THE 3-I RETRIEVAL METHOD
RECENT LOCAL AND GLOBAL APPLICATIONS

A. Chedin
Laboratoire de Météorologie Dynamique du CNRS
Ecole Polytechnique, 91128 Palaiseau Cedex, France

SUMMARY

The major role played by a priori information in the retrieval problem of satellite sounding of the atmosphere can be understood by considering the fact that radiance observed by the sensors integrates the atmospheric thermal structure over relatively thick layers. Such an integration results in the well known problem of the non uniqueness of the solution. By specifying an initial guess solution as close as possible to the final correct solution, any a priori information may help to overcome this difficulty. The "3I" (Improved Initialization Inversion) method, developed at Laboratoire de Météorologie Dynamique, makes systematic use of available a priori information for retrieving the best possible initial guess through a pattern recognition type approach. The 3I algorithm has been designed with the purpose of retrieving geophysical parameters from space radiometric measurements and oriented towards the processing of the TIROS-N series observations made by the TOVS (TIROS-N Operational Vertical Sounding) instruments: HIRS-2, a 20-channel infrared radiometer, and MSU, a 4-channel microwave radiometer. The first purpose of the present paper is to present an overview of the 3I method: initialization of the inversion problem including cloud detection and cloud clearing, atmospheric temperatures retrieval, cloud parameters retrieval, atmospheric water vapour retrieval and surface temperatures retrieval. The second purpose of this paper is to present and discuss the results of recent applications of the algorithm to NOAA data.

1. INTRODUCTION

The retrievals of atmospheric temperatures from the satellite soundings are still confronted with the problem of the non-uniqueness of the solution. The problem can be partially rectified by specifying an initial guess solution which is as close as possible to the correct one. Knowledge of a priori information on other atmospheric conditions observed by the satellite at the time of soundings may therefore be beneficial. Among this kind of information are:

- latitude, longitude, observation and solar viewing angles for each field of view;

- surface properties: ground elevation, percentage of water, albedo, emissivity, temperature (sea surface), presence of snow or ice;

- operational analyses or forecasts, available at the time of the satellite pass;

- some features of the atmospheric structure like the type of the air mass (polar, middle or tropical) or the temperature of the lower stratosphere, which may be obtained, for example, from regressions over well sampled pre-archived data sets.

It is worth pointing out that a priori information belongs to two well separated categories. In the first one are those parameters which actually enter the initial solution through the forward model and the computation of the radiative transfer equation: viewing angle, surface properties (pressure, emissivity, temperature over sea) in particular. In the second one are those which orientate the proximity recognition procedure: air mass type, tropopause information, forecasts, inversion in the boundary layer, etc.

The number and the variety of the parameters mentioned above lead to the idea of establishing a set of characteristic measurements including both the set of observed radiances and a set of local properties like the one we have just seen. Gathering all such possible sets, which will be referred to as "feature" or measurement vectors, following the pattern recognition terminology, gives rise to the feature space which may be considered as an ensemble of templates or prototypes to which the input pattern (the observation) may be compared using a similarity criterion. The "3I" - Improved Initialization Inversion - method developed at LMD in the past few years for application to the satellites of the TIROS-N series has been designed following a pattern recognition type approach: the feature space is a large precomputed data set containing feature vectors made of synthetic radiances and associated local properties for a large number of atmospheric conditions: the so-called "TIGR" (TOVS Initial Guess Retrieval) data set. The similarity criterion is a minimum distance classifier. This procedure produces an initial guess solution of high quality which may start a physical type inversion method requiring only one iteration (one matrix inversion) to get an accurate final solution. To the best of our knowledge, this method, developed in 1983 (Chedin and Scott, 1983) has been the first non iterative or direct physical method.

1.1 The feature space: TIGR data set

The "3I" procedure is based upon the creation, once and for all, of a data set describing physical, radiative and local properties of a large number of atmospheric situations that may plausibly be observed from a satellite platform (here the TIROS-N series). The patterns made of combinations of all these parameters are usually quite complex and the number of features very large, which makes the idea of describing such a complexity in terms of a hierarchy of sub-classes very attractive. Here this hierarchy is made of the air mass type, the viewing angle value, the surface pressure value (which contains the surface elevation), the type of the surface (related to the emissivity value) and finally, the atmospheric conditions themselves. Three types of air masses are presently considered: polar (north and south), mid (north and south) and tropical. This is the first approach to a climatic classification of atmospheric conditions which could be implemented in the future. Ten viewing angles are considered from nadir to a maximum value of approximately 60° which correctly sample the range of viewing angles associated with the scanning instruments on board the TIROS-N series, HIRS-2 and MSU. Nineteen surface pressure values are considered. Two emissivity values representative of over sea and over land observations are also considered. Finally, for all the possible combinations of viewing angles, surface pressures and emissivities, a total of 1207 atmospheric conditions, carefully screened from a much larger set using a sophisticated statistical algorithm, are considered: 525 for polar latitudes, 545 for mid-latitudes and 137 for tropical latitudes.

The radiative properties, actually the transmittance profiles and the brightness temperatures for all the channels of the instruments, associated with each situation archived (including the observing conditions) are computed with a fast line-by-line transmittance and radiance model (Scott, Chédin, 1981). This code is regularly validated against satellite data, here using TOVS radiosonde matchups through the so-called "γ-δ" procedure which consists in replacing the original synthetic radiance $I(\tau)$, for a given channel and the transmittance profile τ, by $I(\tau^{\gamma})+\delta$. Being close to 1, γ takes care of the limitations of the forward radiative transfer model whereas δ removes remaining biases, in particular instrumental. The accuracy of the forward model being very satisfactory, on the one hand, and because $d\tau$ and not τ itself actually enters the inversion process on the other, it has recently been decided to fix the γ's to unity and to remove all biases through the δ's.

This procedure is used not only for controlling the evolutions of the measurements in time but also to adjust the TIGR data set to any satellite of the series.

2. SURVEY OF THE 3I PROCEDURE

The spatial resolution of the 3I algorithm is a compromise between the spatial resolutions of the two major sounders on board the NOAAs: HIRS-2 and MSU. Retrievals are made for boxes of 2 to 4 (according to the viewing angle) HIRS-2 spots along the scan line by three HIRS-2 spots along the sub-orbital track. Such boxes approximately represent a surface of 100 x 100 km^2. However, in order to increase the density and check the coherence of the results, an oversampling may eventually be introduced along the sub-orbital track and one retrieval produced every 100 x 30 km^2, two consecutive boxes sharing two scan lines. See, for example, Chedin et al. (1985) and Chedin and Scott, (1985) and Fig. 1. A review of the essential aspects of the 3I method is presented in the next sections.

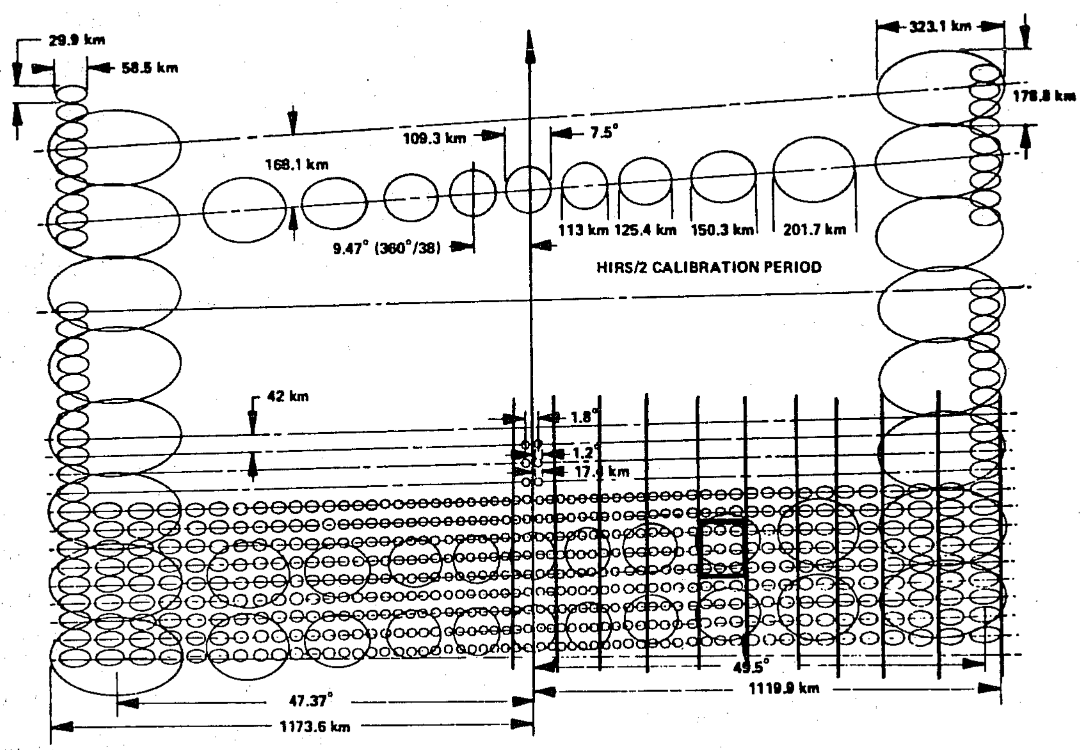

Fig. 1 HIRS-2 and MSU scan patterns projected on the Earth's surface. Retrievals consisting of three scan lines and four (small viewing angles) to two (large viewing angles) HIRS spots are made for each box. There may be an oversampling along the y-axis (suborbital track): two consecutive boxes sharing two scan lines. The ten divisions, from nadir to limb, correspond to the ten viewing angles sampled.

