THE IMPACT OF TOTAL SYSTEM NOISE ON A TOVS RETRIEVAL SYSTEM

M. J. Uddstrom

National Institute of Water and Atmospheric Research
Wellington, New Zealand

1. INTRODUCTION

All physical retrieval algorithms, whether explicit or implicit, have the following general form:-

$$\hat{T} - \overline{T} = D(r(T,q) - R(\overline{T},\overline{q}))$$

Where $\overline{T}$ and $\overline{q}$ indicate an assumed initial state of the atmosphere, R is the radiative transfer model operator, and r is the satellite "measurement" operator on T and q, the true state of the atmosphere. D is the retrieval operator, and $\hat{T}$ the resultant estimate, or retrieval, of the true state of the atmosphere.

The most significant aspect of this equation is that the retrieval operator is applied to a difference signal, which incorporates information from the satellite measurement process and a radiative transfer model. As a consequence, the retrieval operator is susceptible to a number of error sources, that together define the total system noise in the measurement and retrieval processes. Additionally, there is no a priori expectation that the mean bias $\overline{(r(T,q) - R(T,q))}$ will be zero for some sample of collocated satellite and radiosonde measurements.

In this paper, estimates of the total system noise for the High resolution Infrared Radiation Sounder (HIRS) and Microwave Sounding Unit (MSU) instruments of the TIROS Operational Vertical Sounder (TOVS) will be discussed, and the consequential impact upon the accuracy of an explicit, optimal, physical retrieval algorithm indicated.

2. TOTAL SYSTEM NOISE

Estimates of the total system noise in the satellite measurement and retrieval process may be computed from samples of collocated radiosonde and satellite soundings.

2.1 Noise Components

There are four error terms that can be identified as contributing to the "measurement" noise in the retrieval process:-

- Satellite measurement errors,
- Radiative transfer model errors,
- Validation data errors, and
- Non representativness errors.

The first incorporates the fundamental errors arising from the instrument's optical, detector and electronic characteristics, together with errors introduced by ensuing processing, such as, contributions from calibration, cloud clearing and limb correction procedures.

Line by line radiative transfer models also suffer errors due to a fundamental lack of knowledge of molecular collision processes, such as the characteristics of absorption features in terms of their pressure and temperature dependence (Chedin et al. 1988). In addition, practical transmittance models utilise further simplifications (Weinreb et. al. 1981) in order to enable the rapid computation of radiances.

Implicit to all validation experiments is the assumption that the true state of the system can be estimated with appropriate accuracy by independent means. Here, this assumption implies that all collocated radiosondes can be relied upon to provide unbiased, precise and accurate estimates of the temperature and water vapour structures of the atmosphere sounded by the satellite. A further implication is that the extrapolated radiosonde temperature measurements are accurate to 0.1 hPa, the top level for the radiative transfer model used here. Neither of these assumptions is correct (Nash and Schmidlin (1987), Schmidlin and Luers, (1987), Uddstrom (1989), McMillin et. al. (1992)).

The last error source, non-representativness, includes all those error terms that relate to the fundamental differences between satellite remotely sensed and in-situ measurements. Specifically, the differences between point (radiosonde) and areal (satellite) measurements, as well as time and distance mismatch considerations. For sufficiently large samples these errors might however be expected to contribute zero bias to the total system noise.

2.2 Total System Noise Estimates

2.2.1 Data sources

The data used in the experiments reported here were extracted from the NOAA/NESDIS collocation archive (termed the DSD5), and consist of approximately 30000 NOAA10 collocations having the following characteristics:-

- Global distribution.
- Sampled from the period October to April (weighted toward January) and the years 1989 to 1991.
- Each radiance observation, whether clear, nstar or cloudy, is collocated with only the closest radiosonde (in either time or distance).
- Data from both day and night soundings over land and sea surfaces are included.
- Radiosonde shortwave radiation errors have been corrected using the method of McInturff et al. (1979). For the U.S.A. VIZ radiosondes in the sample, a long wave correction is applied above 20 hPa.

