WATER VAPOR PROFILING WITH SSM/T-2 DATA
EMPLOYING A HYBRID INVERSION ALGORITHM

Markus J. Rieder and Gottfried Kirchengast
Institute for Meteorology and Geophysics, University of Graz (IMG/UoG), Austria

ABSTRACT

The feasibility of retrieving water vapor profiles from SSM/T-2 data is demonstrated by usage of a hybrid
inversion algorithm. The SSM/T-2 downlooking sounder data - consisting of brightness temperature
measurements in five microwave bands sensitive to water vapor (frequencies near 90/150/183 GHz) - can
be used, together with total water vapor content data, in order to compute water vapor profiles of about 3-
5 km vertical resolution.

The radiative transfer equation of this profiling problem yields a nonlinear mapping of state space into
measurement space, which is reflected in a significant nonlinearity in the cost functional to be minimized.
We employed a Bayesian approach to nonlinear inversion (linearized optimal estimation), complemented
by several extensions tailored to further minimize the nonlinear cost functional. We investigated, in
particular, problems related to the sensible guess of a priori information, the shape of probability density
functions involved, and the potential of Monte Carlo and simulated annealing techniques.

The resulting hybrid algorithm combines the strengths of library-furnished minimal-cost a priori,
Marquardt-Levenberg optimal estimation, simulated annealing, and iterative a priori weighting.
Based on synthetic state vector data, we show the utility and the characteristic features of each of these
algorithm components. The hybrid algorithm was found to cope well with the subtleties involved in
passive downlooking microwave sounding and to robustly deliver reliable water vapor profiles.

For an initial validation based on genuine data, we compared SSM/T-2 water vapor retrievals to co-
located water vapor profiles from ECMWF (European Centre for Medium-range Weather Forecasts)
 atmospheric analyses, taking the required auxiliary temperature information from the analyses. A
statistical validation indicated an accuracy of about 15 to 20 % for retrieved relative humidity profiles.

The developed algorithm can be readily extended to include all sensible sources of additional information,
on the state as well as from additional measurements. It is, furthermore, fairly straightforward to adapt it
for retrievals of water vapor and temperature from the Advanced Microwave Sounding Unit (AMSU).

Keywords: Inversion, Nonlinearity, Optimization, Bayesian Approach, Simulated Annealing, Monte
Carlo Techniques, Water Vapor Profiling

1. MATHEMATICAL SETUP OF THE PROBLEM

When performing a remote measurement it is usually one in which the quantity actually measured is some
more or less complicated function of the parameter that actually required. Inverse theory refers to the
inversion of this function, thus, inverse theory is an organized set of mathematical techniques for reducing
data to obtain useful information about the physical world on the basis of inferences drawn from observations.

Further author information –
M.J.R. (correspondence): Email: mjr@bimg5.kfunigraz.ac.at; Telephone: +43-316-380-5262
G.K.: Email: gottfried.kirchengast@kfunigraz.ac.at
IMG/UoG: WWW: http://www.kfunigraz.ac.at/imgwww; Fax: +43-316-380 9825

449
The term "inversion" is used in contrast to "forward problem", which is defined as the process of predicting the results of measurements on the basis of some general principles or models under a set of specific conditions relevant to the problem at hand. Inversion, roughly speaking, addresses the reverse problem: starting with data and a general principle (determined by the physics of the experiment), it is to estimate model parameters (states).

Let us consider the measurements assembled into a vector \( y = (y_1, y_2, \ldots, y_m)^T \), the measurement vector, and the unknowns (often referred to as the model parameters, that are to be retrieved) assembled into the state vector \( x = (x_1, x_2, \ldots, x_n)^T \),

describing the state of the system. Regarding the notation in the following, lowercase boldface letters denote column vectors (e.g., \( x \)), and uppercase boldface letters denote matrices (e.g., \( K \)); the symbol \(^T\) denotes a vector or matrix transpose (e.g., \( x^T \)), and the symbol \(^{-1}\) denotes a matrix inverse (e.g., \( K^{-1} \)).

