

The Effect of Collocation Radiosonde Errors on the Assessment of the Performance of a Physical Retrieval Estimator

MICHAEL J. UDDSTROM

*New Zealand Meteorological Service
PO Box 722, Wellington, New Zealand*

10 March 1988

1. INTRODUCTION

A necessary prerequisite to the development of a physical retrieval estimator is the requirement that, for a given atmosphere and observing radiometer, radiative transfer in the atmosphere can be modelled to "sufficient" accuracy and precision. Defining what accuracy is "sufficient" is to some degree dependent upon the way a particular retrieval algorithm uses the satellite observations, but the precision of the radiative transfer model should ideally be no worse than the radiometer instrumental measurement errors. If this prerequisite is satisfied then it is expected that physical retrieval schemes will outperform standard regression estimation, if only for the reason that more a priori information can be incorporated into a physical algorithm.

The problem of the indeterminacy of "errors" in the NESDIS fast transmittance model, and the resulting effects on a particular type of physical retrieval estimator will be demonstrated. The retrieval algorithm employed for the analysis is the Typical shape function Maximum a posteriori Simultaneous (TMS) sequential physical algorithm developed in Uddstrom (1988). This estimator utilizes pattern recognition to determine a first guess profile, and the simultaneous retrieval equations are solved using statistical constraints pertinent to the class of atmosphere described by the chosen first guess profile.

2. EXPERIMENT DESIGN

Two kinds of data were employed in this experiment. As the TMS algorithm requires very large samples of a priori data, the retrieval constraints were developed using the radiosonde data archive of the New Zealand Meteorological Service. For each month of the year more than 5000 solar radiation error corrected profiles, from all southern latitudes can be utilised to determine Typical Shape Function (TSF) (Uddstrom and Wark, 1984) a priori first guess profile estimates and retrieval constraints. From the resulting TSF samples, TMS radiance discrimination equations and radiance dependent retrieval constraints were developed using the NESDIS fast parameterised model for the radiative transfer equation (rte). The coefficients and γ values employed were those supplied in the VAX release of the International TOVS Processing Package (ITPP).

The second data set, provided by Dr L. McMillin of NOAA/NESDIS, consists of a global sample of NOAA 7 and 8 satellite observations collocated with solar corrected radiosonde data. The satellite data have been corrected to clear radiance temperatures, and reduced to equivalent nadir measurements. The associated NESDIS (operational) regression retrievals are also given. The collocated data were collected over a period of one year (1983/84) and consist of

some 6000 profiles. Both the TSF, and collocation sample radiosonde temperature profiles were extrapolated to 0.1 hPa using the zonal regression method supplied in the ITPP.

Given the TMS TSF classes for a particular month (in this case, July) and a specific satellite, these same classes can be duplicated in the collocation data set for the same month, by the method given in Uddstrom (1988). However, because of the smallish size of the monthly samples, data from adjacent months had to be included, and northern hemisphere observations were shifted in time by 6 months. Fifteen percent of the collocated sample data, chosen at random, were then set aside as independent data.

For an example month, the collocation (duplicated) TSF sample means are given in fig 1. Using these samples, air mass dependent corrections for the rte modelled radiance temperatures may be computed, and applied to satellite observations used by the TMS retrieval algorithm.

Consequently, three types of retrieval error statistics can be considered: those arising when the TMS algorithm is applied to the rte radiances (suitably perturbed by measurement noise), and those calculated from retrievals by either the TMS or the NESDIS estimators using the satellite observations.

3. ESTIMATION OF RTE MODEL ERRORS

a. Delta estimation

The usual method employed to determine rte model errors relies on computing differences between satellite observed radiances and those generated by an rte model from the associated collocated radiosonde observations. Hence, for a given atmosphere and rte model, equivalent satellite observations are estimated according to (1).

$$\hat{\mathbf{T}}^{sat} = \mathbf{T}^{rte} + \delta \quad (1)$$

Where $\hat{\mathbf{T}}^{sat}$ is an estimate of the vector of radiance temperatures the satellite radiometers would measure if the atmosphere in question had been observed, \mathbf{T}^{rte} is the rte model radiance temperature vector generated for that atmosphere, and δ is a vector of empirically defined corrections for the rte model. Further, the standard deviations of the δ_i ($i=1,2,\dots,n$, for n channels) estimates indicate the precision with which $\hat{\mathbf{T}}^{sat}$ can be computed and therefore the effective "measurement" noise to be used in any optimal physical retrieval algorithm.

Here, the δ vectors may be calculated as a function of TSF identified air masses. Accordingly (1) can be recast to reflect this additional refinement, in which case;

$$\hat{\mathbf{T}}_{TSF}^{sat} = \mathbf{T}^{rte} + \delta_{TSF} \quad (2)$$

If the individual TSF class δ_{TSF} vectors are not significantly different, then (1) will suffice, otherwise air mass dependent δ corrections should be employed in the retrieval algorithm. In the case of the TMS retrieval estimator, where all retrieval constraints are computed in terms of model radiance temperatures [$\delta_{TSF} = 0$], the satellite radiance temperatures are in fact converted to equivalent model radiance temperatures.

The mean values, standard deviations and standard errors of a selection of individual channel δ_{TSF} estimates are plotted in fig 2. Both day and night results are shown, for each of the TSF classes (as identified in fig 1). Considering the general characteristics of these results first, it is evident that whilst the HIRS δ values are similar for TSF classes having similar mean profiles,

there are significant differences for some of the airmass types, for example, classes 21 and 22. Interestingly though, the MSU δ corrections are in all cases rather small, as are the associated class standard deviations. The between class differences are significant in a number of cases, but the δ estimates are, by and large, well defined having small sample standard deviations relative to the expected (0.3 K per IFOV) channel measurement noise. However, apart from these channels, the expected noise in the δ estimates is disappointingly large, relative to both the measurement noise characteristics of the HIRS and the within TSF class radiance temperature variances.

There are clear and systematic differences between the day and night statistics for some of the classes and channels, e.g., class 311. However the most noticeable features in the day and night δ_{TSF} vector differences relate to the surface and near surface channels. For daylight observations these channels show a positive bias relative to the computed radiances whilst the night values show a bias of the opposite sign (see, e.g., channel 8). There are a large number of continental soundings in the collocation sample so this sign change is presumably due to the skin effect. The night soundings often have substantial surface inversions.

Those channels which have an important upper stratospheric component (i.e. channels; 1, 2 and 17) are susceptible to radiosonde extrapolation errors, as demonstrated by the results for channel 1 in fig 2. It would seem that profiles having warm stratopauses are not well extrapolated by the ITPP equations (i.e., classes 12, 22, 32, 41, 42 and 43). This is surprising, since these atmospheres tend to be more tropical in origin, where the atmospheric temperature variance in the middle stratosphere is smaller than that of mid-latitudes and polar atmospheres.

Two further contributing factors to noise in the δ_{TSF} vectors can be identified, even assuming perfect cloud clearance, correction to nadir and collocation data interpolation. The first is the absence of surface information in the collocation data set (apart from surface height) and the second is the inherent limitation in the standard NESDIS transmittance algorithm, where pressures greater than 1000 hPa are not permitted. The first problem was resolved through estimating the surface pressures, temperatures and mixing ratios from a mean sea level pressure climatology and the radiosonde profiles and surface height information. Although this approach doesn't solve the problem, it should remove some of the noise variance arising from this source. The second problem requires further research, although a simple approach is to employ an extrapolation technique (e.g. piecewise continuous cubic splines) to infer the transmittance at the surface pressure and hence compute appropriate radiance temperature corrections.

However, inherent to the delta approach for determining "errors" in an RTE model is the implicit assumption that the sample of collocated radiosondes suffers only from random, zero mean, measurement errors. This would be a satisfactory assumption if the collocation sample used only one kind of radiosonde, or all radiosondes in the sample had been adjusted in some uniform way so as to effectively reduce them to some "standard" radiosonde type. However, results from the WMO International Radiosonde Intercomparison (Nash and Schmidlin, 1987) indicate that significant between sonde biases exist, in part because radiosondes are affected by short and long wave radiation in different ways. Since these errors tend to be correlated in the vertical, they may contribute as much as 1 degree of more to the δ_{TSF} estimate variances for night time observations alone (see Nash and Schmidlin, Table 11.1), in addition to causing erroneous estimates of the δ_{TSF} vectors. The radiosonde bias errors are not limited to the stratosphere either; significant differences exist at all standard levels. Obviously, the day time differences will be even larger because of the rather more variable solar radiation effects. Uddstrom (1988b) has found between sonde temperature biases as large as 1.5 degrees (see fig 3.), from a comparison of Väisälä RS80 DIGICORA data with Philips RS4 (similar to the VIZ sonde) data, corrected to effective night time values.

b. An alternative correction method

If the radiosonde data in collocation samples cannot be corrected to yield measurements equivalent to some standard radiosonde, because either the particular radiosondes employed are unknown, or consistency corrections do not exist for all of the sondes utilized, then the δ method determination of "errors" in the rte model radiance temperatures is untenable. This is for two reasons. Firstly, the δ corrections will always be erroneous, leading to what appear to be incorrect retrievals, when compared with the collocated radiosondes, and secondly the "apparent" measurement noise, the sum of both instrumental and δ estimation noise, will be so large that all retrievals will lie close to the first guess estimates.

An alternative approach to this problem is to estimate equivalent rte model radiance temperatures from the observed radiances (as required by the TMS estimator) in such a way that the errors to which the δ estimation procedure is susceptible are minimised. This is easily arranged, in principle, by rewriting the correction equation as:-

$$\hat{\mathbf{T}}_{TSF}^{rte} = \mathbf{D}_{TSF} \mathbf{T}^{sat} + \mathbf{C}_{TSF} \quad (3)$$

where \mathbf{D}_{TSF} and \mathbf{C}_{TSF} are respectively the regression coefficient and constant matrices derived from the dependent TSF class samples. In the following discussion, this will be called the **D** method.