- The data have been divided into dependent (90%) and independent (10%) samples by random selection without replacement.

2.2.2 The forward model

The radiative transfer model utilised here is essentially identical to the operational NESDIS model (Weinreb et al. 1981), except that the "gamma" transmittance adjustment factors have been set to unity. Surface skin temperatures are estimated directly from the window channels and associated radiosonde profile (Uddstrom 1991).

2.2.3 Radiance temperature adjustments

Radiative transfer models currently in use do not simulate the satellite measurements particularly well since mean radiance temperature differences $\overline{(r(T,q) - R(T,q))}$ show a strong airmass dependence (Kelly and Flobert (1988)). However, all physical retrieval algorithms are built on the assumption that a zero radiance temperature signal will produce a zero perturbation to the first guess profile, therefore either the measured or modelled radiance temperature measurements must be "adjusted" before use in a retrieval algorithm. Generally the measured radiance temperatures are adjusted to "equivalent" modelled radiance temperatures.

To circumvent the airmass dependence problem, the radiosonde data were sorted into airmass classes by Typical Shape Function (TSF) classification (Uddstrom and Wark, 1985) and the measured radiance temperature data adjusted to model values. The adjustment technique used here is the shrinkage estimator of Oman, (1982), a constrained regression method, i.e.:-

$$\hat{\mathbf{y}} = (\mathbf{S}_{yx} + \gamma\mathbf{C}_0^T)(\mathbf{S}_{xx} + \gamma\mathbf{I})^{-1}\mathbf{x} + \mathbf{C}$$

where **x** is the vector of predictors, which includes all the satellite measured channels to be corrected, together with two additional predictors; a land/sea indicator, and the cosine of the solar zenith angle. The predictands $\hat{\mathbf{y}}$, are the model equivalent radiances. A separate equation is derived for each airmass class, and the constraint $\mathbf{C}_0$ is chosen so as to force the significant predictor channels to lie close, in height, to the relevant predictand channels. $\mathbf{S}$ is used to represent the covariance (subscript *xx*) and cross-covariance (subscript *xy*) matrices generated from the dependent sample, and $\mathbf{C}$ is the constant matrix. γ is a ridge smoothing parameter.

Generating shrinkage adjustment equations for each atmospheric class, and applying these to the independent samples, the total system noise properties of the measurement process can be defined. The root mean square (rms) errors for the independent samples of adjusted radiance temperatures are plotted in Fig 1. The airmass classes have been sorted so that class 1 represents the most polar sample, and class 19 the most tropical. It is these errors that define the total system noise in the "measurements" provided to a retrieval operator, and this noise statistic contains the inseparable components of the four noise sources mentioned earlier. These values differ in a significant way from the NEΔT values specified in Planet (1988) (or inferred from the NEΔN values given).

Also it is apparent from Fig. 1 that the opaque water vapour channels (HIRS 11 and 12) are not well modelled, or else the validating atmospheric water vapour measurements do not describe the atmosphere well.