The state vector, which assembles the quantities to be retrieved, may (in principle) comprise any set of relevant variables suitable to describe the specific problem. The measurement vector should include all of the measured quantities that are functions of the state vector. Measurements are made to a finite accuracy; random error or "measurement noise" will be denoted by the vector \( \varepsilon \).

For each state vector there is a corresponding measurement vector, \( y \), determined by the physical laws. We can describe the physics formally as the forward model \( f \). Therefore, we will write the relationship between the measurement vector and the state vector as

\[
y = f(x) + \varepsilon. \tag{1}
\]

Let us consider a problem where the forward model is a linear function of the state; if this is not the case by nature, a linear problem can be obtained by linearization of the forward model about some reference state \( x_0 \). For this purpose we rewrite Eq.(1) as follows:

\[
y - f(x_0) = \partial f / \partial x (x-x_0) + \varepsilon = K(x-x_0) + \varepsilon, \quad \text{with} \quad K_{ij} = \partial f_i / \partial x_j \tag{2}
\]

This defines the \( m \times n \) weighting function matrix, \( K \), assumed to consist of elements \( K_{ij} \), the (row) index \( i \) ranging from 1 to \( m \), the (column) index \( j \) ranging from 1 to \( n \). If the problem were linear by nature, the form given by Eq.(2) would directly apply with \( K \) independent of the state \( x \).

### 2. THE BAYESIAN APPROACH

The Bayesian approach provides a framework within which one can understand the inverse problem. Given a measurement together with a description of its error statistics, a forward model describing the relation between the measurement and the unknown state, and any kind of a priori information that might be available, it allows one to assign a probability density function (pdf) to the possible states, consistent with the measurement and the a priori.

Following Rodgers [1976], the theorem of Bayes can be used to derive a scalar cost functional \( \chi^2 \), which enables to describe the inverse problem as an optimization problem of finding the minimum of this scalar functional,

\[
\chi^2 = (y-f(x))^T S^{-1}_\varepsilon (y-f(x)) + (x-x_{ap})^T S^{-1}_{ap} (x-x_{ap}) = \text{Min}. \tag{3}
\]

The above equation, Eq.(3), has been obtained using Gaussian-shaped pdf's as measurement and a priori error distributions. This is a good assumption in the former case (\( S_\varepsilon \)), as we nominally deal with unbiased (non-systematic, random) measurement errors. In the latter case (\( S_{ap} \)) it is valid only in some cases. Many problems are rather poorly approximated by a Gaussian a priori distribution; this may (mis-)lead to estimated states, which are influenced by the a priori state in an inappropriate manner.
For a linear problem, using Eq.(2), the term \( f(x) \) is substituted by \( Kx \) and it is possible to construct an explicit formula for the retrieved state \( x_{\text{retr}} \). Using Eq.(3) and minimizing \( \chi^2 \) with respect to \( x \) yields, for the state \( x \), an equation for optimal estimation of the state:

\[
x_{\text{retr}} = x_{\text{ap}} + (K^T S_e^{-1} K + S_{\text{ap}}^{-1})^{-1} K^T S_e^{-1} (y - K x_{\text{ap}}).
\]  

(4)

3. DEALING WITH NONLINEARITY

The term nonlinearity denotes that the forward model that maps state space into measurement space is a nonlinear function of the state. Several inversion problems are intrinsically nonlinear, it is only by making simplifying assumptions (like, e.g., linearization) that we can construct a linear model. Considering a moderately nonlinear case, the linear approach does no longer do a good job. New principles of finding an optimally estimated solution have to be invoked, some of them are represented in this section. In general, the Bayesian approach by no means is restricted to the linearity of the forward model, although some of the specific formulae of optimal estimation are strictly valid only in the case of a linear forward model (cf., e.g., Jackson [1985]).