For a regression on all channels except 9, 20 and 21—implying therefore, perfect cloud clearance of the satellite observed radiances—the resulting rms error (and bias) statistics for (3) applied to the independent collocation data, are given in fig 4a. The same error statistics for the δ_{TSF} method of (2) applied to the independent data are plotted in fig 4b. With only a few exceptions, the rte radiance temperature estimates derived from (3) are superior to those derived from (2).

This result is perhaps a little surprising, given the generally small samples used to derive the **D** equations, and the gross cloud clearance assumption. However, the **D** method of computing rte model "errors" does have a number of advantages when it is known a priori that the "truth" radiosonde profiles are, in reality, variably biased estimates of the true state of the atmosphere. Because a component of radiosonde error is vertically correlated, using more than one channel's information to fit the rte model "error" should yield better error correction equations, when the results are validated against the biased data. Therefore, this method may only be useful for collocation samples where the radiosonde data have not been reduced to some common standard. In practice a combination of the δ and **D** methods could be employed in order to reduce the impact of possible undesirable side effects resulting from the **D** method, such as the propagation of cloud contamination errors into channels not previously affected.

4. RETRIEVAL RESULTS

a. NESDIS method

Plotted in fig 5 are error statistics for the operational NOAA/NESDIS retrieval scheme and the independent data. As expected, the sample standard deviation statistics indicate that the regression retrievals, in general, display less sample variance than that in the "true" sample.

In considering the rms error results, it is important to keep in mind the problem of radiosonde errors discussed in the previous sections. Since the regression retrieval coefficients are derived from the NESDIS collocation archive, the resulting equations will also have the property

that they minimise the effect of errors arising from radiosonde data bias problems. Regression retrievals, although in error, because they faithfully reproduce a proportion of the radiosonde bias errors, will appear to be accurate when compared with the collocated (biased) radiosondes!

b. TMS retrieval results

Three kinds of TMS retrieval results can be constructed from the collocation data; in the first, retrievals are computed from *rte* radiances, suitably perturbed by Gaussian noise, commensurate with the expected measurement noise values of the radiometers. The second kind utilizes satellite radiances, adjusted by the TSF dependent δ_{TSF} correction method, while the third uses satellite radiances adjusted by the D_{TSF} correction scheme described by (3).

Results from the first approach define the inherent accuracy of the retrieval algorithm and remove any effects due to radiosonde biases, because in this case the radiosonde does define the "truth". With the second approach, errors in the ground truth are propagated into the retrieval error statistics in a number of ways. Contributions arise from incorrect estimates of the δ *rte* correction vectors, degradation of the information in the satellite observations through the necessity of using large "measurement" noise estimates relative to those fundamental to the measurement process, and comparison of the retrieved profiles with a biased estimate of the true state of the atmosphere. The third method produces retrievals which are not so severely affected by collocation sample radiosonde biases, in so far as they affect the problem of estimating *rte* and collocation radiosonde bias errors. Also the "apparent" between channel error covariances are reduced.

Results for the first method, using the TMS estimator and two different sets of TOVS radiance observations are given in fig 6. In both examples the a priori information used to constrain the solution is determined from Bayesian discrimination equations defined on channels 24, 2, 3, 4, 5, 23 and 12. In fig 6a, the retrieval estimator made use of only channels 8, 24, 23, 22, 2, 4, 5 and 6, and the solution was iterated twice. Fig 6b demonstrates the results for a more complex four pass estimator where in the first pass a log(*pcw*) representation (Uddstrom, 1988) of the simultaneous retrieval equations is applied for channels 18, 12, 11 and 10, then a simple non log(*pcw*) representation of the equations is employed for channels 18, 12, 10, 1, 2, 24, 3, 23, 4, 5, 6, 7, 8, 16, 15, 14 and 13, and this latter selection is iterated twice more. Clearly, the rms error statistics for the second TMS estimator are superior to those of the first, as would be expected. In particular the water vapour retrievals are substantially improved as are the temperature errors in the lower troposphere. Of course, because there is no error due to skin effect problems in these results, the retrieval of the surface information is good.

When δ_{TSF} corrected satellite observed radiances are employed in the TMS algorithm, and the same channels as the estimator of fig 6a are used, the accuracy of the results, as defined by the differences between the collocated radiosondes and the TMS retrievals is substantially reduced (fig 7). A component of this error is however fictitious, since the TMS satellite retrievals do not produce profiles which include the radiosonde bias effects. Even allowing for this fact, the temperature and thickness errors of the NESDIS and TMS algorithms are not too dissimilar, except near the surface where the NESDIS results are superior (see fig 5). The NESDIS water vapour retrievals are better than the TMS results. The variance in the retrieved samples (not shown) is significantly less than that in the "true" sample.

Additional experiments were attempted in order to improve upon the TMS estimator results given in fig 7. For example, more and different selections of channels were used, and the δ_{TSF} corrections were modified so as to combine day and night estimates for channels where solar radiation was not expected to have an important effect (i.e., excluding channels 7, 8, 13, 17, 18

and 19). However, no major improvements in the error statistics occurred. These experiments only underlined the fact that the physical estimator is, not unexpectedly, sensitive to the choice of the δ_{TSF} vectors.

However, using radiances corrected by the D method of (3), and a TMS estimator identical to that used to derive the fig 7 statistics, results in important improvements in the rms error statistics at nearly all levels. The improvement in the lower troposphere is especially pronounced, as demonstrated in fig 8.

In fig 9 some example retrievals are displayed for NESDIS, TMS from satellite observations (both δ_{TSF} and D_{TSF} corrected) and TMS from noise contaminated rte radiance temperatures. The most noticeable difference between the TMS (δ_{TSF} corrected) and NESDIS retrievals, is that the regression algorithm appears to make better use of the near surface channels. However the situation is reversed for the TMS retrievals computed from the D method corrected satellite observation, even though the TMS retrieval estimators were identical in both cases.

The retrieval at 35°S 58°W, computed from one of the outermost (NESDIS) retrieval boxes, is included to demonstrate the kind of errors that result when the observed radiances simply do not agree with those computed from the rte model. For this example, there is a simple linear shift between the retrieved temperature profiles and the radiosonde measurements. Whether this is due to bias errors in the radiosonde data, or perhaps the nadir correction algorithm, is unclear.

In other regards, the TMS and NESDIS retrievals, for these examples, show striking similarities except perhaps for the retrieval at 66°S 14°E, where the TMS algorithm has more faithfully reproduced the structure of the tropopause. The TMS retrievals computed from the synthetic radiances do however demonstrate that, in principal, there is sufficient information in TOVS radiances to compute accurate temperature and water vapour profiles.

5. DISCUSSION AND CONCLUSIONS

If errors due to undetected or incorrectly identified clouds, nadir correction methods and collocation data interpolation are excluded, two important sources of error remain for physical retrieval estimator algorithms. The first of these arises because rte models are unable to accurately and precisely simulate infrared satellite observations. Consequently, in order to bring the modelled and measured observations into agreement—a prerequisite of all physical retrieval algorithms—some adjustment scheme must be developed so that effective rte radiance temperatures can be estimated from the satellite observed measurements (or vice versa). As a further complication to this problem, it is apparent that these corrections are air-mass dependent. The second, and associated problem relates to what is regarded as the “truth” when determining corrections for rte models. Empirically derived δ radiance temperature adjustments will always yield bad estimates of rte model errors if the collocation sample from which the correction vectors are derived is not homogeneous. Any inhomogeneities in the collocation sample will additionally decrease the precision of the δ estimates, forcing the retrieval algorithm to reduce the weight attached to the observed radiances and so degrade the value of the direct measurements. As a consequence, the variance in retrieved profile samples will be reduced because the retrieval solutions are always forced to lie close to the first guess profiles. Therefore, air-mass dependent δ correction terms (and their variances) can, properly, only be calculated from samples of collocated radiosondes which have systematically been reduced to some effective “common radiosonde”, because the channel measurement and modelling errors are assumed to be independent. Unfortunately there is no space in the present WMO Upper Temps code for radiosonde make, model and processing information, and the WMO catalogue of “Radiosondes

in use by members" (see e.g. WMO, 1982) is invariably out of date and sometimes inaccurate. For these reasons it may be impossible to make consistency corrections to the standard NESDIS collocation archive.

Accordingly, regression retrieval algorithms, which minimise the effects of errors introduced by radiosonde dependent measurement biases, will have an apparent advantage over physical retrieval schemes utilizing δ type rte model corrections, when their respective outputs are compared with collocated radiosondes.

The D correction method, in effect a generalisation of the δ scheme, is, in principle a better estimator of rte "errors", since the corrections are derived from all the available information. However, this regression estimator may also produce undesirable side effects if the observed data do not satisfy the assumptions implied in the development of the D equations.

To further underline the importance of the radiosonde bias problem, it has been possible to demonstrate that these errors have a measurable effect on the performance of a primitive equation NWP model. When all radiosonde data used in the New Zealand Meteorological Service optimal interpolation analysis scheme were adjusted to make them "look like Väisälä RS80" radiosondes (by the method given in Uddstrom, 1988b), the NWP 6 hour verification statistics (against observed data) showed improvements at most levels. The bias error statistics for heights improved by as much as 50% at 200 hPa and above. Even at the surface, the model showed a 0.2 hPa performance improvement (Purnell and Revell (1988) pers. comm.).

Therefore, perhaps proper assessment of the relative merits of both physical and regression retrieval algorithms must await the arrival of a quality controlled set of collocation data, where the coincident radiosonde data have been reduced to some equivalent "standard". It is to be hoped that the Base line Upper Air Network will provide a data set having these properties.