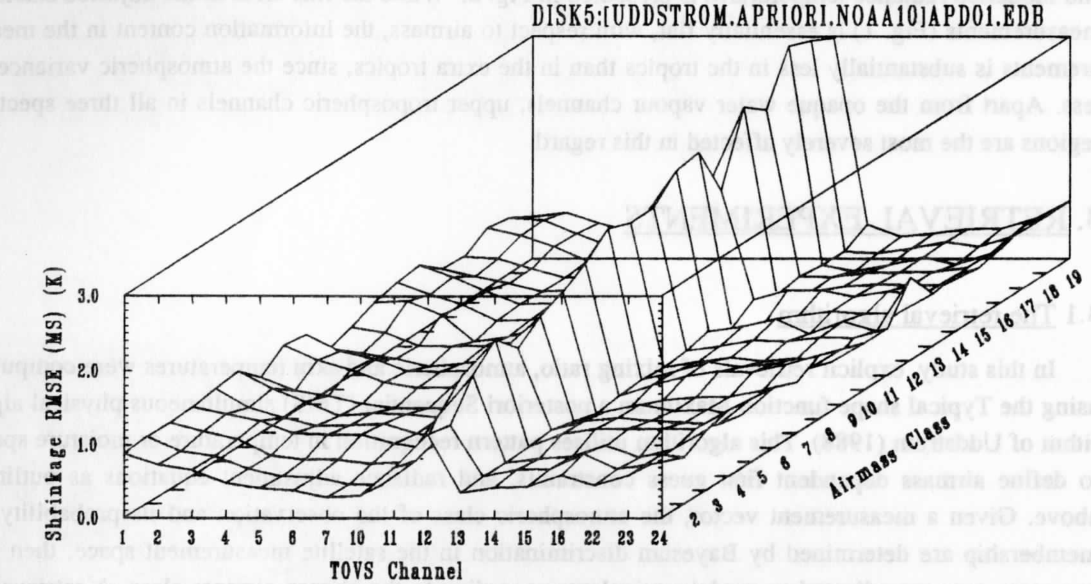

Fig. 1: RMS errors for NOAA10 measured independent radiance temperatures adjusted to model equivalents (by shrinkage estimation).

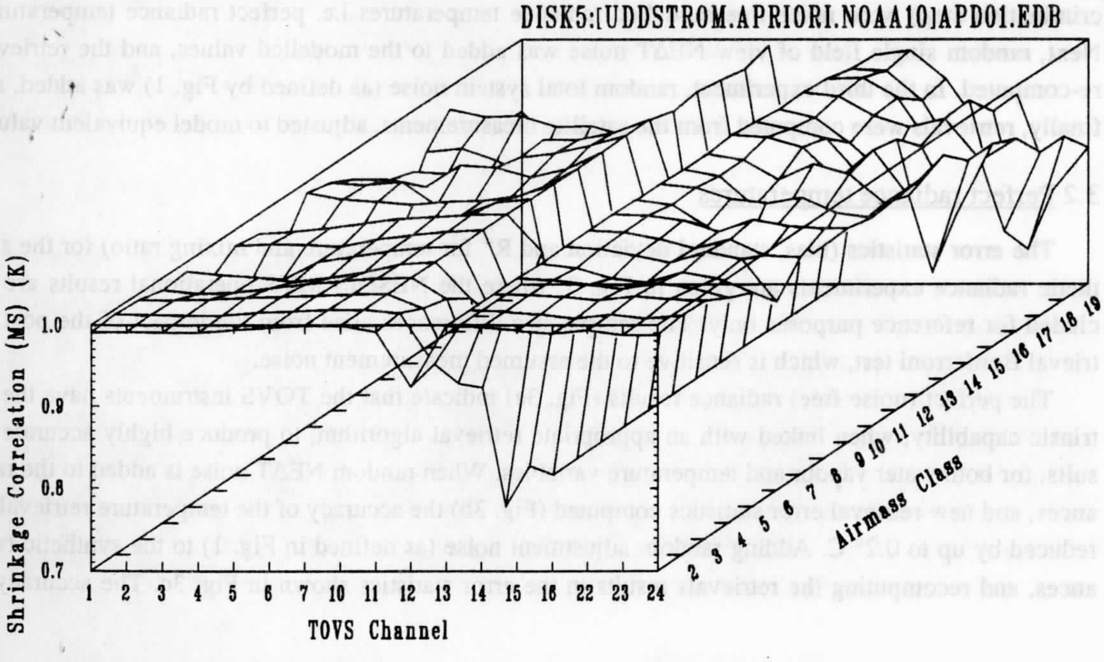

Fig. 2: Correlation between NOAA10 measured, adjusted to model equivalents (by shrinkage estimation), and modelled radiance temperatures for the independent samples.