The Bayesian solution for the linear problem can be modified straightforwardly for the purpose of an inverse problem in which the forward model is a general function of the state, the measurement error is Gaussian, and there is a prior estimate available obeying Gaussian statistics. The Newtonian iteration can be viewed as a proper algorithm for such cases, which uses a (truncated) Taylor expansion up to the first term in the neighborhood of the desired state [cf., e.g., Press et al., 1992; cf. also Rieder, 1998]. The algebraic calculation yields an implicit equation for \( x_{\text{retr}} \) (and \( S_{\text{retr}} \)), if the nonlinearities are not too pronounced (approximation of the forward model \( f \) by a linear term).

The Newtonian iteration method finds the minimum of an exactly quadratic cost function \( \chi^2 \) in one step. However, if the solution is sufficiently far from the state found for a given iteration step it is quite possible that the quadratic assumption represents the surface so poorly that a next iteration step taken is almost meaningless, and may even increase rather than decrease the residual [Rodgers, 1998]. Under these circumstances a steepest descent approach is at least safe, even if agonizingly slow. This gradient method simply steps off from the current trial value in the direction of the negative gradient of \( \chi^2 \). It is necessary to control the step size carefully once the direction of the gradient vector has been established. In a case where the nonlinearities are too gross, and therefore more than one minimum exists in the hypersurface \( \chi^2 \), severe problems arise.

The Marquardt-Levenberg algorithm [Marquardt, 1963] is a way of combining steepest descent and Newtonian iteration as an optimum interpolation by

\[
x(i+1) = x(i) - (V^2 \chi^2(i) + \gamma I)^{-1} V \chi^2(i),
\]  

(5)

where \( I \) denotes the identity matrix in \( n \) dimensions.

If \( \gamma \to 0 \), this tends to the Newtonian iteration, if \( \gamma \to \infty \), it tends to the steepest descent algorithm with a small step size. Thus \( \gamma \) can be chosen to give steepest descent far from the solution, and Newtonian iteration near the solution.

In the notation of optimal estimation, the Marquardt-Levenberg algorithm reads:

\[
x_{\text{retr}}(i+1) = x(i) + (K_i^T S_e^{-1} K_i + S_{\text{ap}}^{-1} + \gamma I)^{-1} (K_i^T S_e^{-1} (y - f(x_i)) - S_{\text{ap}}^{-1} (x(i) - x_{\text{ap}})).
\]  

(6)

The algorithm shares with the gradient methods their ability to converge from an initial guess which may be outside the region of convergence of other methods because of the nonlinearity. The algorithm shares with the Taylor series method the ability to close in on a converged value rapidly after the vicinity of the
converged value has been reached. Thus, the method combines the best features of its predecessors while avoiding their most serious limitations [Marquardt, 1963].

4. MONTE CARLO METHODS

There is another - rather simple - but computationally inefficient way of solving the inverse problem. This is, indeed, the attempt to minimize the cost functional (defined by Eq. (3)) by a random or quasi-random examination of state space. Starting off at a given state, a sequence of trials is generated. In each trial a state is selected from the neighborhood of the current state. If this neighboring state has a lower cost, the current state is replaced by this neighbor, otherwise another neighbor is selected and compared for its cost value. The algorithm terminates when a state is obtained whose cost is no worse than any of its neighbors, and if this criterion is computationally not effective, the algorithm terminates when time runs out. The convergence of such an iterative improvement is obvious, the run time is proportional to the number of trials and thus it is important to concentrate on a sophisticated definition of „neighborhood“.

A handful of guiding principles can be readily understood:

- The starting state (initial guess) of a Monte Carlo algorithm should be motivated from either physical reasoning or a Bayesian point of view. As the infinite members of state space are theoretically all equal before calculating the cost, one should know a priori something on the structure of the probability density of state space.

- A neighborhood region in form of a (Gaussian-weighted) volume in n-dimensional state space is practicable, because of the choice of a starting state that is thought to be not too far from the requested minimum of the cost functional.

- As a transition rule from one state to the next, several empirical and/or computational advises can be used, depending on the particular problem. For example, a sort of covariance can be introduced, in order to correlate certain elements of the state vector. This is more or less the same as the use of a filter or the application of a smoothing function on the states.

- The run time of the algorithm could be controlled using statistical tests, such as a $\chi^2$ test, in order to check for significant improvement of the cost value.

- The cost functional by no means has to be fixed. In some applications it could be advantageous to modify this cost functional during the iteration, in order to more efficiently approach a certain precision limit.

A Monte Carlo iterative improvement algorithm does neither need any theoretical input of inversion theory nor a computation of the gradient of the cost functional (cf. Marchuk et al. [1980]). Thus it is not a physical method, in the sense of constructing a kind of inverse mapping from measurement space into state space, it is a pure statistical method (more or less a brute force method) to minimize a scalar measure of the fit between measurement and state. The algorithm is essentially no more than an iterative application of the forward model, it is blind of recognizing the direction of minimization of the cost function value, hence we have to expect high computer burden and many steps of computation, which a posteriori will turn out to have been useless.

5. SIMULATED ANNEALING

Simulated Annealing is another kind of statistical method, a more smart brother of the plain iterative improvement Monte Carlo method outlined above, which is of particular interest because it furnishes capability of coping with implicit nonlinearities in an optimization problem. The problem of nonlinearities can be understood when we think of the cost functional in a geometrical interpretation. In the linear case, the cost functional $\chi^2$ is a multi-dimensional hypersurface of parabolic shape. This surface clearly has one global minimum and it can be readily found applying standard methods. In the nonlinear
case, this hypersurface might be distorted and thus it is conceivable that there is more than one minimum; a situation, we refer to as the existence of multiple (local) minima.

The simulated annealing algorithm originates from the analogy between two problems: that of finding the ground state of a solid and that of finding a globally minimal configuration in a combinatorial optimization problem. In condensed matter physics, annealing denotes a physical process by which, if carried out sufficiently slowly, the ground state of a solid can be found. The simulated annealing algorithm is based on a simulation of this process [e.g., Leathoven, 1996].

Given the current state, a small, randomly generated, perturbation is applied. If the perturbation results in a lower energy state, the process is continued with the new state. If the difference in the energy before and after the state transition, \( \Delta E \), is greater than zero, then the probability of acceptance of the perturbed state is proportional to the Boltzmann factor, namely \( \exp(-\Delta E/kT) \). This rule for accepting new states is a criterion that furnishes capability of escaping local minima by temporarily allowing, contrary to strict minimization, also for cost increases in the course of the iterations.

We employed this kind of an annealing technique in that we carried out the simulated annealing more than one time (see Rieder [1998] for details). We started with a state, applied the simulated annealing technique following a specific cooling schedule, and then stored the cost function value of the retrieved state. Next we simulated a „heat up“ of the state, in that we applied some perturbation to the best estimate of all previous steps. Thus we could escape from a spurious minimum. Subsequently, the annealing algorithm started again. The best \( \chi^2 \) value of all these annealings was taken to be the one indicating the best estimated state.

6. POTENTIAL OF A HYBRID ALGORITHM

Having described the properties of an iterative approach (Marquardt-Levenberg algorithm) to minimize the cost functional, and having described the characteristics of a simple Monte Carlo and a simulated annealing search, we now want to combine these tools to a hybrid algorithm. Such an algorithm shall make use of the strengths and advantages of each specific tool, while at the same time overcoming their specific shortcomings.

The hybrid algorithm has to

- find a proper a priori; this is best accomplished by a search in a catalogue of representative states [cf. Rieder and Kirchengast, 1998].
- assure that we go into the correct direction when starting at some a priori state; this is best accomplished by a „physical“ part in the algorithm. