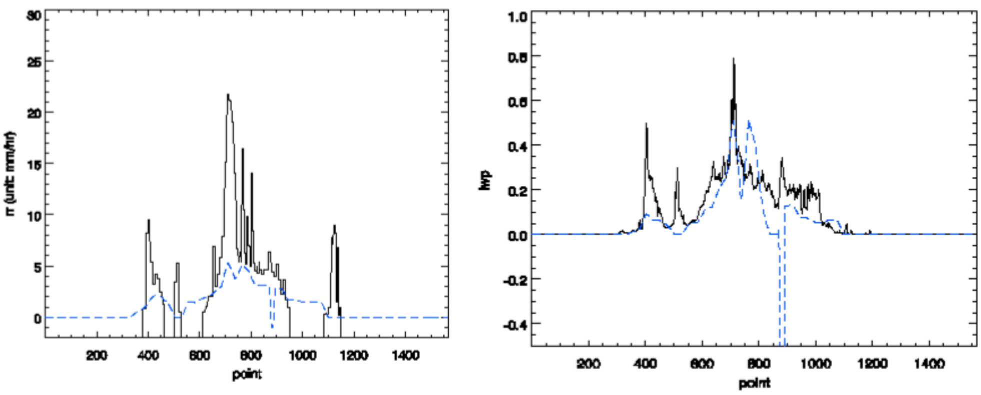

Fig.11. The comparison of retrieved water material (dotted) to CloudSate (in solid) of rain rate (left, unit in mm/hr) and IWP (right, unit in mm)

4. Conclusion

For the reason of simulation error which contains in observation error changes in unclear sky while the change is not considered in retrieval, satellite microwave sounding data retrieval under unclear sky such as cloudy and precipitation are not so accurate as that in clear sky. Observation error could be diagnosis and tuned according to the hypothesis that the relationship between background and observation error are unrelated in data assimilation. In this paper, after using quality control to divide observation into three different weather conditions such as clear sky, cloudy and precipitation, Desrozier's approach of diagnosis and tuning of observation error is used in retrieval of AMSU sounding data at all sky conditions under the basis of MIRS. Comparing of retrieval result between original and tuned observation error shows that the error tuning could get observation error of more accuracy and make retrieval improved at all weather conditions, in details the retrieved temperature profiles of tuned observation error are more close to verification data than original. Besides retrieved water materials such as ice water path and rain rate are accord with CloudSate data also.

What should be stated out here is firstly the satellite visible light figure observation is used in verification to the effect of quality control in dividing weather condition, but the figure data tends to be personal error in practice, numerical data such as cloud top height may be more suitable. Secondly is that the error diagnosis and tuning has no improvement effect in humidity retrieval at this experiment, more researches will be carried out in future.

5. References

Courtier,P. and Talagrand, O. 1987 Variational assimilation of meteorological observations with the adjoint vorticity equation. II: Numerical results. *Q. J. R.Meteorol. Soc.*, 113, 1329–1347

Talagrand,O. 1997. Assimilation of observations; an introduction. *J. Meteorol.Soc.Japan*, 75, 191–209

Hollingsworth, A. and Lonnberg, P. 1986 The statistical structure of short-range forecast errors as determined from radiosonde data. Part I: The wind field. *Tellus*,38A, 111–136

Desroziers, G. and Ivanov, S. 2001 Diagnosis and adaptive tuning of observation-error parameters in a variational assimilation. *Q. J. R. Meteorol. Soc.*, 127, 1433–1452

Talagrand, O.1999, 'A posteriori verification of analysis and assimilation algorithms'. p. 17–28 in Proceedings of Workshop on diagnosis of data assimilation systems, 2–4 November 1998, ECMWF, Reading, UK

Chapnik, B., Desroziers, G., 2006 Diagnosis and tuning of observational error statistics in a quasioperational data assimilation setting. *Q. J. R. Meteorol. Soc.*, 132, 543–565

Dee, D. P. and da Silva, A. 1999, Maximum-likelihood estimation of forecast and observation error covariance parameters. Part I: Methodology. *Mon. Weather Rev.*, 127, 1822–1834

Liu Z-Q, Rabier F. 2003. The potential of high-density observations for numerical weather prediction: A study with simulated observations. *Q. J. R. Meteorol. Soc.* 129: 3013–3035.

Kalnay E. 2003. *Atmospheric modelling, data assimilation and predicatability*. Cambridge University Press.

Desroziers G, Berre L, Chapnik B, Poli P. 2005. Diagnosis of observation, background and analysis-error statistics in observation space. *Q. J. R. Meteorol. Soc.* 131: 3385–3396.

Rutherford ID. 1972. Data assimilation by statistical interpolation of forecast error fields. *J. Atmos. Sci.* 29: 809–815.

Hollingsworth, A. and Lonnberg, P. 1986 The statistical structure of short-range forecast errors as determined from radiosonde data. Part I: The wind field. *Tellus*, 38A, 111–136

Bormann N, Saarinen S, Kelly G, Th´epaut J-N. 2003. The spatial structure of observation errors in Atmospheric Motion Vectors from geostationary satellite data. *Mon. Weather Rev.* 131: 706–718.

Niels Bormann and Peter Bauer Estimates of spatial and inter channel observation-error characteristics for current sounder radiances for numerical weather prediction. I: Methods and application to ATOVS data, *Q. J. R. Meteorol. Soc.* 136: 1036–1050, April 2010 Part B1036-1050

Bormann N, Collard A, Bauer P. 2010. Estimates of spatial and interchannel observation-error characteristics for current sounder radiances for numerical weather prediction. II: Application to AIRS and IASI. *Q. J. R. Meteorol. Soc.* 136(.): Submitted. (Paper II.).

Garand L, Heilliette S, Buehner M. 2007. Interchannel error correlation associated with AIRS radiance observations: Inference and impact in data assimilation. *J. Appl. Meteorol.* 46: 714–725.

Daley, R. 1991 *Atmospheric data analysis*. Cambridge atmospheric and space series. Cambridge University Press, UK

Fisher M. 2003. Background error covariance modeling. In *Proceedings of the ECMWF Seminar on Recent Developments in Data Assimilation from Atmosphere and Ocean*. ECMWF: Reading, UK; pp 45–64.

Stewart L, Cameron J, Dance S, English S, Eyre J, Nichols N. 2009. '*Observation error correlations in IASI radiance data*'. Mathematics Report Series 1/2009, 26pp. Univ. Reading: Reading, UK.