2.1 Air mass classification

Knowledge of the air mass type should be very useful because the average temperature profile and associated variance are very different for the three main classes of air mass: polar, middle and tropical latitude types. Providing full description of the underlying atmospheric column, observations from the vertical sounders on board the TIROS-N series may be used to determine the type of the air mass observed. The method which has been developed (Moine et al., 1987) is based upon a classification scheme known as the 'dynamic cluster algorithm'.

For application to satellite observations, the input to the classification scheme are the observed radiances (or equivalent brightness temperatures). The great differences in the observing conditions (viewing angle, surface characteristics) must be taken into account prior to applying the classification scheme: the objects are divided into nine groups resulting from the consideration of three viewing angles (small, medium and large) and of three terrain elevations for observations over land (low, medium, high).

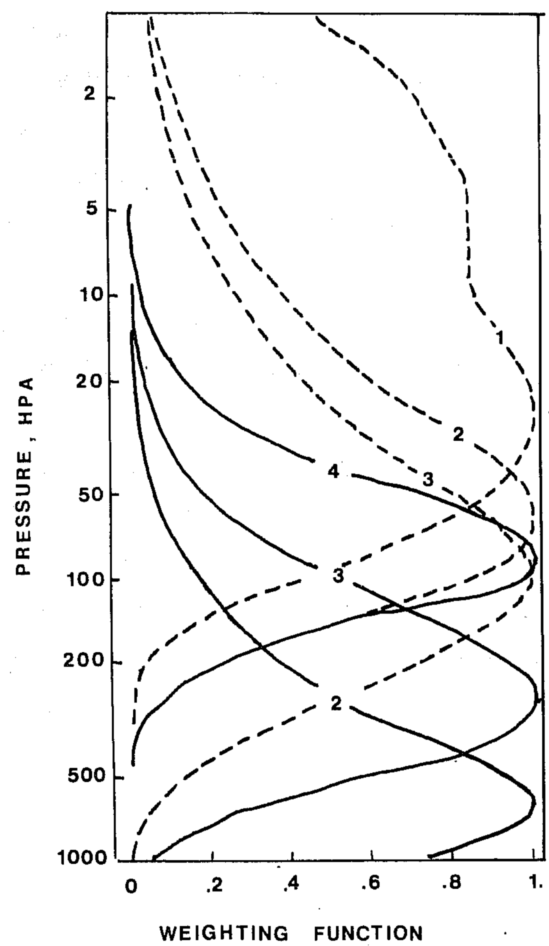

Fig. 2 Weighting functions (normalized) of the channels used in the classification of satellite observations. —— MSU channels; ----- HIRS-2 channels.

Each group of objects is then classified separately. Another difficulty comes from the fact that observed radiances are often contaminated by the presence of clouds in the field of view. Fortunately, some of the TOVS channels are almost insensitive to clouds while giving a good description of the whole atmospheric column. These descriptors are channels 1, 2 and 3 of HIRS-2, 'looking' in the atmosphere higher than the region where most clouds are, and channels 2, 3 and 4 of MSU, benefiting from the low sensitivity to clouds of microwave measurements (see Figure 2).

2.2 Cloud detection

The presence of clouds over the observed area influences the retrieval process. Consequently, the role of a cloud detection algorithm appears as essential since from its performance will depend the quality of the retrieved results. Applied to each HIRS spot in a box, the algorithm consists of a threshold test on the albedo, several coherence tests between various channels of the HIRS and MSU instruments and a threshold test on the spatial coherence of the HIRS spots into the box. Moreover, additional tests aim at an unambiguous distinction between clouds and special surface conditions (warm surfaces, snow). A particular problem at high latitudes is the presence of sea ice covering large areas which is not always easy to discriminate from an extended cloudiness. This problem can be at least partly solved by cloud detection algorithm. As explained in Claud et al. (1988) such a discrimination is attainable through the determination of the microwave surface emissivity, at the very beginning of the retrieval process, directly from earth-located calibrated radiances (see also Section 4.2).

Originally close to the tests defined by McMillin and Dean (1982) the present series of tests involved in 3I represents a substantial evolution with respect to this reference as it can be seen from the Appendix which describes all tests. For validation purposes, many comparisons between the results of the tests and either AVHRR high resolution pictures or nephanalyses or both have been carried out, apparently satisfactorily. In particular, snow covered areas, as well as sea ice, are well identified.

2.3 Initialization of the inversion problem

The feature space consisting of the TIGR data set just described may be used as a set of reference pattern vectors or prototypes to which the input pattern, made of the full set of parameters describing the real observation and a priori information, may be compared if a resemblance measure has been defined. The 3I algorithm makes use of a minimum distance criterion. A full set of parameters describing the observation (the feature vector) is made of the brightness temperatures observed in each channel plus the extra parameters describing the local properties of the observation: air mass, viewing angle, surface pressure and temperature (over sea), percentage of water in the terrain observed. If the two first parameters are easy to get, the latter two require knowledge of ancillary data. Use is made of a high resolution topography file and of operational forecasts. From these parameters the classification proceeds through the hierarchy organizing the TIGR data set and finally reaches the proper set of m atmospheric conditions and their associated set of computed brightness temperatures, Jacobians, etc. Proper means that the computations have been made in the exact conditions of the observations. Each of the m sets of brightness temperatures is then compared to the observation set which is classified as from class i if the quantity

$$D_i = \frac{1}{n} \sum_{k=1}^{n} \left(T_{obs}(k) - T_i(k)\right)^2 / S_k^2 \qquad i=1,\ldots,m$$

is the minimum. In this expression, n is the number of observations which have been selected, the T's are observations values (T_{obs}) or the corresponding values for the set i of the feature space and S_k is the variance of T(k) over the sample of m situations. Auxilliary data may also be included in the definition of D_i like, for example, operational forecasts of the temperature of the lowest layer. This approach to the distance between two situations is equivalent to that based upon a principal component analysis of the sample, at least when the same set of channels is used to express the idea of proximity in the two approaches. This method for retrieving the initial guess solution, from which the final solution is obtained, is particularly flexible since as many extra parameters as wanted, provided they are useful, may be added to the full set of observations, or equivalent, to each feature vector. Moreover, the ensemble of channels comprised in D_i may be easily varied, for example from a set of channels insensitive to the cloud cover in case of cloudy areas, to a more complete set in case of clear areas or as soon as cloud cleared brightness temperatures have been obtained. This approach leads to finding, within the TIGR data set, the element the closest to the input pattern. This particular situation may eventually bear signatures of too local phenomena (boundary layer, lapse rate, tropopause, etc.). This difficulty is overcome by modifying the proximity recognition in the following way: the above single element class is enlarged by considering, in the feature space, all the elements whose distance, d, to the input pattern is such that:

$$d \leqslant d_{MIN}\ (1 + \alpha) = d_R$$

where d_{MIN} is the distance to the closest element. In the present application, α has been given the value 0.25. The initial guess solution is obtained by averaging the elements within the "distance circle" of radius d_R, the "center" of which is the element the closest to the input pattern. Experience with the method has shown that, on the mean, about 10 elements are averaged to provide the initial guess solution. The minimum value is 1 (no average) to about 40 elements being averaged at maximum. Important improvements in retrieval accuracies have been demonstrated (Chedin and Scott, 1985).

2.4 Cloud Clearing

The procedure for retrieving the initial guess solution in the proper sub-set of the TIGR data set (the one corresponding to the observation conditions: viewing angle, surface pressure, emissivity, ... etc.) is direct in case of cloud free areas, whereas two steps are required for partly cloudy or cloudy areas (see Chedin et al., 1985). In the first step, comparisons are made between the observations and the archived brightness temperatures on the basis of a restricted set of channels insensitive to clouds. Operational forecasts of the temperature at the lowest level are also included in this proximity search. From this preliminary initial guess is started the cloud clearing algorithm based upon the so-called "ψ-method" (Chedin and Scott, 1983; Chedin et al., 1985). Following this method, a cleared infrared brightness temperature is obtained by adding to the initial guess value of the channel considered, the difference between the observed and the initial guess values of a microwave channel, almost insensitive to clouds and peaking approximately at the same level:

$$\psi^{HIRS}(i) = T^{HIRS}_{closest}(i) + \left[T^{MSU}_{OBS}(j) - T^{MSU}_{closest}(j)\right]$$

where $T^{HIRS}_{closest}(i)$ and $T^{MSU}_{closest}(j)$ are, respectively, the brightness temperatures of HIRS-2 channel i and MSU channel j associated with the first initial guess and $T^{MSU}_{OBS}(j)$ is the observed brightness temperature of MSU channel j. When index i takes the values 4, 5, 6 and 15, index j respectively takes the values 3, 2, 2 and 2. For i = 3, the reference MSU value is half the sum of $T^{MSU}(3)$ and $T^{MSU}(4)$. The ψ-method simply relies upon the difference between two channels peaking at approximately the same level and obtained for a situation (the initial guess) relatively close to that observed. This difference is applied to a microwave observation that is almost completely insensitive to the clouds and provides the clear infrared observation. Peaking in the lower troposphere (920 hPa) the brightness temperature of pseudochannel 14, $\psi^{HIRS}(14)$, is obtained through a regression of the type

$$\psi^{HIRS}(14) = b_0 + \sum_i b_i T^{HIRS}(i) + \sum_j c_j T^{MSU}(j),$$

where i takes the values of 2-6 and 15, and j the values 2 and 3. The coefficients of this regression are obtained from the data archived in each TIGR data subset. Applied to the synthetic data of the TIGR data set, the regression provides values of $\psi^{HIRS}(14)$ which, when compared to the true $T^{HIRS}(14)$ values, show no bias and a standard deviation varying from 0.1 to 0.3 K according to the viewing angle and surface pressure.