REFERENCES

- Nash J. and F.J. Schmidlin, 1987: WMO International radiosonde comparison (U.K. 1984, U.S.A. 1985) Final Report, *Instruments and Observing Methods Rep No. 30*, 103pp
- Richner, H. and P.D. Phillips, 1982: The radiosonde intercomparison SONDEX Spring 1981, Payerne. *Pure and Appl. Geophys.*120, p852-1198
- Uddstrom M.J., 1988: Retrieval of atmospheric profiles from satellite radiance data by typical shape function maximum a posteriori simultaneous retrieval estimators. In press *J. Appl. Meteor.*
- Uddstrom M.J., 1988b: A comparison of Philips RS4 and Väisälä RS80 Radiosonde Data. Submitted as a Note to *J. Atmos. Oceanic Tech.*
- WMO, 1982: WMO catalogue of radiosondes in use by members. *Instruments and Observing Methods, Rept No 11, Upper-Air data compatibility* 19pp

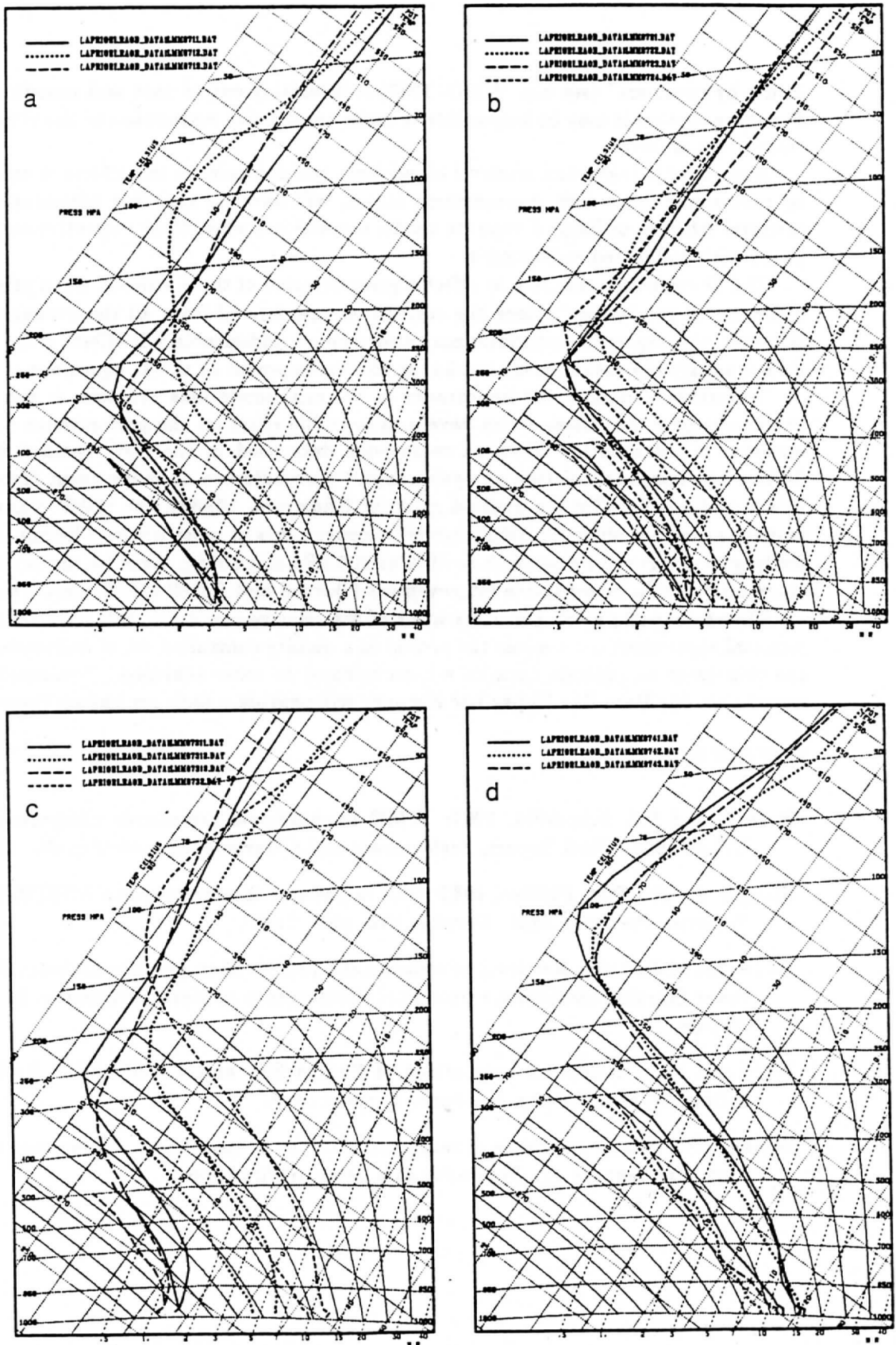


Fig.1. Mean profiles for winter, TMS duplicated, TSF samples of collocation radiosonde data. Panel (a) is classes 11, 12 and 13; panel (b) is classes 21, 22, 23 and 24; panel (c) is classes 311, 312, 313 and 32; and panel (d) is classes 41, 42 and 43.

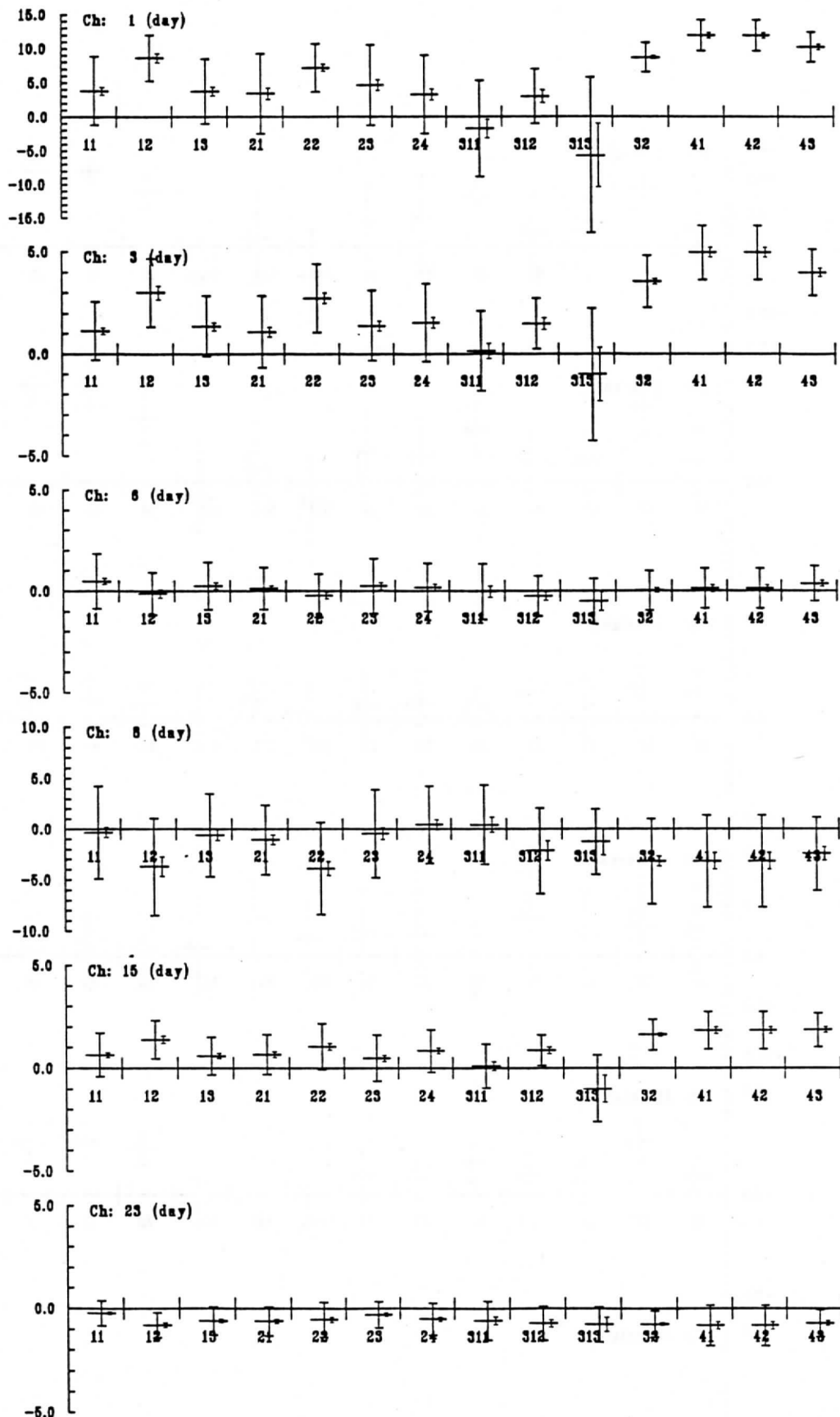
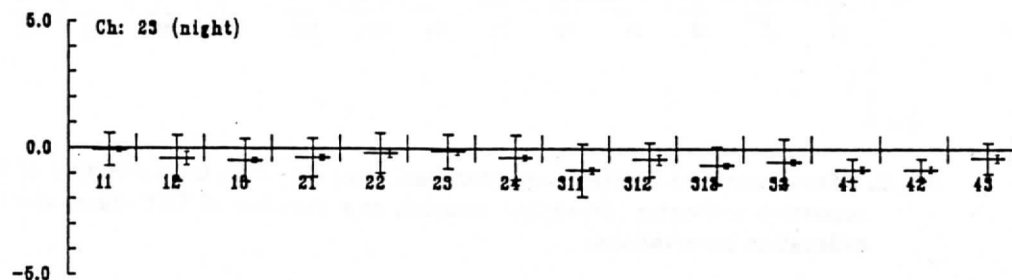
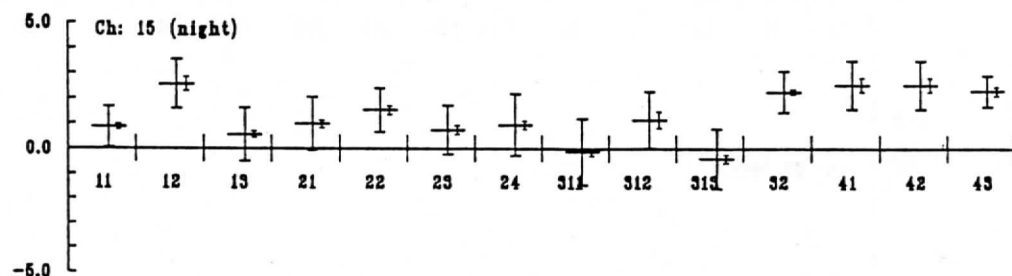
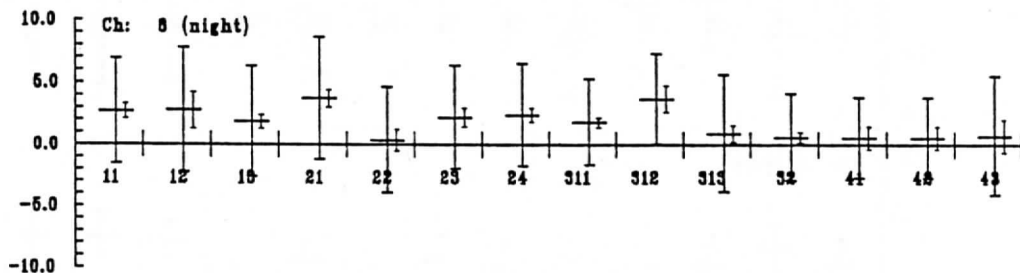
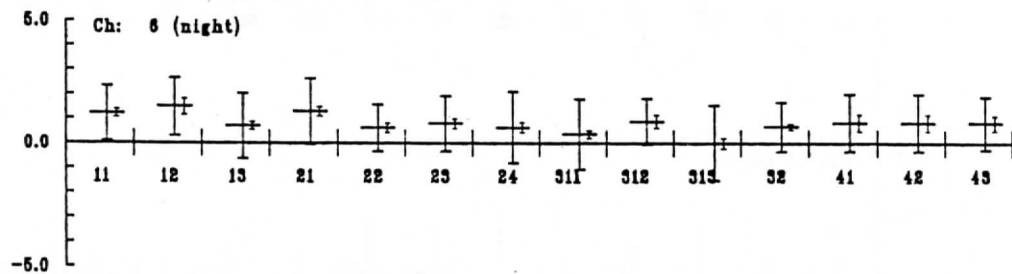
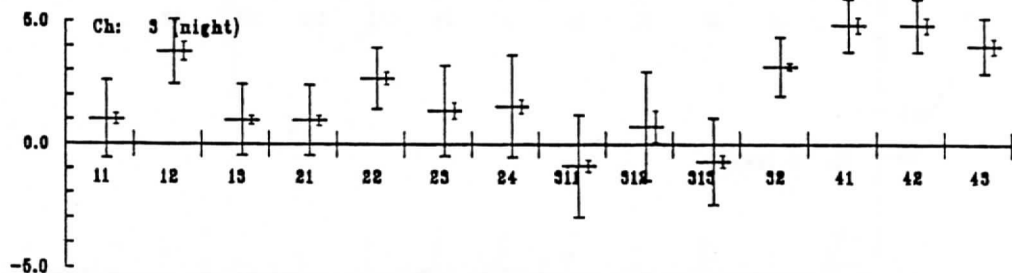
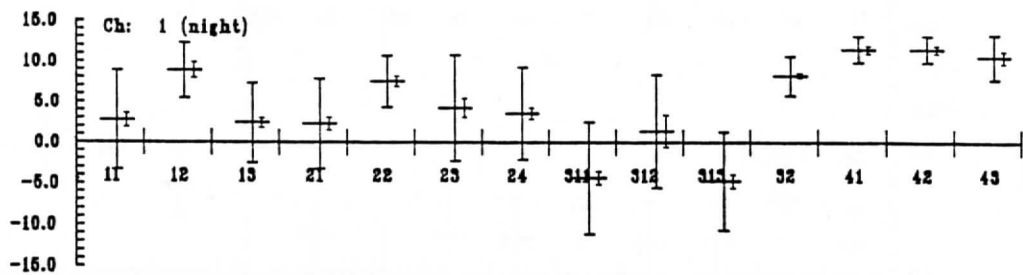


