The Regression Retrieval of the Clear TOVS Radiance

Jing Shan Hong* **Dongsoo Kim ****

* Central Weather Bureau, Taipei, Taiwan, R.O.C.
** Cooperative Institute for Research in Environmental Sciences (CIRES)
University of Colorado/NOAA, Forecast Systems Laboratory, Boulder, Colorado U. S. A.

1. Introduction

Satellite data are playing an increasing important role in Numerical Weather Prediction (NWP). For example, the retrieval profiles (SATEMs) generated by NOAA/NESDIS are used in many operational NWP center, e.g. the Central Weather Bureau (CWB) in Taiwan. Though investigations have shown that the SATEMs have the positive impact on the southern hemisphere and neutral or negative impact on northern hemisphere, the data also show a systematic, air mass dependent bias because the SATEMs contain the statistical or climatological component (Andersson et al. 1991). The retrieval of the satellite radiance, the TIROS Operational Vertical Sounder (TOVS), has a tendency in the recent year to use the NWP forecast as the background informations which provide a better estimate of the atmospheric state than the climate (Eyre et al. 1993). In this paper we present a new scheme, statistic regression based, to retrieve TOVS radiance increment to the atmospheric profile increment under the cloud-free cases. The radiance increment is calculated from the computed radiance which is simulated by the RTTOVS radiative transfer model (Eyre 1991) using the background based on 12 hours forecast produced by the operational Limited Area Forecast System (LAFS) in CWB. The retrieved profile compared statistically with the colocated RAOB measurements and the case study will be presented in the following sections.

2. Methodology

Fig. 1 is the flow chart of the procedure to retrieve the TOVS radiance in the paper. The procedure is mainly composed by three major components. First, the computed radiances are simulated at the footprint of the satellite observations from the background profiles provided by the NWP forecast. The informations of the cloud fraction used in the forward model are estimated by the observed AVHRR pixels within the HIRS field of view (*fov*). For considering the clear case only in the paper, the complexity of the estimation of the cloud information will be greatly reduced. Second, take the difference between the observed radiances and computed radiances as the observation increment. Third,

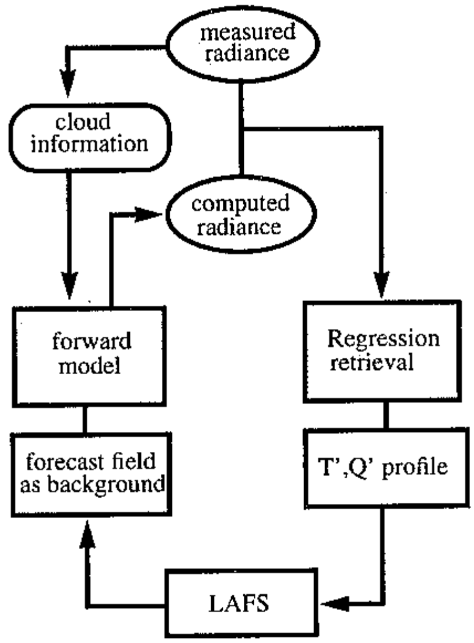

Fig. 1: The procedure to assimilate the TOVS radiance to the LAFS (Limited Area Forecast System).

retrieve the observation increment of the radiances to the profile increment and assimilate to the forecast system. This paper only focuses on the performance of the regression retrieval algorithm under the clear sky. The impact to assimilate the retrieval products to the forecast model will be discussed in future.

Here we use the constrained least square estimation proposed by Kim (1993) to retrieve the observed radiance increment:

$$T' = M(y) + S_{yx}\left[\lambda E + S_{xx}\right]^{-1}\left[x - M(x)\right], \quad (1)$$

where T' is the retrieved increment profile. x and y are the observation increments in brightness temperature (HIRS channel 3 - 8 and 10 - 15), and in profile. There are 22 elements in the retrieved/observed increment profile vectors which are composed of 11 mandatory pressure levels in temperature (1000 - 100 hPa), 7 levels in mixing ratio (1000 - 300 hPa), the surface temperature, mixing ratio, pressure and skin temperature. Thus, (1) provides a simultaneous retrieval of the temperature and humidity profile as well as the surface informations. The matrix opera-