Contrary to the well-known N* method (Smith, 1968), which uses at least two spots to eliminate the effects of the clouds, the ψ-method uses only one spot and consequently avoids the assumption that the difference in radiance for a given channel in the two spots is due only to a variation of the amount of clouds of the same type (same height in particular). The result is that the N* method is subject to noise when this assumption does not hold. Moreover, the N* approach applies to radiances that have been corrected for the viewing angle (equivalent nadir viewing radiances) through statistical regressions whereas the ψ-method is made within the TIGR sub-set corresponding to the real conditions of observation.

The strong correlation between the initialization of the inversion process and the cloud clearing technique must also be pointed out: the quality of the initialization and that of the cleared brightness temperatures are directly connected. As explained above, if the proximity recognition is direct in case of clear areas, two steps are necessary in case of cloudy areas. More details are now given on the proximity search in each case.

For cloudy areas, the first step relies upon channels not or almost not sensitive to the clouds, plus extra parameters (temperature at the lowest standard level, temperature of the lower stratosphere, T(LS), obtained from a regression over the proper TIGR subset, in particular). The second step starts from this first guess and activates the ψ-method for declouding channels 3 to 6, 14 and 15 of HIRS-2. This two step procedure is iterated once. Table 1 gives the list of parameters used in the proximity recognition for the clear case and each step of the cloudy case. The results are: a set of channels, mostly sensitive to the temperature profile, decontaminated from the influence of the clouds and an initial guess solution selected among those archived in TIGR to start the temperature inversion. It is important to point out that the initial guess is the result of the averaging of the elements found in TIGR as being the closest to the input pattern (observations and extra parameters). No modifications are made to the situations extracted from TIGR at the eventual exception of the surface temperature for observations over sea when reliable independent data are available. The absence of channels not sensitive to the clouds and having absorption properties similar to the HIRS-2 channels that respond to water vapour - channels 10, 11, 12 but also 7, 8, ... etc. precludes applying the ψ-method to the declouding of these channels. As explained below, they are cleared after the temperature retrieval and the determination of reliable cloud parameters: mean top pressure and amount. The high sensitivity of these parameters to errors in the temperature profile justifies the use of the final profile instead of the initial guess for their determination.

Type of area observed	Parameters used for the recognition HIRS-2 Channels	 MSU Channels	 Extra Parameters
Clear	1, 2, 3, 4	2, 3, 4	T_{1000mb}
Cloudy 1st step	1, 2	2, 3, 4	T_{1000mb}, T(LS)
2nd step	1,2,4,14,15	2, 3, 4	T_{1000mb}, T(LS)

Table 1 Parameters used for the proximity recognition in the feature space TIGR

2.5 Temperature inversion algorithm

The radiative transfer equation may be written as:

$$y(\nu_i) = \int X(\nu_i, z)\, \beta(z)\, dz \tag{1}$$

where y represents the observations, β the atmospheric temperatures, the ν_i's summarize the conditions of observations (frequencies, viewing angles, surface properties, etc.), $X(\nu_i, z)$ is the kernel including transmittances and z is directly related to the altitude. A linear model may be established using an appropriate quadrature formula giving the matrix equation:

$$y_e = X\,\beta + e$$

where $y_e = y + e$, e being the associated error. A Bayesian estimator of β (maximum probability estimator: see Rodgers, 1970) may be written as:

$$b_{MAP} - \beta_o = [X'\, S_e^{-1}\, X + S_\beta^{-1}]^{-1}\, X'\, S_e^{-1}\, (y_e - y_o) \tag{2}$$

where β_o and y_o are the mean values of β and y respectively, S_e and S_β being the covariance matrices of e and β. In Eq. (2), MAP stands for "maximum a posteriori".

Since it starts from an optimum initial guess, which is the situation closest to the observed situation among those archived in the TIGR data set, the "3I" method aims at estimating the difference between the parameters and their values for the closest situation rather than the parameters themselves as in Eq. (2). This can be rewritten as:

$$\beta^* - \beta_o^* = [X'\, S_e^{-1}\, X + S_{\beta^*}^{-1}]\, X'\, S_e^{-1}(y^* - y_o^*) \tag{3}$$

where $\beta^* = b_{MAP} - \beta_{IG}$, $y^* = y_e - y_{IG}$, β_{IG} is the initial guess solution associated with y_{IG} and S_{β^*} is the covariance matrix of β^*. We have verified that β_o^*, mean of β^*, and y_o^*, the mean of y^*, are both equal to 0, which means that the initial guess retrieval procedure is unbiased. Consequently, Eq. (3) is equivalent to:

$$b_{MAP} - \beta_{IG} = [X' \, S_e^{-1} \, X + S_{\beta*}^{-1}]^{-1} \, X' \, S_e^{-1} \, (y_e - y_{IG}) \tag{4}$$

$S_{\beta*}$ may be easily computed from the "TIGR" data set, as explained in Chedin et al. (1985) where more details on the inversion algorithm are also given. In particular, it is shown that $S_{\beta*}$ is much more diagonal than S_β which renders the latter way of computing b_{MAP} more attractive than the usual one.

2.6 Cloud Parameters Retrieval: the "Coherence of Effective Cloud Amounts" Method

Considering a partially cloudy field of view (cloud amount N), the radiance measured by the satellite in channel i can be expressed as (Smith, 1968):

$$I_{m_i} = (1 - N\varepsilon_i)I_{Cl_i} + N\varepsilon_i I_{Cd_i}(P_c) \tag{5}$$

where I_{Cl_i} is the clear radiance, $I_{Cd_i}(P_c)$ the radiance arising from a black cloud at level P_c and ε_i the cloud emissivity. $N\varepsilon_i$ is called the "effective cloud amount" for channel i. From Eq. (5), N and ε_i cannot be obtained separately. The use of Eq. (5) (written for several channels) to determine $N\varepsilon_i$ and P_c requires knowledge of I_{Cl_i} and the function $I_{Cd_i}(P_c)$. These values can be obtained from the TIGR data set. More precisely, the TIGR data set provides an estimation of I_{Cl_i} and of $I_{Cd_i}(P_c)$ in all HIRS-2 channels for a set of cloud levels P_c (30 values from the surface to 100 mb) and for all the possible conditions of observations.

The method is based upon the consideration of a plausible variation of the cloud emissivity $\varepsilon_c(\nu)$ with frequency. From the conclusions of a study by Yamamoto et al. (1970), we assume that $\varepsilon_c(\nu)$ is roughly constant in the longwave region (5 channels: 4 to 8) and in the shortwave region (5 channels: 13, 14, 15, 18, 19) and, moreover, that the shortwave value is slightly lower than the longwave one. The algorithm can be divided into two steps:

1) Selection, among the initial set of 23 pressure levels, of a sub-set of plausible cloud top levels. A level is plausible if the following test is verified: we consider the function:

$$S(P_c) = \sum_{(i,j)} \left[\left(\frac{I_{m_i} - I^o_{Cl_i}}{I_{m_j} - I^o_{Cl_i}} \right) - \left(\frac{I^o_{Cd_i}(P_c) - I^o_{Cl_i}}{I^o_{Cd_j}(P_c) - I^o_{Cl_j}} \right) \right]^2 \qquad (6)$$

for $(i,j) = (4,5), (5,6), (6,7)$

where $I^o_{Cl_k}$ and $I^o_{Cd_k}(P_c)$ are respectively the estimations of I_{Cl_k} and $I_{Cd_k}(P_c)$. Eq. (6) derives from Eq. (5), written for each pair of channels considered, after $N\varepsilon_i$ has been eliminated.

The test consists in verifying that $S(P_c)$ is close to its minimum value S_{min}. In the pairs of adjacent channels technique, the cloud top level P_c would be the one for which $S(P_c) = S_{min}$. Here the test limit is: $S(p_c)/S_{min} \leqslant 5$, which, on the mean, eliminates one half of the 23 a priori possible cloud levels.

2) Tests on the coherence of effective cloud amounts. For each plausible cloud level P_c, corresponding values of $N\varepsilon_i$ are computed in each channel from Eq. (5). This step is itself divided into two parts:

- Elimination of obviously impossible levels: various tests are carried out, in each spectral region, to make sure that the individual values $N\varepsilon_i$ are not too much scattered, and that both mean values (respectively $N\varepsilon_{LW}$ and $N\varepsilon_{SW}$ for the longwave and the shortwave region) range between 0 and 1, and that $N\varepsilon_{SW}$ is slightly lower than $N\varepsilon_{LW}$.