The correlation between the measured, adjusted to model equivalent radiance temperatures, $\hat{r}(T,q)$, and modelled radiance temperatures is presented in Fig. 2. While the rms error in the adjusted satellite measurements (Fig. 1) is essentially flat, with respect to airmass, the information content in the measurements is substantially less in the tropics than in the extra tropics, since the atmospheric variance is less. Apart from the opaque water vapour channels, upper tropospheric channels in all three spectral regions are the most severely affected in this regard.

3. RETRIEVAL EXPERIMENTS

3.1 The retrieval algorithm

In this study, explicit retrievals of mixing ratio, atmospheric and skin temperatures were computed using the Typical shape function Maximum a posteriori Sequential (TMS) simultaneous physical algorithm of Uddstrom (1988). This algorithm utilises pattern recognition in temperature or moisture space to define airmass dependent first guess constraints, and radiance adjustment equations as outlined above. Given a measurement vector, the atmospheric class of the observation and its probability of membership are determined by Bayesian discrimination in the satellite measurement space, then the measurements are adjusted to model equivalents according to the chosen airmass class. A retrieval is performed using a sequential, simultaneous, physical retrieval estimator. The retrieval portion may be iterated, and a Bonferroni radiance temperature convergence test (Neter et al., 1989) is applied as a final quality control measure.

Four experiments were conducted in order to estimate the impact of system measurement noise on the retrieval algorithm. In the first experiment, the radiances used in the retrieval step (but not the discrimination step) were noise free modelled radiance temperatures i.e. perfect radiance temperatures. Next, random single field of view NEΔT noise was added to the modelled values, and the retrievals re-computed. In the third experiment, random total system noise (as defined by Fig. 1) was added, and finally, retrievals were computed from the satellite measurements, adjusted to model equivalent values.

3.2 Perfect radiance temperatures

The error statistics (bias, standard deviation and R^2 for temperature and mixing ratio) for the synthetic radiance experiments are given in Fig. 3, where the NESDIS MVS operational results are included for reference purposes only. The sample size differences arise from the impact of the post retrieval Bonferroni test, which is sensitive to the assumed measurement noise.

The perfect (noise free) radiance results (Fig. 3a) indicate that the TOVS instruments have the intrinsic capability, when linked with an appropriate retrieval algorithm, to produce highly accurate results, for both water vapour and temperature variables. When random NEΔT noise is added to the radiances, and new retrieval error statistics computed (Fig. 3b) the accuracy of the temperature retrievals is reduced by up to 0.2° C. Adding random adjustment noise (as defined in Fig. 1) to the synthetic radiances, and recomputing the retrievals results in the error statistics shown in Fig. 3c. The accuracy of the retrievals is now much like that generated by the NESDIS operational processing, with temperature