Corresponding author address:
Jing-Shan Hong, Computer Center, Central Weather Bureau, 64 Kuang-Yuan Road, Taipei, Taiwan, R.O.C.
e-mail: rfs14@cwb.gov.tw

tor are defined as follows:

$$M(y) = \langle \mathcal{P}_r - \mathcal{P}_b \rangle$$
$$M(x) = \langle \mathcal{R}_o - \mathcal{R}_b \rangle$$
$$E = \mathrm{cov}\,(\mathcal{R}_r - \mathcal{R}_b, \mathcal{R}_r - \mathcal{R}_b)$$
$$S_{xx} = \mathrm{cov}\,(\mathcal{R}_o - \mathcal{R}_b, \mathcal{R}_o - \mathcal{R}_b)$$
$$S_{yx} = \mathrm{cov}\,(\mathcal{P}_r - \mathcal{P}_b, \mathcal{R}_o - \mathcal{R}_b)$$

where

$\mathcal{R}_o$: measured radiance

$\mathcal{R}_b$: computed radiance from NWP profiles

$\mathcal{R}_r$: computed radiance from RAOB profiles

$\mathcal{P}_b$: NWP profiles of T and q

$\mathcal{P}_r$: RAOB profiles of T and q

$\langle\ \rangle$ stands for smaple mean and cov (a, b) is the covariance of variable a and b for a large sample which are all collocated at RAOB site. The scalar λ in (1) is the weighting of E relative to S_{xx}.

The regression retrieval algoritm presented in the paper has lots of advantage, such as

(1). Matrix E and S_{xx} are symmetric and positive so that the inverse of $S_{xx} + \lambda E$ is numerically stable.

(2). Since the calculation in (1) is only a matrix multiple, the high computing performance can be reached once the covariance matrices are given.

(3). E and S_{xx} implicitly contain the systematic error in NWP forecast as well as the forward radiative transfer model. It is important especially when cloud exists and the great uncertainty of the surface information.

(4). Since the solution in (1) is the increment profiles, it is the most suitable to ingest them to the OI (Optimal Interpolation) module as an independent measurement and is able to avoid the correlations between the background and observation error.

3. Data processing

In order to construct the covariance matrix in (1), it is necessary to collect the measured radiances and profiles from NWP forecasts and RAOB measurements long time enough to ensure the statistical stability. Since the error covariance matrix in (1) are collocated at the RAOB site, the NWP forecasts are interpolated from the model grid system to RAOB site and the represented observed brightness temperatures at the RAOB site are the average of the possible clear *fovs* which locate within the area of $\pm 0.5°$ centered at the RAOB site.

Incorporating the assimilation cycle of LAFS, which is twice per day, the satellite paths over Taiwan are chosen within a time window of $\pm$ 3 hours at synoptic time, 00Z and 12Z. It requires the multi-channel measurements in (1) to retrieve the the satellite sounding data. In this study only the HIRS channels 3-8 and 10-15 have been considered. Channel 1 and 2 are excluded because their peak level of the weighting function are far away from the interested retrieval levels (lower than 100 hPa) in (1). because of the bias due to the reflection of of the sunlight in daytime, the channel 18 and 19 are not used,too.

One of the most concerning on the satellite measurement is the effects of the cloud. What we concerned in the paper is to determine which HIRS pixels are cloud-free. In the paper we adopt the method proposed by Aoki (1980) to match and extract the AVHRR pixels in each HIRS *fov*. The methodology to estimate the cloud information is referred to Kim (1996). The original design in Kim (1996) is to compute the smoothed probability density function of AVHRR pixels in each HIRS *fov* and find the local minimum as breaking point to partition cloud mass. Thus each cloud mass will have a unique representative value implying a unique cloud-top height. The simulation results in Kim (1996) show the method has the ability to estimate the cloud information extending to multi-levels. Following Kim's algorithm, the density function of the AVHRR pixels will exhibit a single peak pattern as the atmosphere is cloud-free or fully overcast by single level cloud. In order to distinguish the clear or fully overcast case, the difference between the observed and computed TOVS (from the NWP forecast background) channel 8 brightness temperatures will be a criterion. As the differences are greater than 4 K, the case will be treated as fully overcast by single level cloud and rejected to go through the retrieval procedure.