Fig.2. Mean, standard deviation and standard error statistics for a selection of NOAA-7 rte 6 channel correction estimates (dependent sample), as a function of TSF class, and the time of day of the collocation observations.



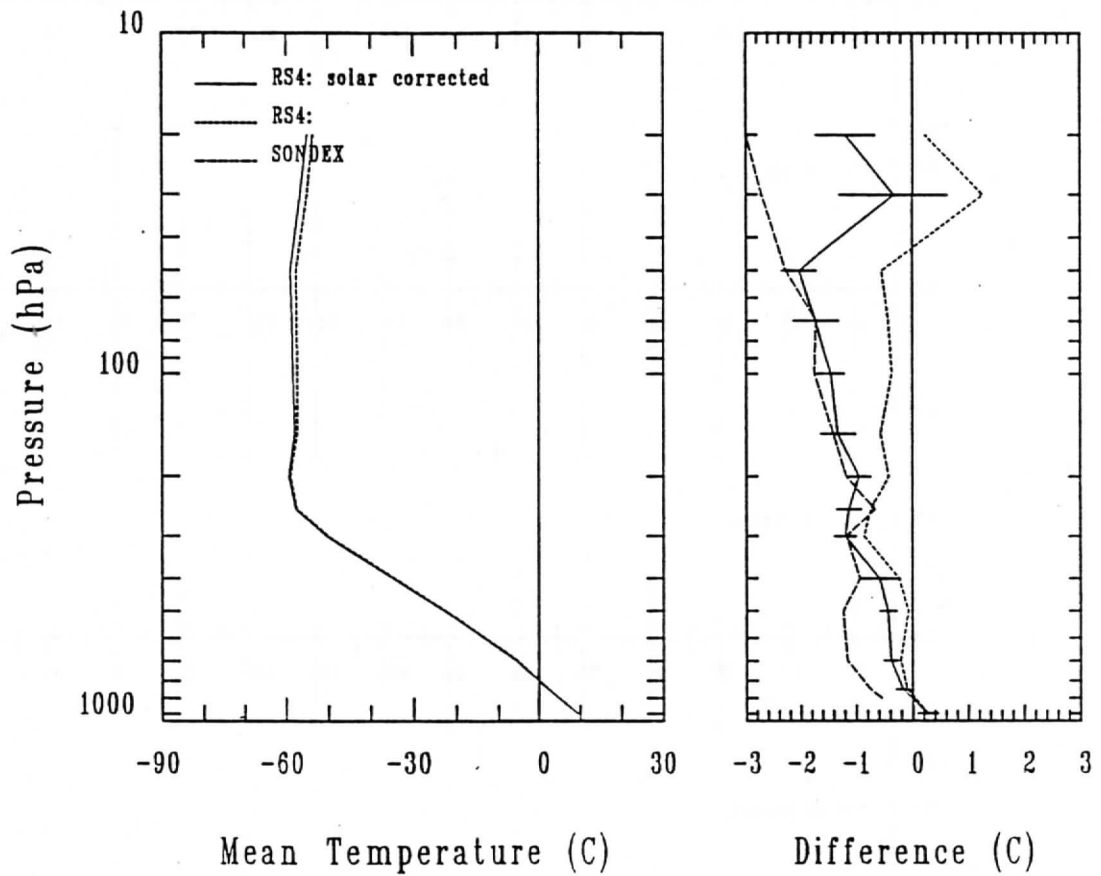
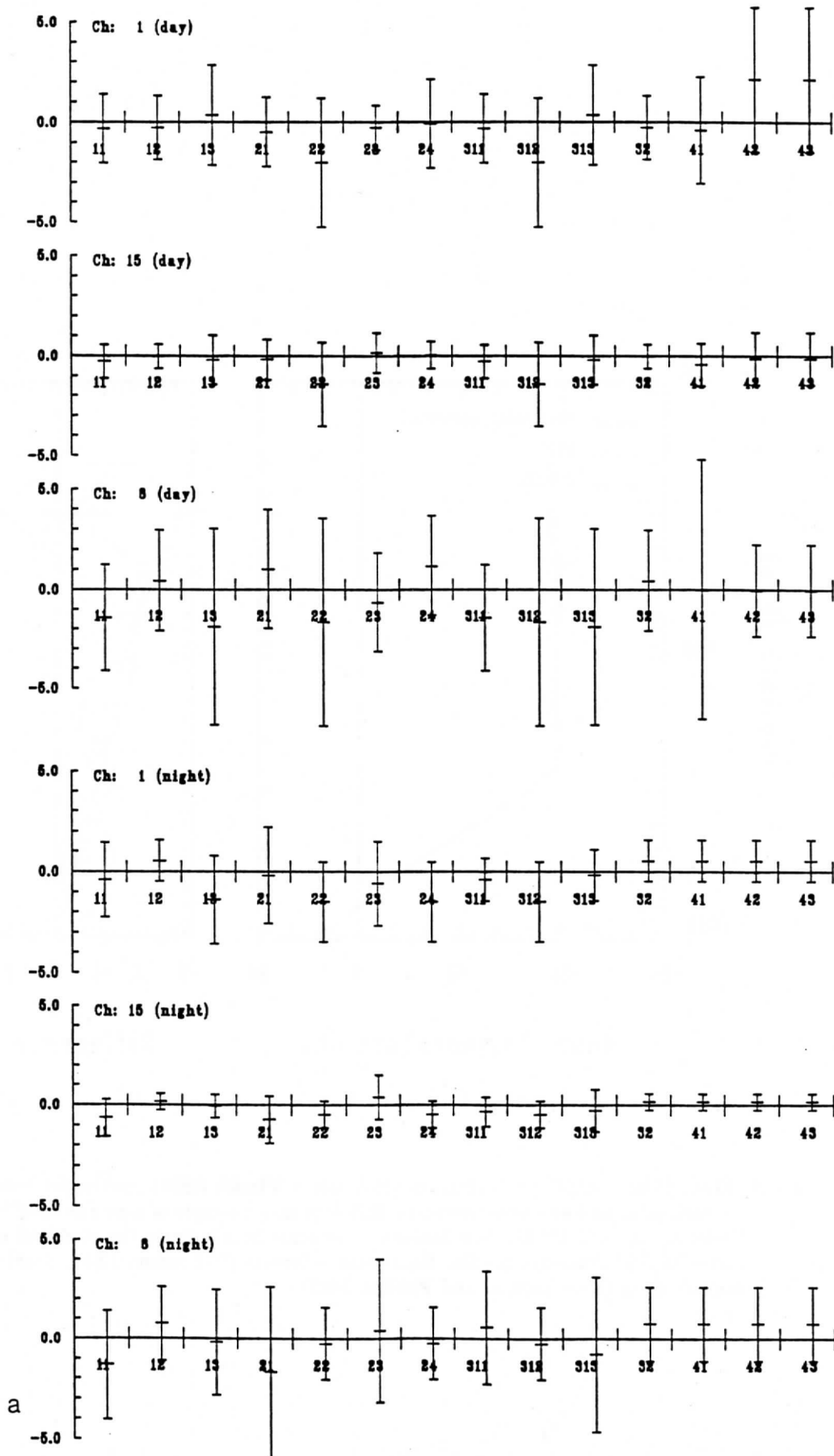
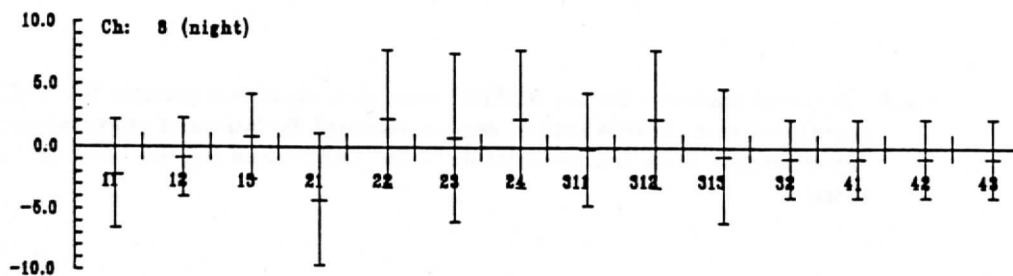
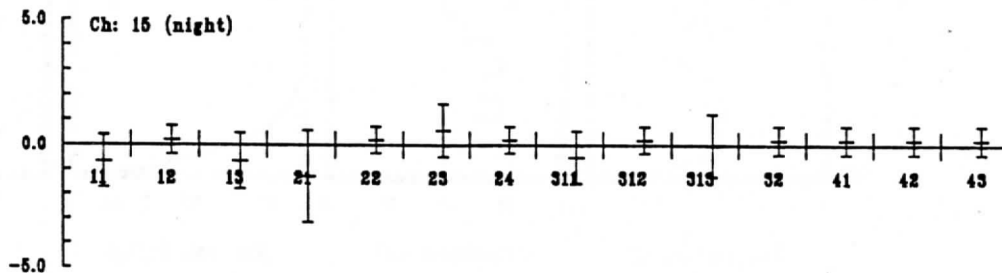
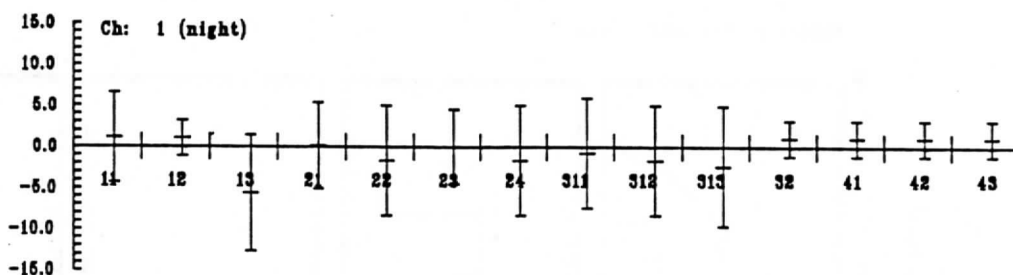
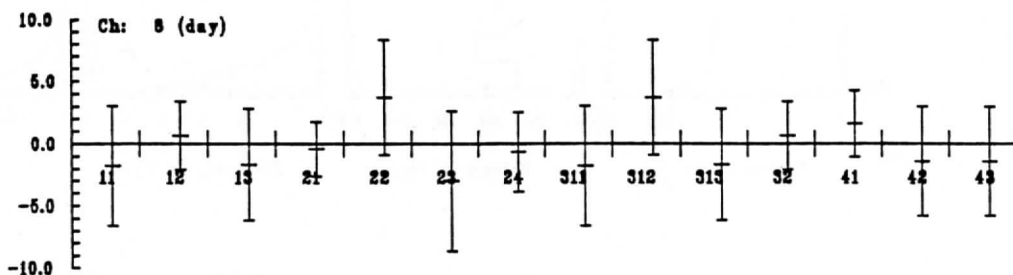
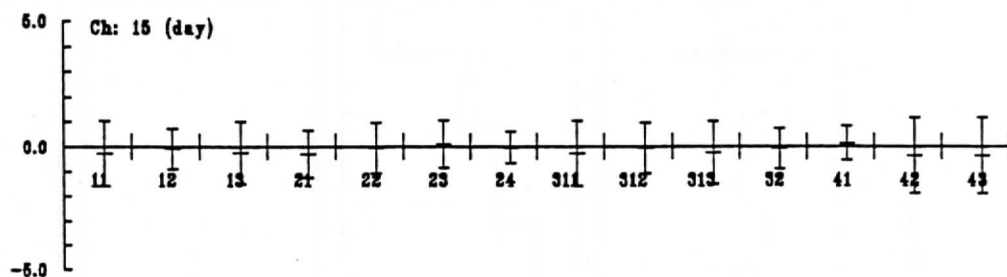
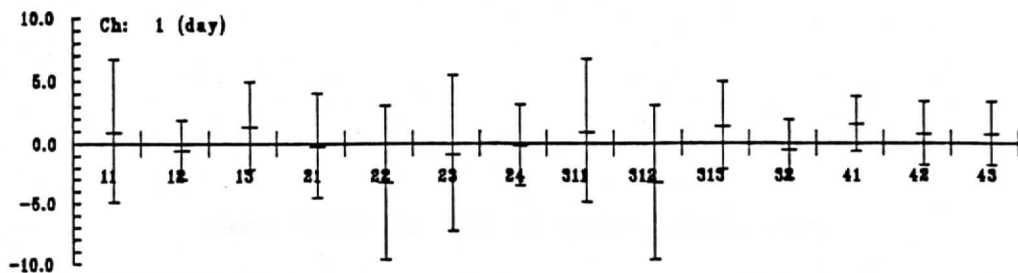


Fig.3. Mean (Philips RS4) and difference (RS4 minus Väisälä RS80) profiles for temperature profiles, for both solar and non solar corrected RS4 data and a sample of eighteen 0000 UTC soundings at Christchurch (44S 173E), New Zealand. The error bars indicate the standard errors of the solar corrected RS4 difference profile. Equivalent difference (VIZ minus RS80) profiles from SONDEX are also given (after Richner and Phillips, 1982).



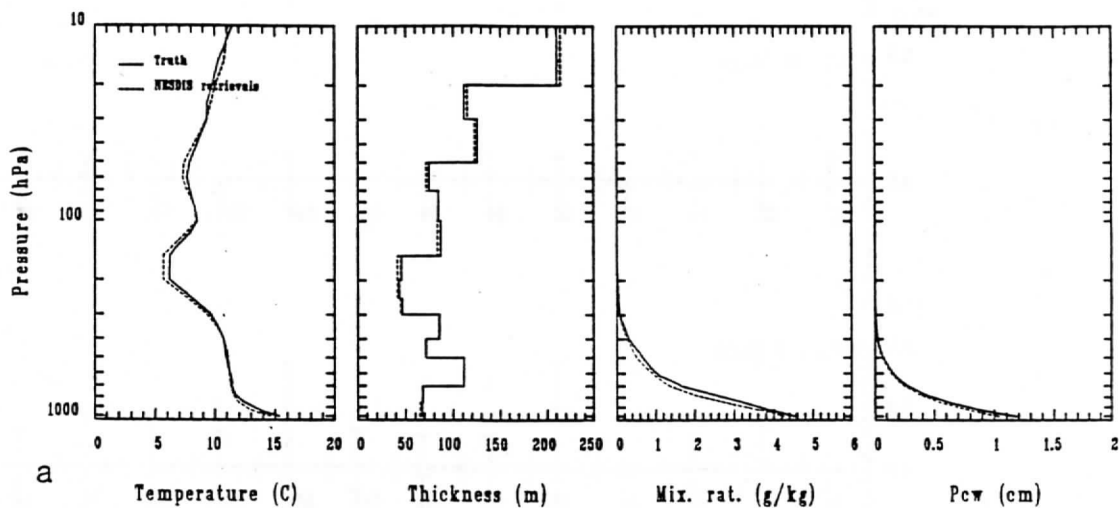
a

Fig.4. (a) Rms and bias errors for NOAA-7 satellite observations corrected to equivalent rte model values by the D method (independent sample), as a function of TSF class, and time of day of the collocation observations. (b) Rms and bias errors for the same data as (a), but corrected to equivalent rte model values by the δ_{TSF} method.