- Choice of the final cloud top level: the ratio of the standard deviation of the effective cloud amounts $N\varepsilon_i$ by the mean value $N\varepsilon_{LW}$ in the longwave region must be minimum, which ensures the smallest scattering.

For daylight observations, shortwave channels cannot be used because of their contamination by reflected solar radiation. In that case, a simplified algorithm is applied in which all tests using shortwave channels are excluded.

The algorithm has been validated on synthetic data and compares favourably, in particular for low clouds, with those obtained by Wielicki and Coakley (1981) with the pair of adjacent channels technique. Knowledge of the cloud parameters makes it possible to "clear" all the channels influenced by clouds and not corrected by the ψ-method. More details are given in Wahiche et al. (1985).

2.7 Water Vapour and Surface Temperature Retrievals

Retrievals of water vapour amounts are made for three layers delimited by the levels 1000 mb, 800 mb, 500 mb and 300 mb in addition to the total amount of precipitable water vapour. Following the temperature profile retrieval, the brightness temperatures associated with the initial guess are corrected for the deviations between the initial temperature profile and the final solution giving rise to the initial guess for water vapour and surface temperature retrievals. The method used is based upon a simultaneous physical inversion of the water vapour amounts and of the surface temperature.

Four HIRS-2 channels are presently used: 8, 10, 11 and 12 for day-time observations and 18, 10, 11, 12 for night-time observations. A ridge type estimation procedure is used:

$$\Delta\beta = (X'X + \gamma I)^{-1} X'\Delta Y$$

where X (X', transpose of X) is the matrix of the partial derivatives of the brightness temperatures with respect to relative humidities and surface temperature, ΔY the difference between observed (cleared) brightness temperatures and the initial guess (corrected as explained above), $\Delta\beta$ the difference between the final and initial values of the parameters considered, γ a smoothing parameter (Lagrangian parameter).

2.8 Microwave Surface Emissivity

Microwave surface emissivity is an important parameter containing information on the type of the surface (ice and snow cover in particular). It is directly extracted from the measurements in the first channel of the microwave sounding unit, a window channel (see Wahiche, 1984).

Comparisons have been made between emissivity values obtained at the term of the retrieval process and values obtained at its very beginning (see Section 2.2). The agreement is satisfactory with differences less than 0.1 in emissivity.

2.9 Quality Control

An important aspect of any retrieval system is its ability to detect and to reject bad satellite data (due to technical errors like transmission, for example) and bad retrievals (due, for example, to heavy cloudiness or precipitations, to the occurrence of very marginal meteorological situations etc).

From the past experience with 3I it appears that bad satellite observations are well detected and rejected. The method used first compares MSU observations to MSU synthetic brightness temperatures calculated from the analysis (or the forecast) and rejects too big differences. In a second step, the comparison between the set of observations and corresponding sets archived in TIGR leads to a large "proximity" distance in case of bad data, thus allowing for their rejection.

For quality control purposes, all retrievals pass a series of tests aiming at detecting bad or doubtful retrieved profiles. A first class of tests consists of comparing either the proximity distance or the difference between the observed and the initialization values for such or such channel to limit values (varying with the air mass type). A second class of tests consists of comparing the initial solution for the temperature profile to the retrieved one. Big differences are not allowed as being in contradiction with the principle of optimised initialization. A third class of tests, presently being evaluated, consists of rejecting the retrievals for which not enough closest situations were found to be in the so-called "distance circle", revealing quite a marginal case. A statistical analysis of the sample in this distance circle should also be implemented soon.

3. CONCLUSIONS ON THE 3I METHOD

The 3I method for retrieving geophysical parameters from observations of the US operational weather satellites (TIROS-N series) is based upon a pattern recognition type approach aiming at optimizing the initial guess solution to the inversion process. This research of the initial guess, from among a vast pre-established library, takes care of the observing conditions and a priori information like the viewing angle, the surface properties, the air mass type, and estimate of the lower stratosphere temperature or also forecasts available at the time of the processing of the observations. The actual initial guess is the average of the elements of the library the closest to the 'input pattern' made of satellite observations and a priori information. None of the information external to the library enter the initial guess at the unique (eventual) exception of the surface temperature for over sea observations. It is also to be noted that the initial guess for temperature retrieval is not identical to the one used for water vapour and surface temperature retrievals. From the former to the latter a correction is made to take into account the differences between the initial and retrieved temperature profiles.

This direct (non-iterative) step by step approach to the retrieval problem is very general and answers to a large extent the problem of how involving more physics in the process without coming to prohibitive computation times. The solution proposed here is: sampling, pre-computing (once) using a sophisticated forward model, archiving, and retrieving the information through pattern recognition.

Numerous applications of the 3I system have established its accuracy and sensitivity to mesoscale meteorological features (gradients, orientation of troughs or ridges, etc.). A few recent ones are presented in the next section.

4. RECENT APPLICATIONS OF THE 3I SYSTEM

4.1 Global use of TOVS Retrievals at ECMWF

Since December 1986, the 3I code has been implemented, validated and run in global experiments using a data set of raw radiances (level 1B data) provided to the European Centre for Medium Range Weather Forecasts (Reading, U.K.) by NOAA/NESDIS (National Oceanic and Atmospheric Administration/National Earth Satellite Distribution and Information Service, Washington DC, USA). With the 3I system the complete processing from the level 1B raw counts for two satellites (NOAA-9 and NOAA-10) for a period of six hours requires 23 minutes CPU, using one processor of the CRAY XMP48.

The data assimilation at ECMWF (Kelly and Pailleux, 1988) is done every 6 hours and uses conventional observations, cloud track winds and satellite determined profiles (Satems). NESDIS Satems consist of geopotential thicknesses and precipitable water contents at approximately 250 km resolution. The 3I profiles have been collocated with radiosondes, to assess the quality of the retrievals, and a similar comparison has been done for the NESDIS "Satems". The maximum time difference for a comparison is 3 hours, with a maximum distance difference of 100 km. Radiosondes of questionable quality (transmission problems, etc.) have been eliminated for differences greater than 10 K. Results are shown for a period of 5 days (30 January to 3 February 1987), for 3 latitude bands and all longitudes. Figure 3 is for latitudes 60°N to 90°N with NESDIS statistics on the left and 3I statistics on the right. Comparisons are made between in-situ and retrieved mean layer virtual temperatures for 11 layers between the levels: 1000, 850, 700, 500, 400, 300, 200, 100, 70, 50, 30, 10 hPa. On Fig. 3, 3I biases (dashed lines) are smaller, and standard deviations are roughly similar except for the lowest layer for which 3I is significantly better. The number of collocations are different for the two codes as a result of different spatial resolutions and screening procedures. Here a factor close to 5 is observed that should result for 3I in a lower bias - this is the case and a larger standard deviation - this is not the case. Figure 4 is for latitudes 20°N to 60°N. Very large numbers of collocations are observed with a ratio of about 3.5 between 3I and NESDIS. However, 3I biases are slightly larger, which could possibly be explained by remaining validation problems, whereas the 3I standard deviations are slightly smaller, in particular for the lowest layers. Figure 5 is for latitudes 60°S to 90°S. The ratio between 3I and NESDIS collocation is about 3 and here again 3I biases are equal to or larger (lowest layers) to those of NESDIS whereas 3I standard deviations are smaller. The period considered was characterized by highly perturbed areas like, for example, over Europe and China-Siberia and variable conclusions may be obtained for more local comparisons. However, conclusions arising from Figs. 3-5 generally hold. It must also be pointed out that comparing a physical retrieval algorithm to the NESDIS statistical one on the basis of collocations with radiosondes is not an easy task since the latter is, by definition, heavily tuned towards radiosounding measurements.