4. Results

Two series of experiments are designed to demonstrate the performance of the regression retrieval. The first experiment, hereafter referred to as June case, we collocate the data from 15 May to 20 June, 1997 to construct the error covariance matrices. There are total 248 clear samples are collected at RAOB site during the period. The sensitivity (not shown) shows the statistics of the error covariance matrix is getting stable as the sample is more than 150. Based on the database, the retrieval procedures are undergoing from 20 June to 1 July, 1997 and verified by the RAOB measurements. The second experiment, hereafter referred to as August case, the covariance matrices are constructed by the data collected from 20 June to 1 August, 1997 (239 cases) and verified from 1 August to 15 August, 1997. The two experiments represent two kind of weather regime. The June case is sampled during the Mei-Yu season which is the seasonal transition and the atmosphere is more convective unstable. The August case is a typical summer one which is mainly dominated by the subtropical High and tends more stable. Because of seasonal dependency of the regression retrieval, the two experiments provide a chance to see the difference under the various weather type.

Fig. 2 is the mean bias of the 12 hour forecast error and the retrieved profile for the temperature and mixing ratio in June case (86 samples). Fig. 3

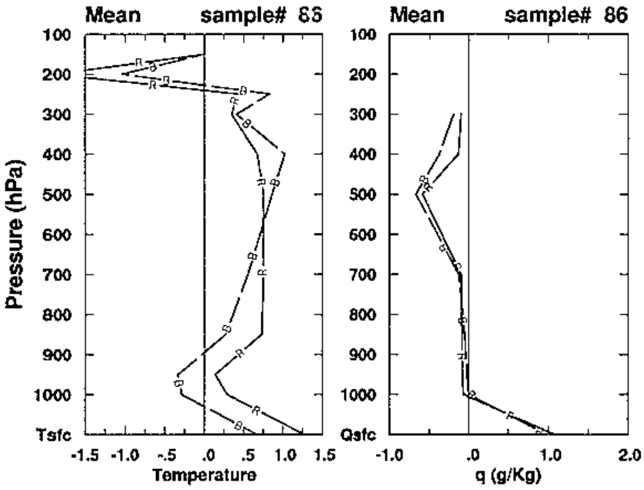

Fig. 2: The mean bias (June case) of the 12 hour forecast error and the retrieved profile for the temperature and mixing ratio. There are total 86 samples collected from 20 June to 1 July, 1997.

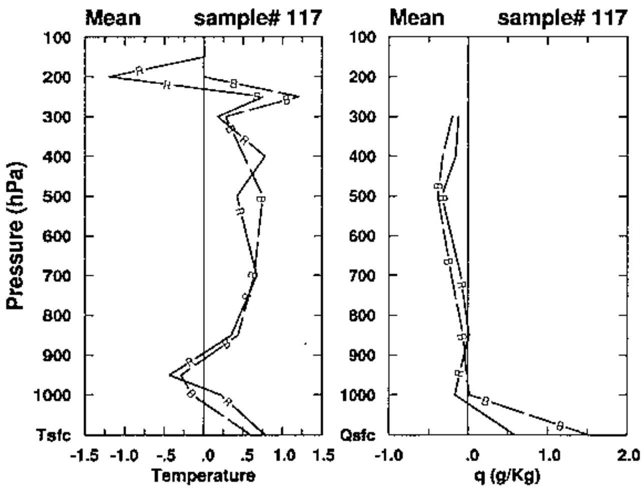

Fig. 3: As Fig. 2 but for August case, there are 117 samples collected from 1 August to 15 August.

is the August case (117 samples). The bias of the retrieved temperature profile in Fig. 3 is of the same order as the background but the June case in Fig. 2 is somehow larger below 500 hPa. The mean bias of the mixing ratio profile are quite similar in Fig. 2 and 7. These results are not surprised because the retrieval procedure is firstguess dependent so that the mean bias of the retrievals has similar behavior as the background.

Fig. 4 and 5 are the root mean square error of June and August case. The reference ground truth is supposed to the RAOB measurements. The figure shows the root mean square error of the forecast temperature error is about 1 ~ 1.5 K within the troposphere and up to 2.1 K at surface. Comparatively, the retrieved profile is only 1 K through the troposphere and 1.4 K at surface, which is significantly smaller than the background error. The mixing ratio profiles also have the improvement of 0.2 ~ 0.4 g/Kg in troposphere and up to 1.2 g/Kg at surface.

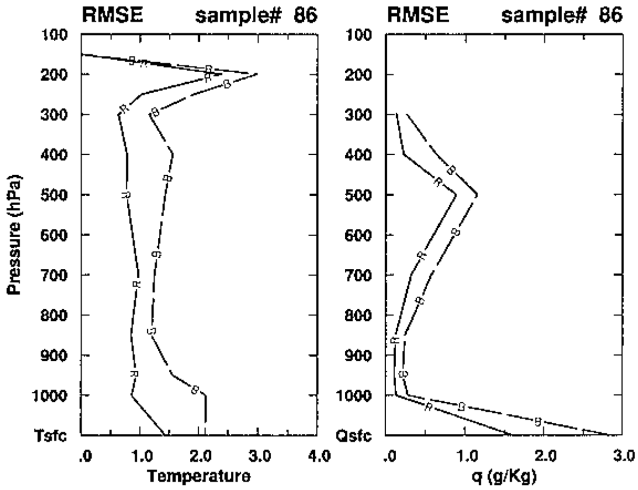

Fig. 4: The root mean square error of the 12 hour forecast error and the retrieved profile for the temperature and mixing ratio for June case.

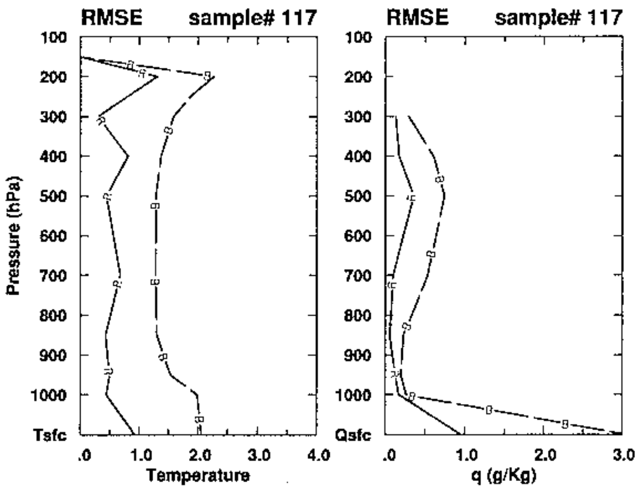

Fig. 5: As Fig. 4 but for August case.

The improvement implies the retrieved profiles are closer to the RAOB profiles than the 12 hour forecast profiles do and it may have the positive corrections as ingest the retrieved products around the sparse observation area.

Because of the unification of the summer weather system, the forecast errors in Fig. 5 reduce a lot, especially the mixing ratio in the mid-troposphere. The figure shows that the smaller root mean square error of the retrieved increments up to 0.5 K in temperature and less than 0.4 g/Kg in mixing ratio are achieved. The results are noticeably better than the June case.

Fig. 6 shows the contour field of the retrieved temperature on 700 hPa for every 3 HIRS *fov* for the case on 0928 LST 27 Aug 1997. As expected, the pattern of the retrieved field (solid line) and the background field (12 hour model forecast, dashed line) are quite similar. It is reasonable that the improvement of the retrievals on the background will be limited if the forecast error is small (Eyre 1989).

In spite of the limited correction, the figure shows that the retrieved field on 700 hPa enhances the temperature perturbation near Japan and Korea. By the way, there exits larger difference between the retrieval and background field in the sub-tropical area (south of 30° N)and is also shown in other cases (not shown here). Such a geographic dependency of the retrievals may be due to the different air mass characteristic between the mid-latitude and subtropical area.

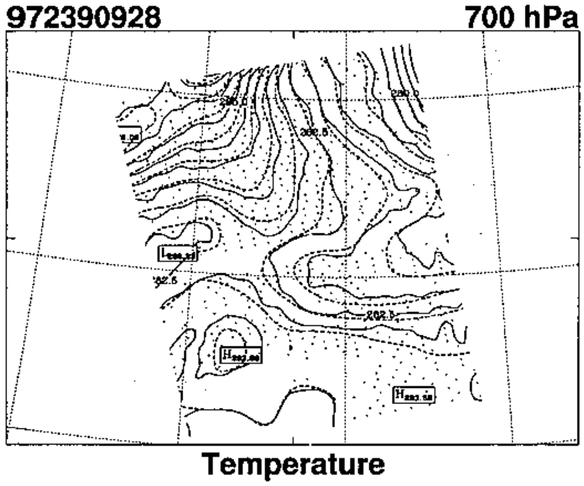

Fig. 6: The temperature field of 700 hPa for the path on 0928 LST, 27 Aug. 1997. The dashed lines is the 12 hour forecast field and solid lines are the retrieved field. The contour interval is 0.5 K. The dots present the HIRS *fov* available in the path.

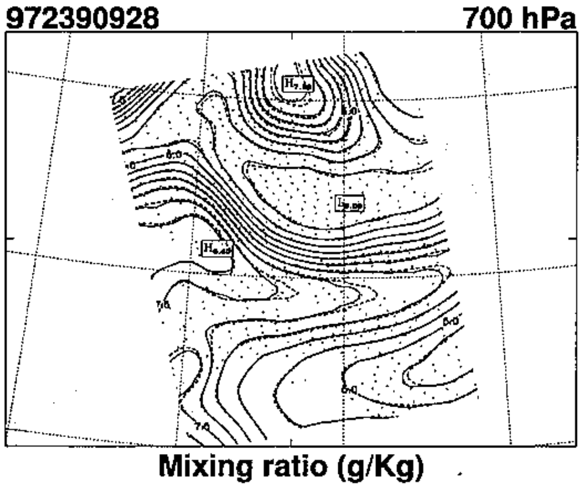

Fig. 7: As Fig. 6 but for mixing ratio. The contour interval is 0.5 g/Kg.

The differences of the mixing ratio field in 500 hPa (not shown) are more prominent than 700 hPa (Fig. 7). It is consistent with the results in Fig. 3-4 that both the retrieval and background have the larger humidity errors in mid-troposphere. The figure also shows that the retrieval field tend to weaken the moist tongue near 30° N. The results from the case study are encouraged that the regression retrieval algorithm exhibits the characteristics which is local coherent in horizontal and vertical. It is important as ingesting the retrieval products to the analysis module.

5. Conclusions

There are many efforts have to be added to the current retrieval framework. For example, first, we didn't discuss further the possible impact due to the effects of the reflected solar radiation though we exclude the HIRS channel 18 and 19 which is most sensitive to the sunlight reflection. Second, the retrieval performance under different weather regime has to be further study. Third, to include the cloud informations to the regression retrieval procedure to inverse the cloudy profile will be the next step and finally, the impact to assimilate the retrieval products to the forecast model is undergoing now.

Reference

Andersson, E., A. Hollingsworth, G. A. Kelly, P. Lönnberg, J. Pailleux, and Z. Zhang, 1991: Global observing system experiments on operational statistical retrievals of satellite sounding data. *Mon. Wea. Rev.,* **119,** 1851 - 1864.

Aoki, T., 1980: A method for matching the HIRS/2 and AVHRR pictures of TIROS-N satellites. *Tech. Note* No. 2, Met. Sat. Center, Japan, 15 - 26.

Eyre, J. R., 1989: Inversion of cloudy satellite sounding radiances by nonlinear optimal estimation. I: Theory and simulation for TOVS. *Quart. J. Roy. Meteor. Soc.,* **115,** 1001 - 1026.

Eyre, J. R., 1991: A fast radiative transfer model for satellite sounding system. *ECMWF Tech. Memo.,* **176,** 28pp.

Eyre, J. R., G. A. Kelly, A. P. McNally, E. Andersson, and A. Persson, 1993: Assimilation of TOVS radiance information through one dimensional variational analysis. *Quart. J. Roy. Meteor. Soc.,* **119,** 1427 - 1463.

Kim, D., 1993: Application of cluster analyses of radiance data measured by satellite and computed from forecast profiles. *Technical Proceedings of the Seventh International TOVS Study Conference,* 10 - 16 Feb, Igls, Austria, 209 - 220.

Kim, D., 1996: Multilevel cloud parameter estimation combing HIRS and AVHRR data - a simulation study. *Preprints, Eighth Conf. on Satellite Meteorology and Oceanography,* 28 Jan - 2 Feb, Atlanta, GA.