b

Standard deviation statistics for "TRUE" and "NESDIS" samples



NESDIS profile RMS errors

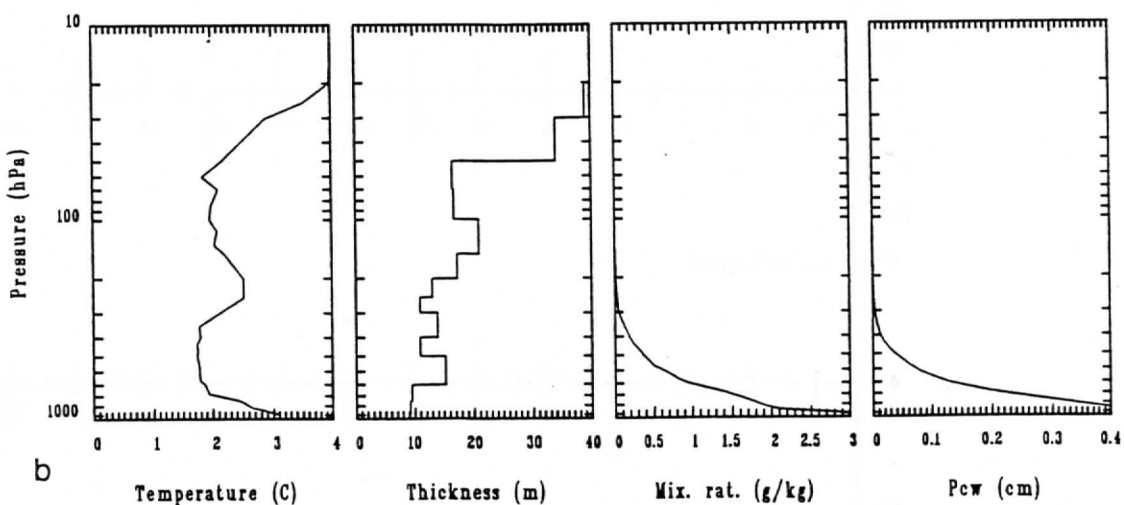


Fig.5. Retrieval statistics for the NESDIS regression algorithm (sample size 192). (a) Retrieved (solid curve) and true (broken curve), sample standard deviations of temperature, standard layer thicknesses, mixing ratio and precipitable water profiles, and (b) rms error profiles for the same quantities.

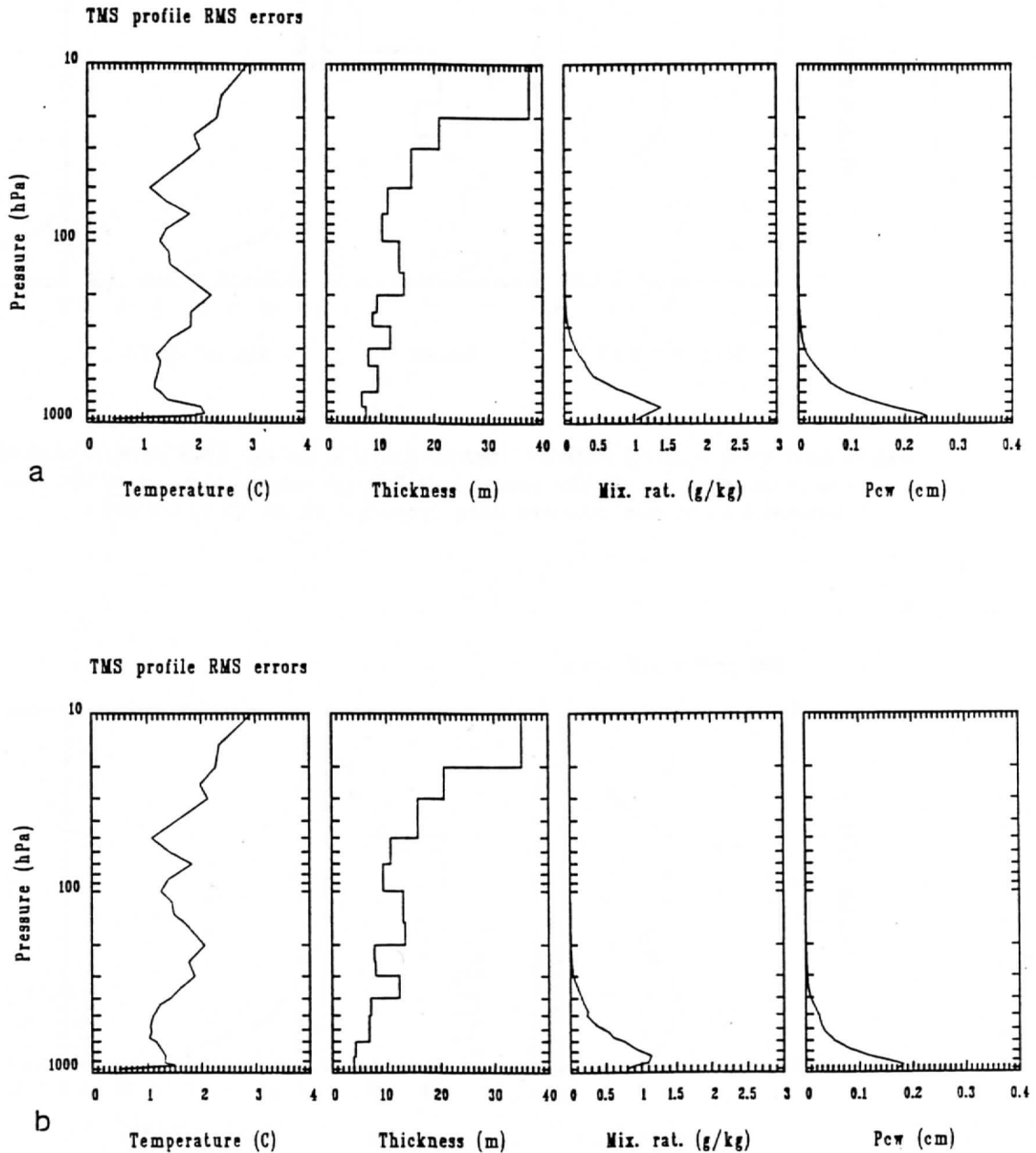


Fig.6. Rms error retrieval statistics (sample size 192) for the TMS physical algorithm applied to noise perturbed rte radiance temperatures, for temperature, thickness, mixing ratio and precipitable water for (a) a two pass estimator using channels 8, 24, 23, 22, 2, 4, 5 and 6, and (b) a 4 pass estimator (see text).

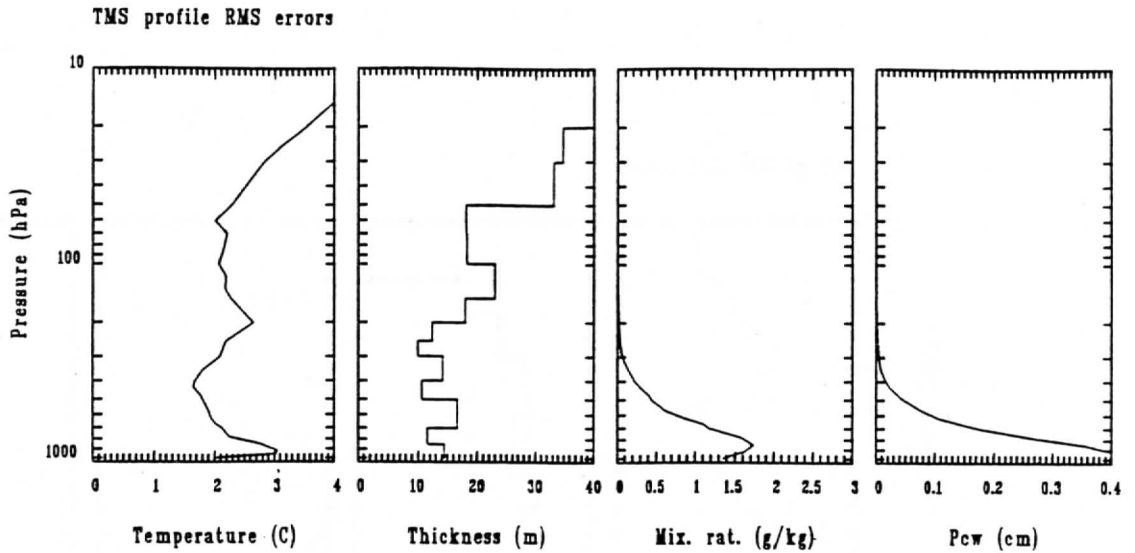


Fig.7. Rms error retrieval statistics (sample size 192) for the TMS physical algorithm applied to δ corrected collocated satellite observations, for temperature, thickness, mixing ratio and precipitable water and a two pass estimator using channels 8, 24, 23, 22, 2, 4, 5 and 6.

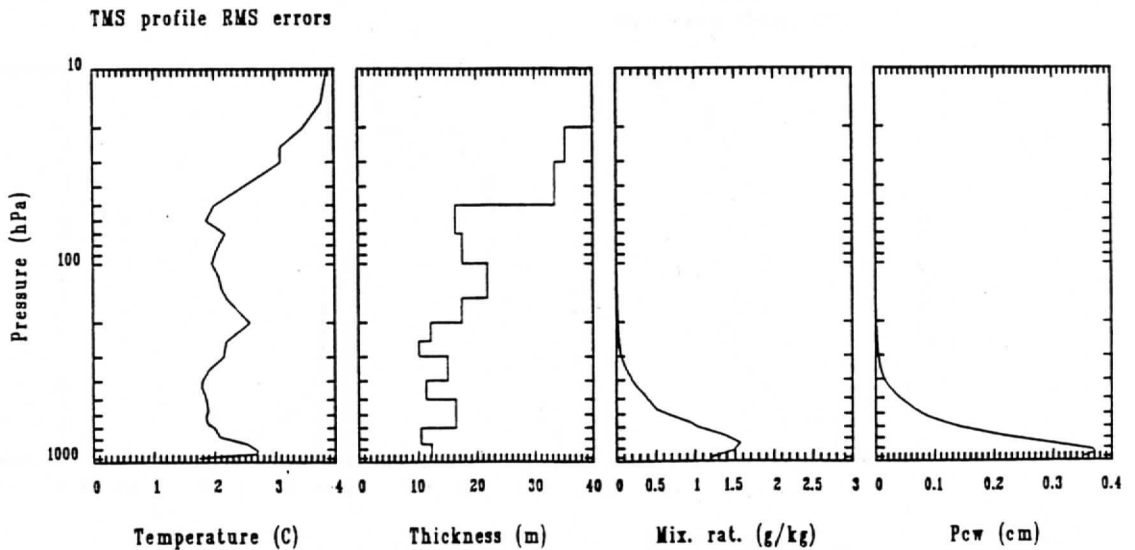
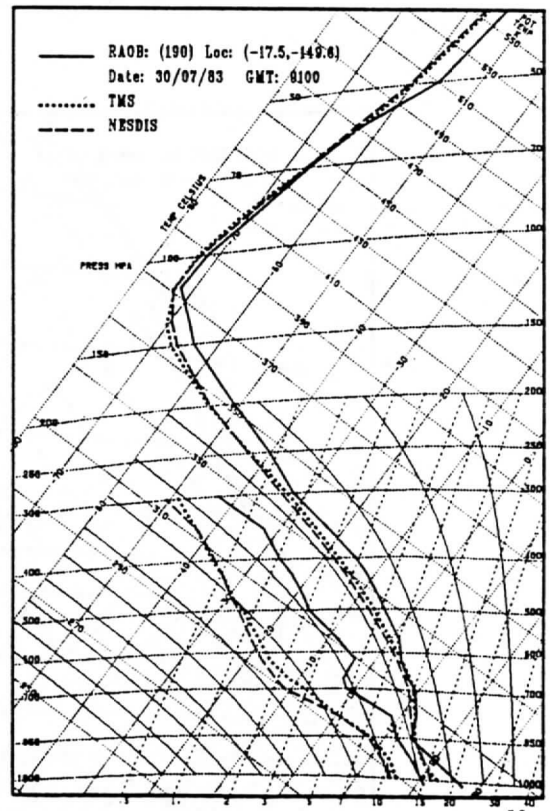
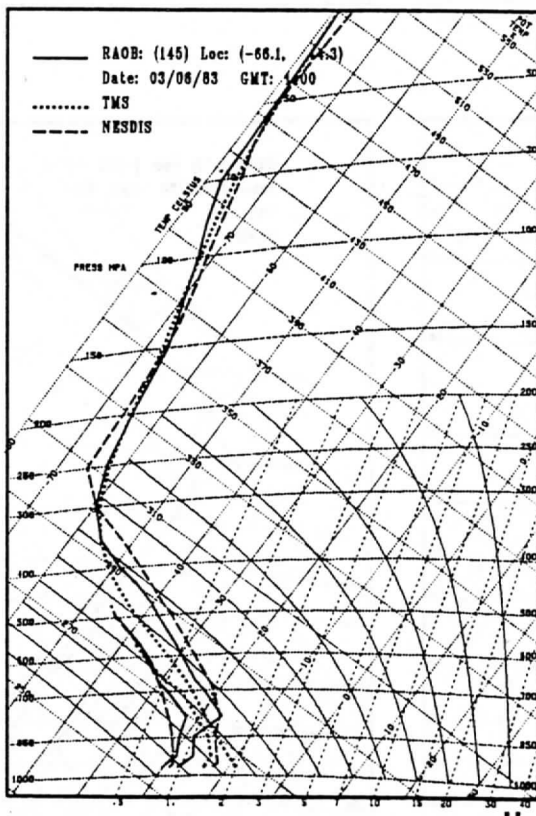
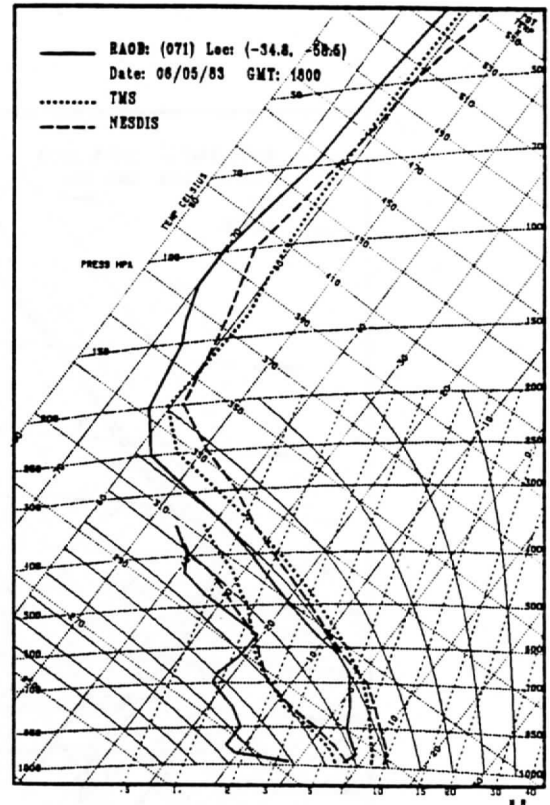
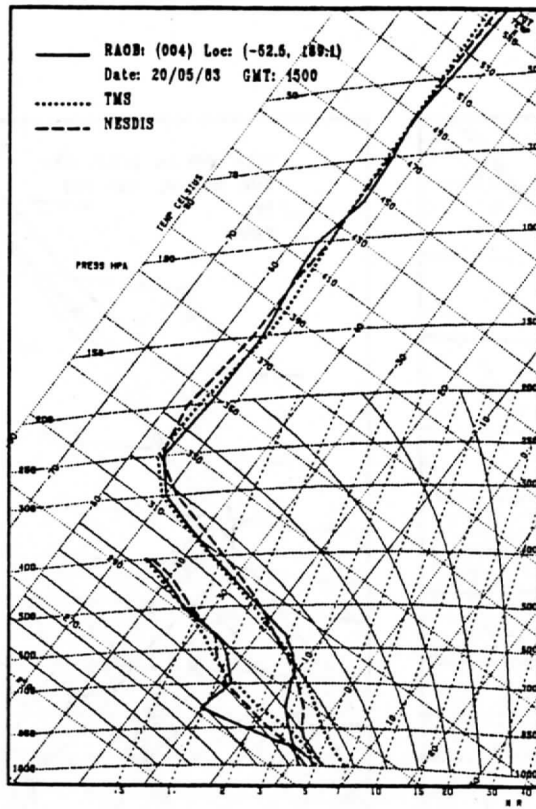
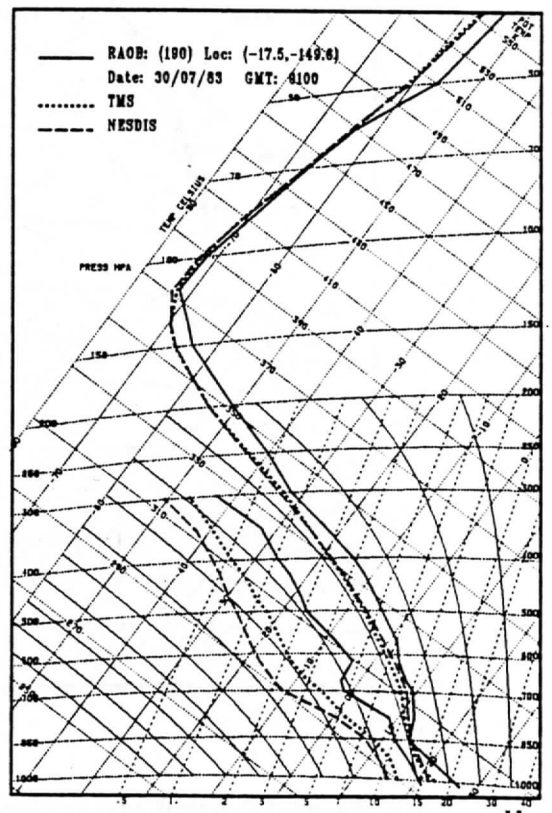
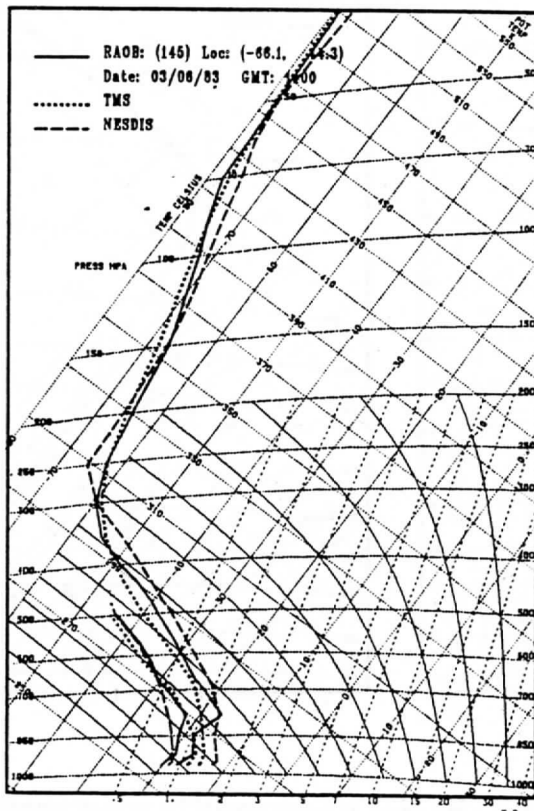
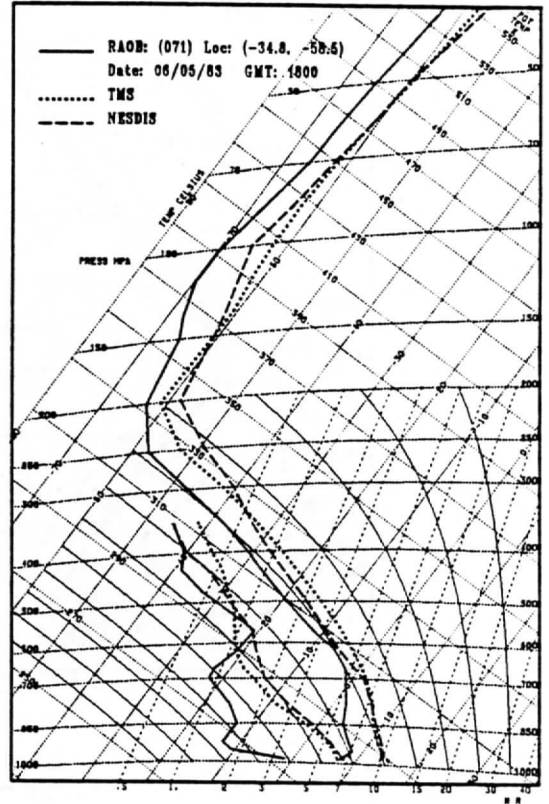
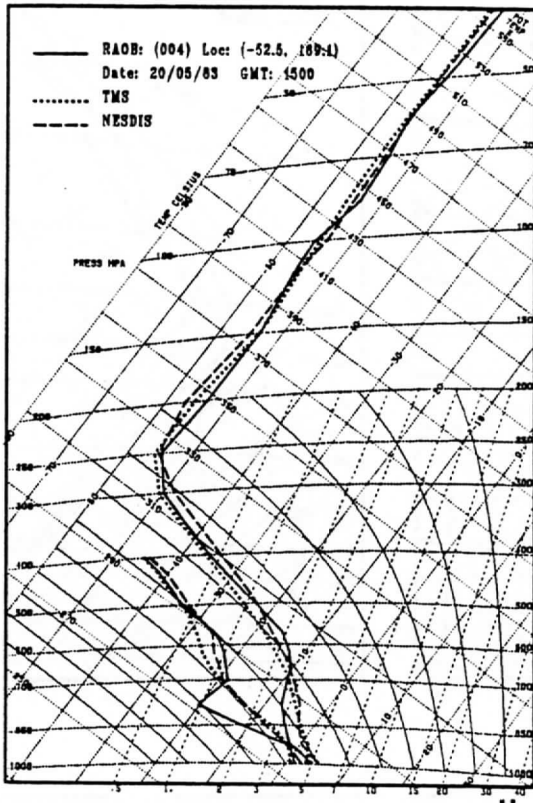


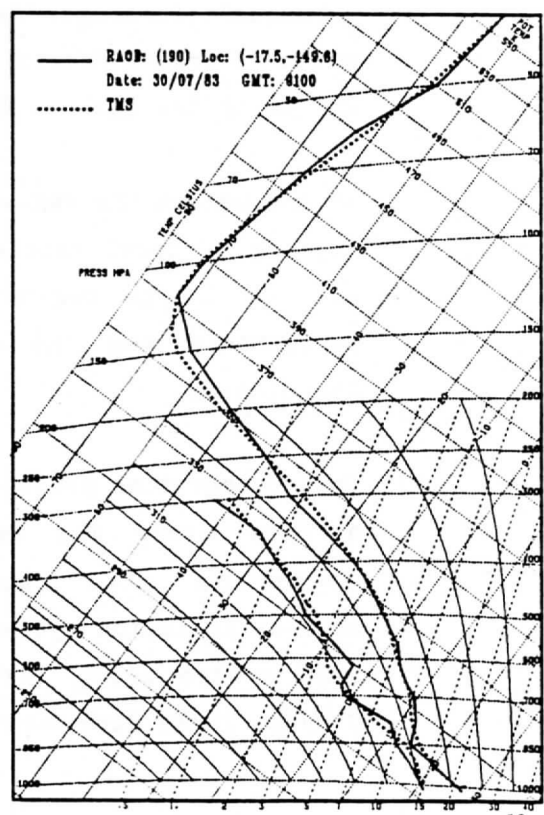
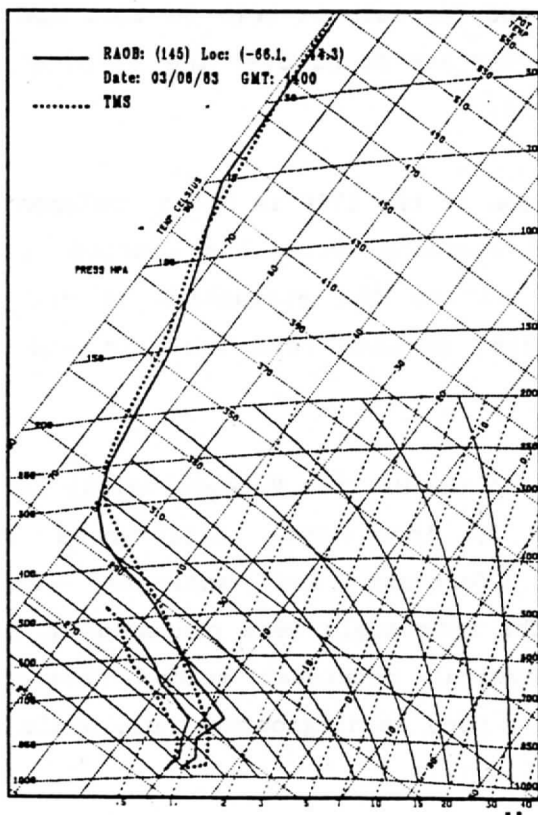
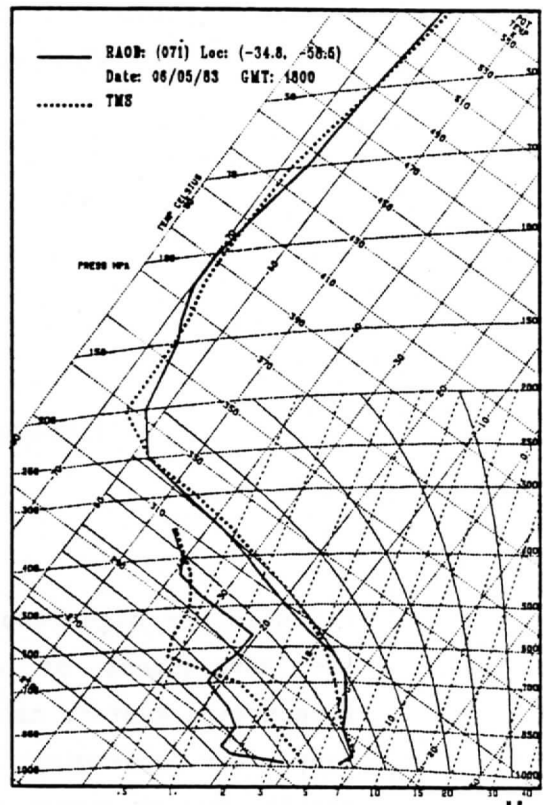
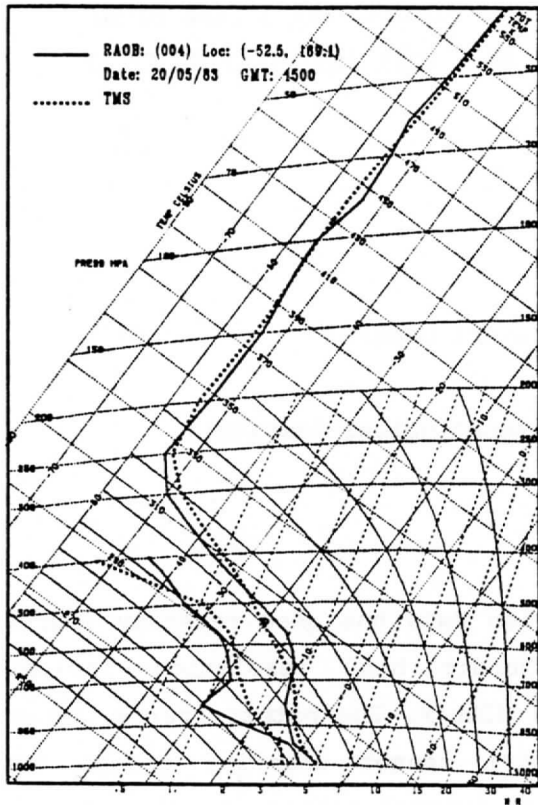
Fig.8. Rms error retrieval statistics (sample size 192) for the TMS physical algorithm applied to D corrected collocated satellite observations, for temperature, thickness, mixing ratio and precipitable water and a two pass estimator using channels 8, 24, 23, 22, 2, 4, 5 and 6.



a

Fig.9. Example retrievals for, (a) NESDIS and TMS δ_{TSF} corrected satellite measurements, (b) NESDIS and TMS D_{TSF} corrected satellite measurements, and (c) TMS from noise perturbed rte radiance temperatures.





The Technical Proceedings of
The Fourth International TOVS Study Conference

Igls, Austria

March 16-22, 1988

Edited by

W. P. Menzel

Cooperative Institute for Meteorological Satellite Studies
Space Science and Engineering Center
University of Wisconsin
1225 West Dayton Street
Madison, Wisconsin 53706
(608) 262-0544

October 1988