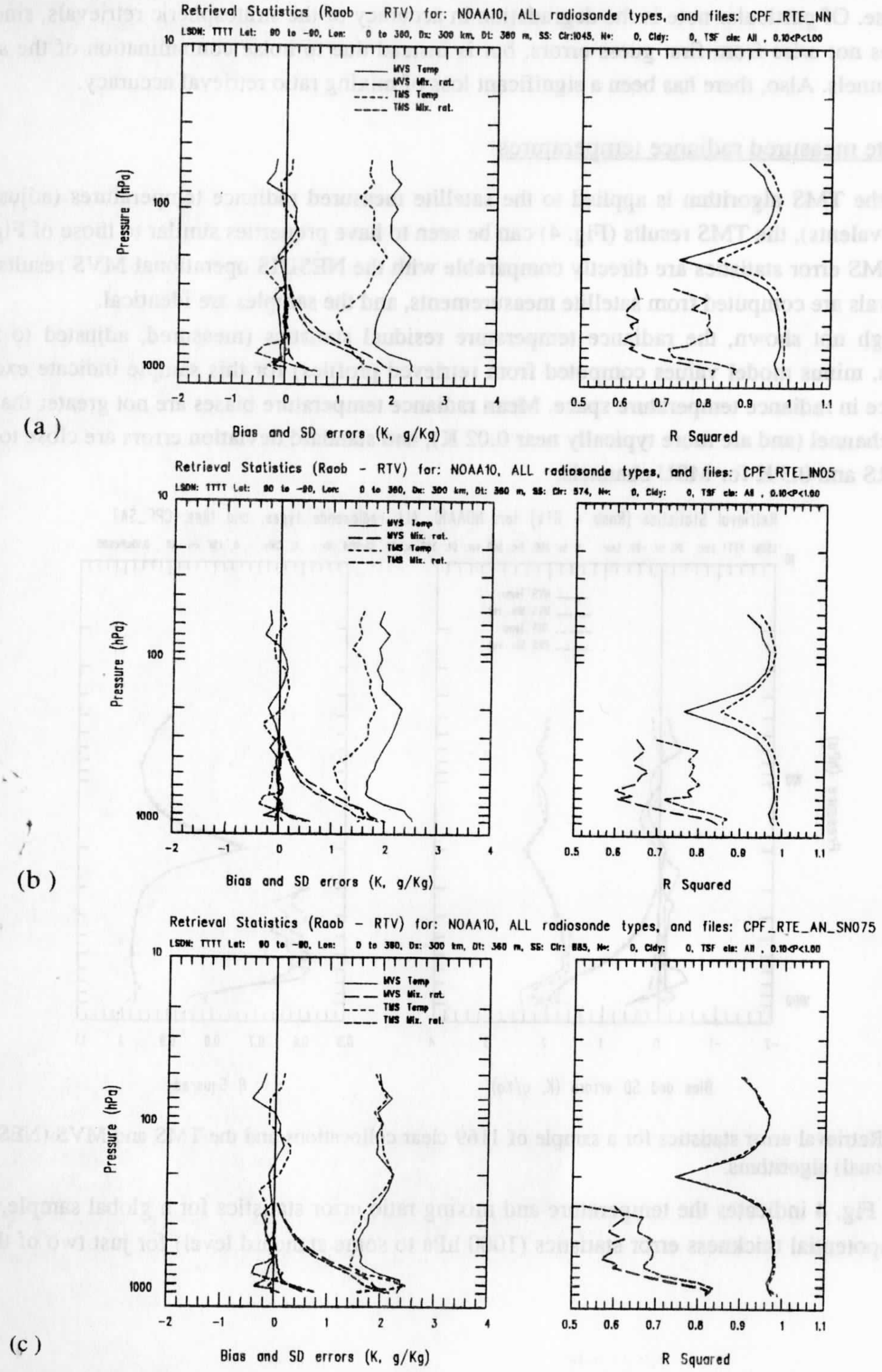

Fig. 3 Retrieval error characteristics calculated from RTE radiances and (a) no noise, (b) radiometer noise, (c) total system noise.

retrieval accuracy at all levels being reduced by approximately 0.5 C from the perfect radiance temperature case. Of particular note is the degradation in accuracy of the stratospheric retrievals, since the change does not arise from first guess errors, but is instead due to noise contamination of the stratospheric channels. Also, there has been a significant loss of mixing ratio retrieval accuracy.

3.2 Satellite measured radiance temperatures

When the TMS algorithm is applied to the satellite measured radiance temperatures (adjusted to model equivalents), the TMS results (Fig. 4) can be seen to have properties similar to those of Fig 3(c). Here the TMS error statistics are directly comparable with the NESDIS operational MVS results since both retrievals are computed from satellite measurements, and the samples are identical.

Although not shown, the radiance temperature residual statistics (measured, adjusted to model equivalents, minus model values computed from retrieved profiles) for this sample indicate excellent convergence in radiance temperature space. Mean radiance temperature biases are not greater than 0.08 K for any channel (and are more typically near 0.02 K), and standard deviation errors are close to 0.1 K for the HIRS and 0.3 K for MSU channels.

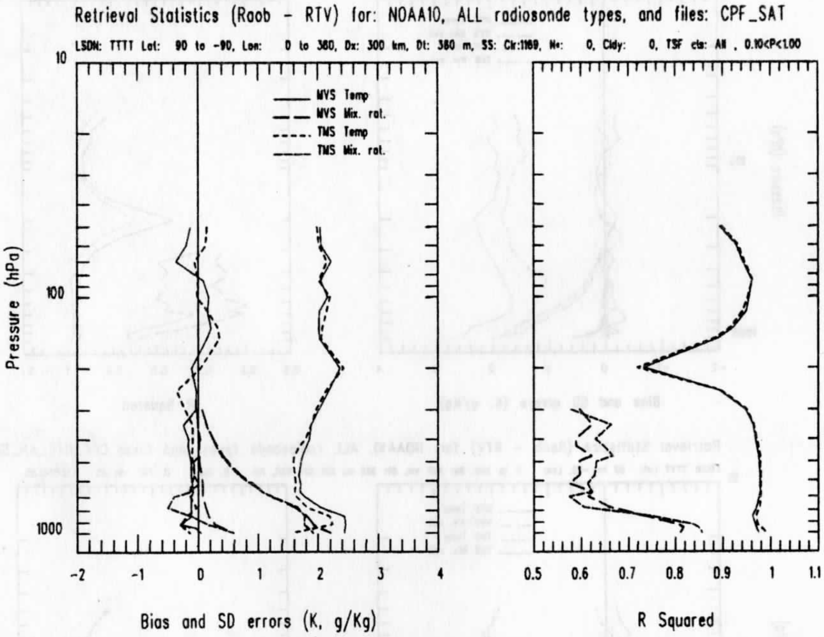

Fig. 4 Retrieval error statistics for a sample of 1169 clear collocations and the TMS and MVS (NESDIS operational) algorithms.

While Fig. 4 indicates the temperature and mixing ratio error statistics for a global sample, Fig. 5 shows geopotential thickness error statistics (1000 hPa to some standard level) for just two of the embedded TSF class samples. Fig. 5(a) is a tropical sample (Class 19 in Figs. 1 and 2), while Fig. 5b is a typical winter mid-latitudes sample.

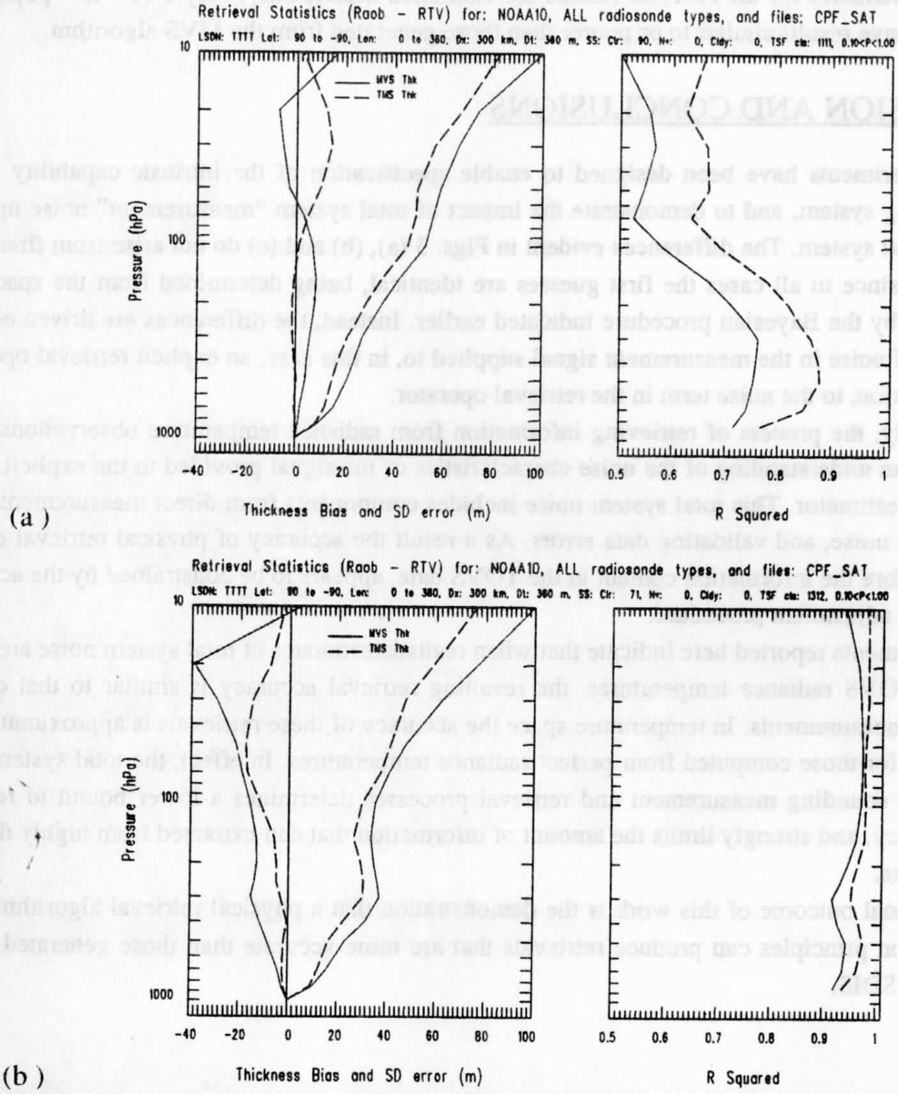

Fig. 5 Geopotential thickness error statistics, bias, standard deviation (SD) and R^2 for standard levels (1000 to 850 hPa, 1000 to 700 hPa etc.) and two airmass classes (a) Tropical, sample size; 90, and (b) Mid-latitudes, sample size; 71. TMS and NESDIS operational results are shown.

The error statistics shown in Fig. 5 indicate a number of important points:-

- The bias and standard deviation error characteristics of retrievals are not uniform across airmasses.
- There is much less information (as indicated by the R^2 statistic) in the tropical retrievals than in the mid-latitudes samples, although the absolute error characteristics are similar.
- There are compensating errors above and below the tropopause.

- The TMS algorithm has error characteristics that are comparable too or better than those of the the NESDIS operational scheme.

When the statistics for all 19 TMS classes are examined individually, only a few low population class samples have results similar to or poorer than those generated from the MVS algorithm.

4. DISCUSSION AND CONCLUSIONS

These experiments have been designed to enable specification of the intrinsic capability of the TOVS sounding system, and to demonstrate the impact of total system "measurement" noise upon an optimal retrieval system. The differences evident in Figs. 3 (a), (b) and (c) do not arise from first guess dependencies, since in all cases the first guesses are identical, being determined from the spacecraft measurements by the Bayesian procedure indicated earlier. Instead, the differences are driven entirely by the effect of noise in the measurement signal supplied to, in this case, an explicit retrieval operator, and by implication, to the noise term in the retrieval operator.

Accordingly, the process of retrieving information from radiance temperature observations is dependent upon an understanding of the noise characteristics of the signal provided to the explicit or implicit retrieval estimator. This total system noise includes components from direct measurement noise, forward model noise, and validating data errors. As a result the accuracy of physical retrieval estimators, and therefore the information content in the TOVS data, appears to be constrained by the accuracy of the radiance adjustment procedure.

The experiments reported here indicate that when realistic estimates of total system noise are added to modelled TOVS radiance temperatures, the resulting retrieval accuracy is similar to that derived from satellite measurements. In temperature space the accuracy of these retrievals is approximately 0.5 K poorer than for those computed from perfect radiance temperatures. In effect, the total system noise in the satellite sounding measurement and retrieval processes determines a lower bound to retrieval product accuracy, and strongly limits the amount of information that can extracted from highly dependent observations.

An additional outcome of this work is the demonstration that a physical retrieval algorithm based on classification principles can produce retrievals that are more accurate than those generated operationally by NESDIS.

References

Chedin, A., 1988: Editor; Report on the ITRA (Intercomparison of Transmittance and Radiance Algorithms) Campaign and Workshop. Pub. International Radiation Commission.

Luers, J. K., 1990: Estimating the temperature error of the radiosonde rod thermistor under different environments. *J. Atmos. Oceanic Tech.,* 7, 882-895.

Kelly, G.A., and J.F. Flobert, 1988: Radiance Tuning. In the technical *Proceedings of the Fourth International TOVS Study Conference.* Igls, Austria, Pub. CIMSS, University of Wisconsin, 99-117,

McInturff, R.M., F.G. Finger, K.W. Johnson and J.D. Laver 1979: Day - night differences in radiosonde observations in the stratosphere and troposphere. NOAA Technical Memorandum NWS NMC **63**, 47 pp.

McMillin, L.M., M.J. Uddstrom and A Coletti, 1992: A procedure for correcting radiosonde reports for radiation errors. *J. Atmos. Oceanic Tech.*, **9**, 801-811.

Nash, J. and F.J. Schmidlin, 1987: Final Report of the WMO International Radiosonde Comparison (U.K. 1984, U.S.A. 1985), Instruments and Observing Methods, Rep No. **30**, 103 pp.

Planet, W.G., 1988: Data extraction and calibration of the TIROS-N/NOAA Radiometers. NOAA Technical Memorandum NESS 107 - Rev. 1, 58 pp.

Neter, J., W. Wasserman and M.H. Kutner, 1989: Applied linear regression models, Second Edition. Pub. Irwin, Homewood, IL, U.S.A. 667 pp.

Oman, S.D., 1982: Shrinking towards subspaces in multiple linear regression. *Technometrics,* **24,** 307-311.

Schmidlin, F.J. and J.K. Luers, 1987: US Radiosonde temperature errors, methods of correction and evidence of data improvement. In Extended Abstracts to *6th AMS Symposium on Methods of Observation and Instrumentation*, New Orleans, LA, U.S.A. Pub American Meteorological Society, 463-466.

Uddstrom, M.J. and D.Q. Wark, 1985: A classification scheme for satellite temperature retrievals. *J. Clim. Appl. Meteor.*, **24,** 16-29.

Uddstrom, M.J. 1988: Retrieval of atmospheric profiles from satellite radiance data by typical shape function maximum a posteriori simultaneous retrieval estimators. *J. Appl. Meteor.*, **25,** 515-549.

Uddstrom, M.J. 1989: A comparison of Philips RS4 and Vaisala RS80 radiosonde data. *J. Atmos. Oceanic Tech.*, **6,** 201-207.

Uddstrom, M.J. 1991: Forward model errors. In the technical *Proceedings of the Sixth International TOVS Study Conference*, Airlie, VA, U.S.A., Pub CIMSS, University of Wisconsin, 501-516.

Weinreb, M.P., H.E. Fleming, L.M. McMillin and A.C. Neuendorffer, 1981: Transmittances for the TIROS operational vertical sounder. NOAA Technical Report NESS **85,** 60 pp.

TECHNICAL PROCEEDINGS OF THE SEVENTH INTERNATIONAL TOVS STUDY CONFERENCE

Igls, Austria

10-16 February 1993

Edited by

J R Eyre

European Centre for Medium-range Weather Forecasts
Shinfield Park, Reading, RG2 9AX, U.K.

July 1993