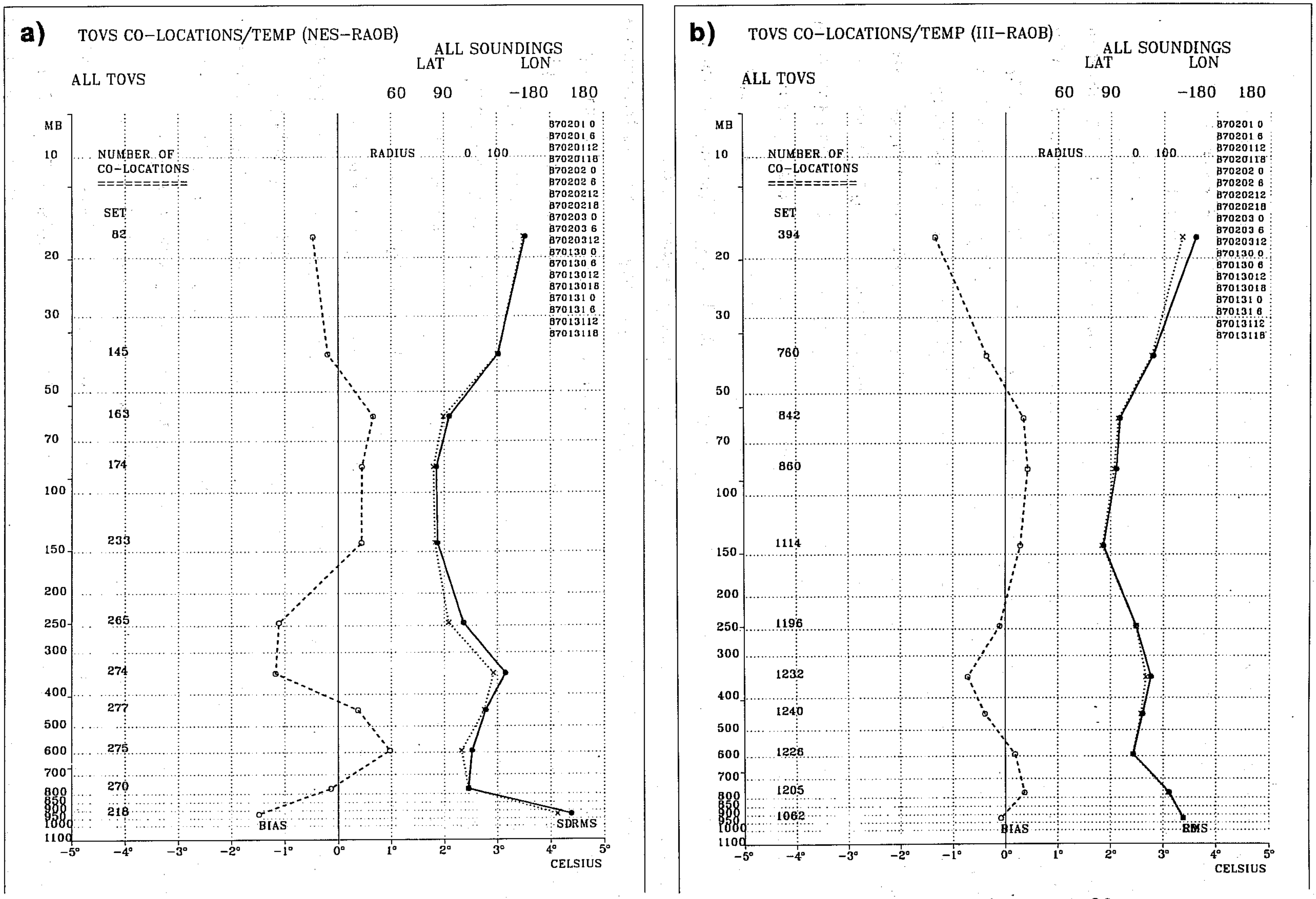

Fig. 3 Statistics with collocated radiosondes: area between 60°N and 90°N; the maximum distance difference is 100 km. Maximum time difference of 3 h.
Solid line: RMS, dotted line: standard deviation, dashed line: bias.
a) "Satems" produced by NESDIS; b) "Satems" produced with the 3I method.

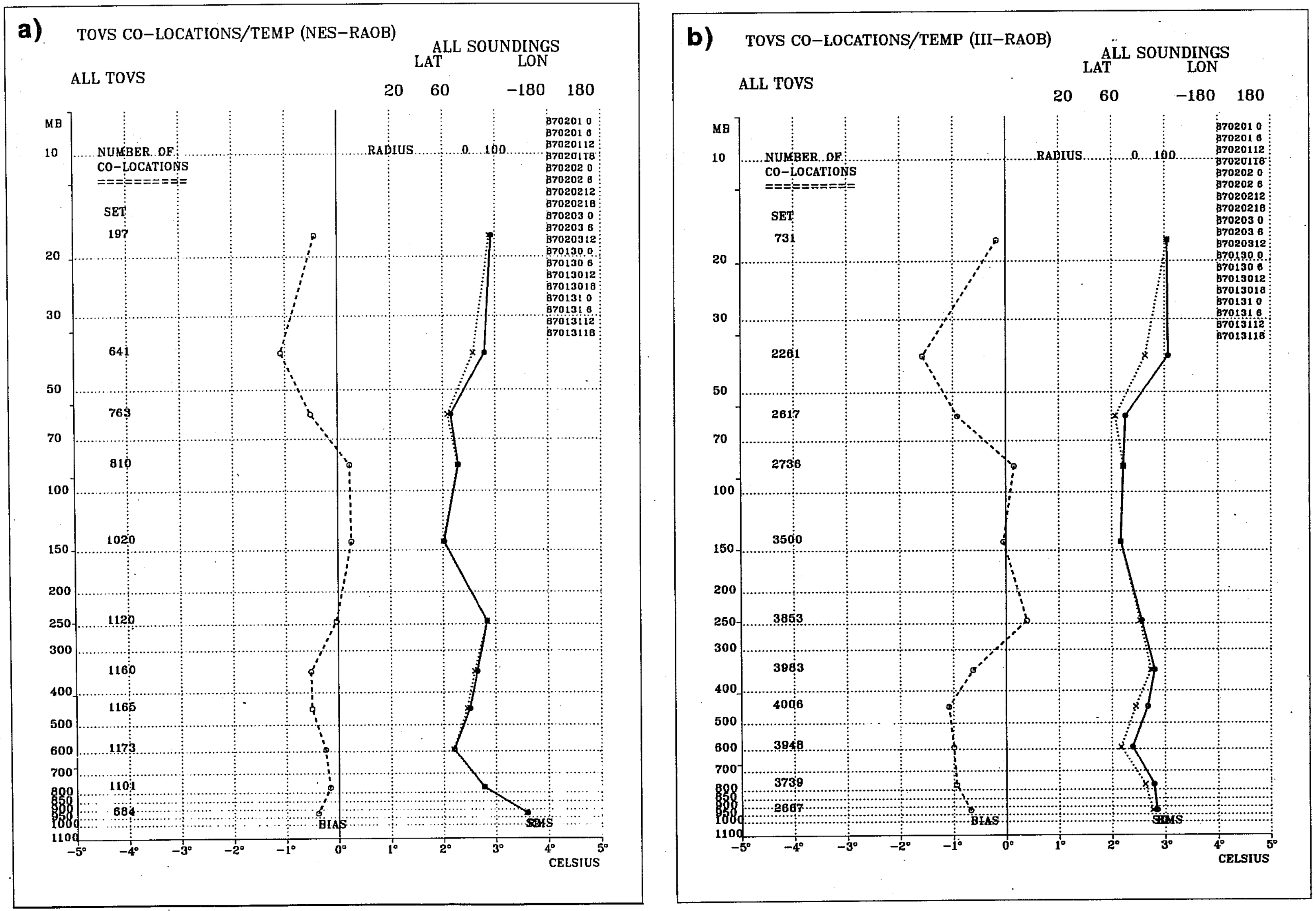

Fig. 4 Same as Fig. 3, area between 20 °N and 60 °N.

Comparisons have also been made between water vapour relative humidity profiles and collocated radiosondes (3 hours in time; 100 km in space). Figure 6 illustrates the results currently obtained for 5 layers between 1000 hPa and 300 hPa. The standard deviation varies between 15% and 20%, roughly constant. The positive bias has been voluntarily introduced (for a greater impact in the assimilation step). On Fig. 6, collocations correspond to 5 days: 30 January 1987 to 3 February at 0.00 Z.

Numerous other comparisons have been made, leading to the general conclusion that 3I results are at least as good as NESDIS results. Although the quality of the statistics between retrieved profiles and collocated radiosondes appears relatively similar, the assimilation of these temperature and water vapour profiles in the analysis leads to significant changes, particularly in the stratosphere and the low troposphere (see Flobert, 1988).

These significant changes in the analyses are likely to cause changes in the forecasted fields. Details are given elsewhere in the present issue.

4.2 Retrieval of mesoscale meteorological parameters for polar latitudes

Observations from the satellites of the TIROS-N series on polar regions, where in-situ data are very scarce are essential for the study of the interactions between ocean, ice and atmosphere through the retrieval of atmospheric and surface parameters. The MIZEX (Marginal Ice Zone EXperiment) and ARCTEMIZ (MIZ of the European ARCTic) campaigns, conducted respectively during the summer 1984 in the Fram Strait and since 1986 north and south of Fram Strait, gave the opportunity of applying the 3I retrieval algorithm to high latitude observations. Associated to numerous in situ data, these campaigns also allowed for comparisons with the retrieved products.

Due to particularities of polar regions, refinements were brought to the method and are presented in Claud et al. (1988). One important point concerns the cloud detection algorithm and the problem of discriminating between sea ice and extended cloudiness since they may have similar radiative properties. To overcome this difficulty, two different methods were developed, which discriminate sea ice from open water directly from the satellite observations, and more specifically from the MSU instrument which channel 1, an atmospheric window located at 50 GHz, displays an interesting property related to its

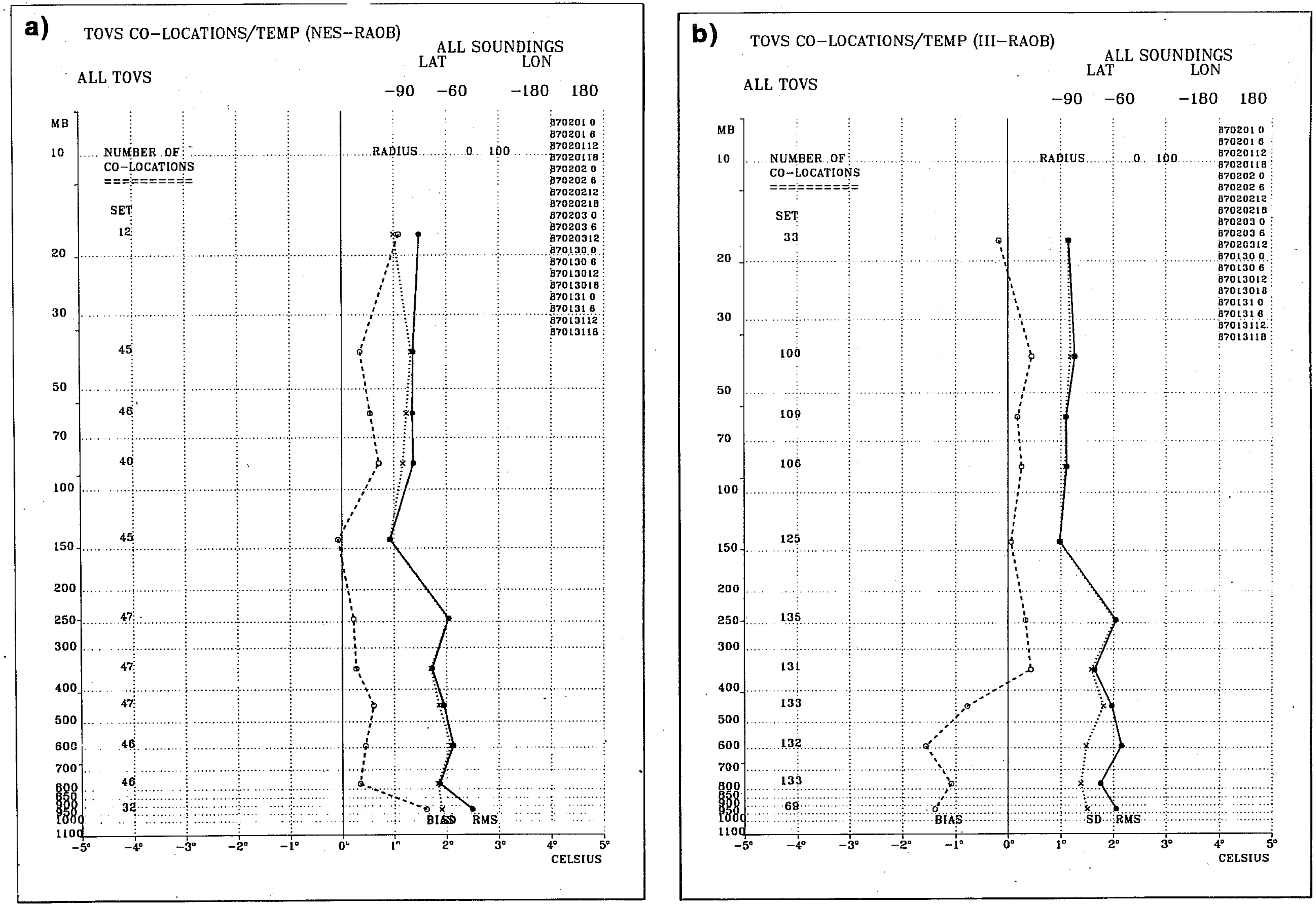

Fig. 5 Same as Fig. 3, area between 60°S and 90°S.

sensitivity to the emissivity of the surface, itself varying rapidly with the type of the surface. Currently admitted values range from 0.45 to 0.60 for open water, to more than 0.75 for multi-year sea ice and to 0.95 for young sea ice. Determining surface emissivity is consequently a way to detecting the presence of sea ice.

4.2.1 Direct statistical method

This method relies upon the fact that channel MSU1 is sensitive to surface emissivity and to temperature in the lower atmosphere and down to the surface whereas channel MSU2 is mostly a function of the temperature of the low atmosphere and relatively not sensitive to the surface because of its low overall transmission (close to 0.1). It is thus possible to extract the surface emissivity from a combination of these 2 channels. Such a particularity has already been used by Grody (1983).

A relation between the surface emissivity ε, the brightness temperature of MSU channel 1, TBMSU1, and MSU channel 2, TBMSU2, has been established under a form similar to the one found by Grody:

$$\varepsilon = a + b \, TBMSU1 + c \, TBMSU2$$

In Grody's approach, the coefficients have been obtained from a sample of data restricted in space (United States) and in time (April 1979).

In our study, the coefficients a, b, c have been regressed independently for the polar situations and for the mid-latitude situations archived in TIGR. They are given in Claud et al. (1988).

4.2.2 Direct physical method

Microwave surface emissivity (ε) may also be extracted by considering the radiative transfer equation itself. The measured radiance can be written as follows:

$$I = \varepsilon \, [\tau B_s - \tau \int^{\downarrow} B d\tau] + [\tau \int^{\downarrow} B dz + \int^{\uparrow} B d\tau]$$

For a given class (polar or mid latitude), it has been found that the quantities τ, $\int\uparrow B d\tau$ and $\int\downarrow B d\tau$ may be considered as constant to within a good approximation, at least with the purpose of determining ε to the required accuracy. Actually, the ratio of the standard deviation by the mean of each of these three quantities varies from 1% to 2%.

As a consequence and using the Rayleigh Jeans approximation, the preceding equation may be written as:

$$TBMSU1 = \varepsilon\ [\tau\ T_s - \alpha] + \beta$$

where τ, α and β are constants (one set for each air mass type). Knowledge of MSU channel 1 brightness temperature directly leads to that of ε.

For each of the orbits, selected from MIZEX and ARCTEMIZ campaigns, surface microwave emissivities have been calculated by the two direct methods described above. Results are very similar apart from a systematic bias of about 0.1 (larger values for the physical method). However, such a bias does not prevent either method from identifying sea ice covered areas. The mean of the values given by these two methods is very close to the value obtained through the full physical inversion method (Wahiche, 1984). Results were also favourably compared to the sea ice maps currently prepared by the Norwegian Meteorological Institute.

4.2.3 Impact of this information on cloud detection

Two tests, which are also part of the operational NOAA/NESDIS retrieval algorithm (McMillin et al., 1982) may benefit from an a priori knowledge of the presence of sea ice: the so-called "frozen sea" test and the albedo test. Details on the modifications introduced are given in the Appendix and more details on the methods reported here are given in Claud et al. (1988). Results are illustrated on Fig. 6 for a pass of NOAA-9 (5 June 1986, 9.55 Z). Fig. 6a shows the results of cloud detection tests before any modifications and Fig. 6b after the modifications. One can see in an area close to North Pole that low clouds, not detected previously, are now detected.

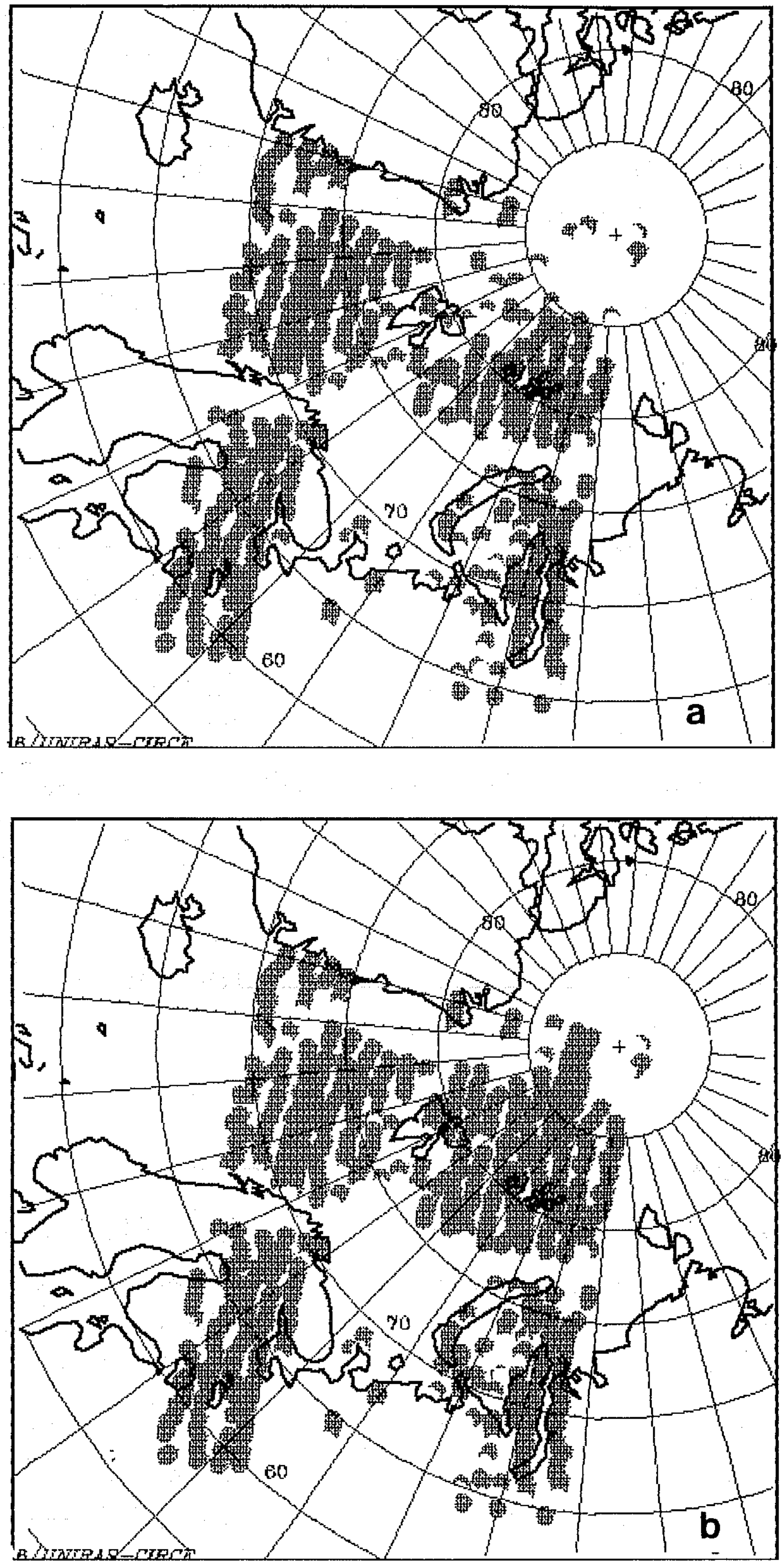

Fig. 6 Detection of clouds in 3I boxes for a pass of NOAA-9 (5 June 1986, 9.55Z). a) before modification; b) after modification. Dark areas correspond to cloudy areas. At the upper right of Fig. 6b), low clouds over sea ice have been detected.

4.3 Detection of snow from HIRS-2

Because of its high albedo, snow cover is often seen as cloud by cloud detection algorithms during day time. Moreover, snow and low cloud top temperatures may be similar. It has already been reported in the literature (see for example Kidder et al., 1984) that comparisons between channels at 11 μm and 3.7 μm could help discriminating between snow and clouds. Computation of the ratio:

$$\{T_B(\text{HIRS},19) - T_B(\text{HIRS},8)\} \,/\, \cos\theta_s,$$

where $T_B(\text{HIRS},n)$ is the brightness temperature of channel n of HIRS-2 and θ_s the solar zenith angle, leads to high values for the clouds and to much lower values for snow covered areas. This is due to the low albedo of snow at 3.7 μm as compared to clouds and to the fact that the ratio given above isolates (and normalizes) the solar contribution to the brightness temperature of channel 19.

The following test has consequently been implemented in the cloud detection algorithm: if, simultaneously, the albedo is greater than 20%, and an estimate of the surface temperature (regression based upon channel 8) smaller than 273 K and the ratio

$$\{T_B\,(19) - T_B\,(8)\} \,/\, \cos\ _s,$$

smaller than 14 K, the field of view considered is supposed to be clear, with snow covering it. Several NOAA-9 passes have been processed, for which manual nephanalyses had clearly identified snow covered areas. Figure 7 illustrates one example. Snow is in white and the areas concerned are in perfect agreement with manual analysis.

5. CONCLUSIONS

The well known problem of the non uniqueness of the solution to the non linear physical satellite retrieval problem may be greatly simplified by a proper initialization of the inversion scheme. Knowledge of a priori information on the situation observed may greatly help in selecting an initial guess solution close to the final correct one.

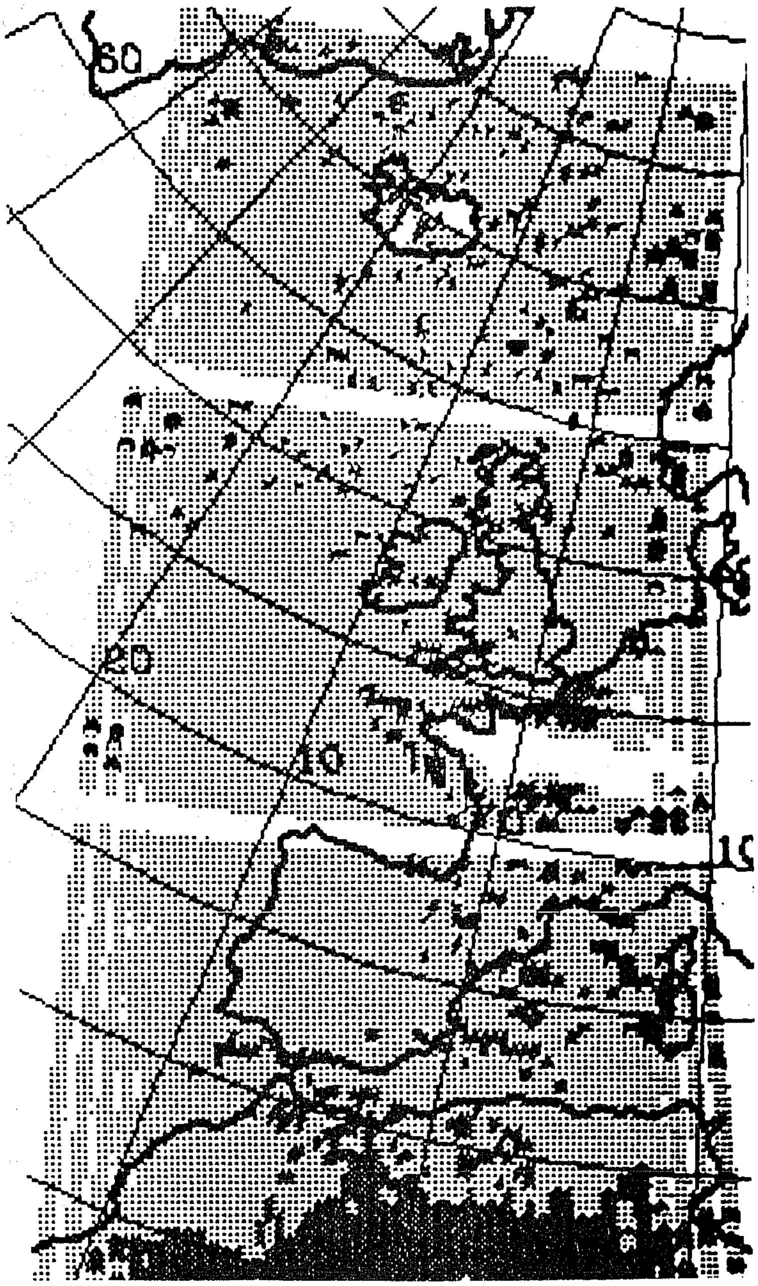

Fig. 7 Snow detection from HIRS-2. Snow in white; clear areas in dark cloudy areas in grey; NOAA-9 orbit 6209 (25 February 1986).

This is the reason for which the Improved Initialization Inversion procedure is based upon the optimization of the initial guess solution from which the final solution is obtained by a simple Bayesian estimation method. The initial guess is obtained from a pattern recognition type approach by comparing the satellite observations - associated with other parameters like the air mass type, the viewing angle, the type of the surface and the surface pressure obtained from a high resolution topography file and geopotential heights operational forecasts, actually the input pattern - to similar data precomputed for a large set of atmospheric conditions, the feature space or TIGR (TOVS Initial Guess Retrieval) data set. The quality of the initialization is illustrated by the fact that only one inversion is enough to provide an accurate "final" solution. An improved cloud detection algorithm, including a discrimination between either sea ice or snow and low clouds has been satisfactorily tested. Numerous local and now global applications have demonstrated its quality. In the latter case (global) work is in progress towards an improved and more operational validation of the forward model and a reinforced quality control.

Acknowledgements

The 3I system has been developed in close cooperation with Dr. N.A. Scott and with the members, past or present, of the ARA (Atmospheric Radiation Analysis) group at LMD: N. Husson, C. Wahiche, P. Moulinier, J.F. Flobert, P. Moine, C. Claud, C. Levy, F.M. Breon, Y. Tahani. J.F. Flobert succeeded in installing 3I at ECMWF and run the global applications. All of them are gratefully acknowledged. I am particularly happy to thank J. Pailleux and G. Kelly (ECMWF) for fruitful discussions. We are deeply indebted to ECMWF and its staff: Drs. L. Bengtsson, D. Burridge, A. Hollingsworth who helped us at important stages of this work. Thanks are also due to Mrs M. Simpson (ECMWF) who typed the manuscript and Mrs E. Kooij-Connally for help and assistance.

References

Chedin, A. and N.A. Scott, 1983: The improved initialization inversion procedure. Internal Note LMD No. 117.

Chedin, A., N.A. Scott, C. Wahiche and P. Moulinier: The improved initialization inversion method: a high resolution physical method for temperature retrievals from the TIROS-N series. J.Clim.Appl.Meteor., 24, 124-143.

Chedin, A. and N.A. Scott, 1985: Initialization of the radiative transfer equation inversion problem from a pattern recognition type approach. In: Advances in Remote Sensing Retrieval Methods. A. Deepak Publishing, 495-515.

Claud, C., A Chedin, N.A. Scott and J.C. Gascard, 1988: Retrieval of mesoscale meteorological parameters for polar latitudes. Annales Geophysicae. To be published.

Flobert, J.F., 1988: Global use of TOVS retrievals in NWP at ECMWF. Proc. of the 4th International TOVS Study Conference. Igls, Austria. To be published.

Grody, N.C., 1983: Severe storm observations using the Microwave Sounding Unit. J.Clim.Appl.Meteor., 22, 609-625.

Kelly, G. and J. Pailleux, 1988: Use of satellite vertical sounder data in the ECMWF analysis system. Proc. of the 4th International TOVS Study Conference. Igls, Austria. To be published.

Kidder, S.Q. and Huey-Tz'u Wu, 1984: Dramatic contrast between low clouds and snow cover in daytime 3.7 μm imagery. Mon.Wea.Rev., 112, 2345-2346.

McMillin, L.M. and C. Dean, 1982: Evaluation of a new operational technique for producing clear radiances. J.Appl.Meteor., 21, 1005-1014.

Moine, P., A. Chedin and N.A. Scott, 1987: Automatic classification of air mass type from satellite vertical sounding data. Application to NOAA-7 observations. Ocean Air Interactions, 1, 95-108.

Rodgers, C.D., 1970: Remote sounding of the atmospheric temperature profile in the presence of clouds. Quart.J.Roy.Meteor.Soc., 96, 654-666.

Smith, W.L., 1968: An improved method for calculating tropospheric temperature and moisture from satellite radiometer measurements. Mon.Wea.Rev., 96, 387-396.

Wahiche, C., 1984: Contribution au problème de la détermination de paramètres météorologiques et climatologiques à partir des données fournies par les satellites de la série TIROS-N. Impact de la couverture nuageuse. Thesis University of Paris - VII. 207 pp.

Wielicki, B.A. and J.A. Coakley Jr., 1981: Cloud retrieval using infrared sounder data: error analysis. J.Appl.Meteor., 20, 157-169.

Yamamoto, G., M. Tanaka and S. Asano, 1970: Radiative transfer in water clouds in the infrared region. J.Atmos.Sci., 27, 282-292.

APPENDIX: CLOUD DETECTION ALGORITHM

The cloud detection algorithm aims at assigning the flag clear of cloudy to each 3I box (see Fig. 1). There are two types of tests: the first ones may declare a box as clear, and, in that case, remaining tests are not executed. The second ones are executed sequentially and only those boxes having passed all tests are declared clear. There are two tests of the first type and eight of the second one. All tests are performed for each HIRS-2 spot.

A1 TESTS FOR CLEAR SKY DETECTION

A1.1 Hot surface test

A spot for which TB8, the brightness temperature measured for HIRS-2 channel 8, is greater than 300 K is declared clear. This test is of particular interest over desertic areas (but not only) which sometimes display anomalous emissivities resulting in a wrong cloudy flag assignment by the tests of the second type.

A1.2 Snow covered areas

For channel 19, at 3.7 μm, contrary to the clouds, snow solar reflectivity is low. In case of snow, the quantity $[TB_{19} - TB_8]/\cos(\theta_s)$, where θ_s is the local solar zenith angle, is much lower than in case of clouds. It is here assumed that the thermal signals of both channels are roughly the same.

The test is executed in daylight, over land. If albedo (estimated from channel 20) is greater than 25%, an estimate of surface temperature (see below) smaller than 275 K and the above quantity smaller than 14K, the spot is declared clear (presence of snow).

A2 CLOUD DETECTION TESTS

A2.1 Microwave window test

This test relies upon an estimate of the microwave surface emissivity (see section 2.2) and is executed over sea. If the emissivity is smaller than 0.65 (no ice) and an estimate of the surface temperature (see below) smaller than 271 K (ice) the spot is declared cloudy, the presence of clouds explaining the contradiction.

A2.2 Albedo test

Executed in daylight.

Over land: executed only if TB_8 is smaller than 285 K (to avoid hot desertic surfaces of high albedo). If albedo is larger than 20%, the spot is declared cloudy.

Over sea: a spot is declared cloudy if albedo is greater than 15%. For latitudes greater than 65(N) or smaller than -55(S) a test is made on the microwave emissivity value to avoid sea ice covered areas: no albedo test if the emissivity is greater than 0.65.

A2.3 Window channels test

Three estimates of the surface temperature are obtained from regressions over the TIGR data subset corresponding to the actual conditions of observation. The three regressions are:

$$(Tsurf)_{regr}(8) = a_1\ TB_5 + a_2\ TB_6 + a_3\ TB_7 + a_4\ TB_8 + a_5\ TB_{11} + a_6\ TB_{MSU2} + a_7\ TB_{MSU3} + a_8$$

$$(Tsurf)_{regr}(18) = a_1\ TB_8 + a_2\ TB_{10} + a_3\ TB_{18} + a_4\ TB_{MSU2} + a_5$$

$$(Tsurf)_{regr}(19) = a_1\ TB_8 + a_2\ TB_{10} + a_3\ TB_{11} + a_4\ TB_{19} + a_5\ TB_{MSU2} + a_6\ TB_{MSU3} + a_7$$

The test consists in comparing these different estimates. At night, a spot is declared cloudy if:

$$(Tsurf)_{regr}(18) - (Tsurf)_{regr}(8) > 2.5\ K$$

or

$$(Tsurf)_{regr}(8) - (Tsurf)_{regr}(18) > 4\ K$$

or

$$(Tsurf)_{regr}(19) - (Tsurf)_{regr}(18) > 2.5\ K$$

or

$$(Tsurf)_{regr}(18) - (Tsurf)_{regr}(19) > 4\ K$$

At day, the test is executed if albedo is smaller than 20%. A spot is declared cloudy if

$$|(Tsurf)_{regr}(18) - (Tsurf)_{regr}(8)| > 12 \text{ K}$$

In the regression based upon channel 8 it should be noted that channel 10 of HIRS-2 is not included. This regression is used in daylight for providing surface temperature estimates and the presence of channel 10, which may be severely contaminated by the so-called "Reststrahlen" effect (at day, over desertic areas) could lead to estimates very far from the reality.

A2.4 <u>Interchannel regression test</u>

Clouds may have very different effects on TOVS channels according to their wavelengths. Their presence may consequently be detected from regressions combining predictors and predictands corresponding to different spectral domains.

Three regressions are used:

$$\langle TB_7\rangle = a_1\ TB_{13} + a_2\ TB_{14} + a_3\ TB_{15} + a_4$$

$$\langle TB_{MSU2}\rangle = a_1\ TB_3 + a_2\ TB_4 + a_3\ TB_5 + a_4\ TB_6 + a_5\ TB_7 + a_6$$

$$\langle TB_{MSU3}\rangle = a_1\ TB_3 + a_2\ TB_4 + a_3\ TB_5 + a_4\ TB_6 + a_5\ TB_7 + a_6$$

the coefficients of which are derived from the proper subset of the TIGR data set. A spot is declared cloudy:

if: $|TB_7 - \langle TB_7\rangle| > 2.7$ K at night;

if: $|TB_7 - \langle TB_7\rangle| > 3.2$ K in daylight;

or:

if: $|TB_{MSU2} - \langle TB_{MSU2}\rangle| > 2$ K.

A2.5 Surface temperature test

For clear conditions, surface temperature estimates should not be too far from conventional analysis values. A spot is declared cloudy if:

$$(T_{surf})_{analysis} - (T_{surf})_{regr} > 2 \text{ K over sea}$$

or:

$$(T_{surf})_{analysis} - (T_{surf})_{regr} > 6 \text{ K over land.}$$

The estimate is based upon channel 8 regression in daylight and channel 18 regression at night. This test is not executed if analysis values are not available.

A2.6 Low clouds test

At night, a spot is declared cloudy if:

$$TB_8 - TB_{19} > -0.5 \text{ K}$$

A2.7 Adjacent spots test

A spot adjacent to a given spot is declared "similar" if: the difference in the mean altitudes (over land) of the two spots is less than 250 m and if the difference in the percentage of water in the two spots is less than 30%. A spot is then declared cloudy if there is at least one similar adjacent spot such that:

$$[(Tsurf)_{regr}]_{adj} - (Tsurf)_{regr} > 1.5 \text{ K over sea}$$

or:

$$[(Tsurf)_{regr})_{adj} - (Tsurf)_{regr} > 3 \text{ K over land.}$$

If the spot remains clear and if there are at least three similar adjacent spots, a coherence test is executed. The spot will be declared cloudy if all the adjacent spots are such that:

$$|[(Tsurf)_{regr})_{adj} - (Tsurf)_{regr}| > 1.5 \text{ K over sea}$$

or:

$$|[(Tsurf)_{regr})_{adj} - (Tsurf)_{regr}| > 3 \text{ K over land.}$$

A2.8 Maximum value test

The maximum value, T_{max}, is that of $(Tsurf)_{regr}$ over the box which comprises the spot considered. The test is executed if that spot and the one corresponding to the maximum value have similar percentages of water (difference less than 30%). The spot is declared cloudy if

$$T_{max} - (Tsurf)_{regr} > 4 \text{ K}$$

A3 CLEAR/CLOUDY FLAG FOR A BOX

After all spots of a box have passed all tests, two cases may arise:

- if at least one spot is declared clear (two clear spots required when surface temperature analysis is not available), the box is declared clear. Variance tests are carried out on the mean of the brightness temperatures of channels 5, 7 and 13 and "out of statistics" spots are rejected. Brightness temperatures for the remaining spots are averaged;

- if all spots are declared cloudy (only one clear spot when surface temperature analysis is not available), the half corresponding to the largest values of channel 8 (day) or channel 18 (night) brightness temperature is retained. Variance tests are then applied to channel 7 brightness temperatures and "out of statistics" spots are rejected. Brightness temperatures for the remaining spots are averaged.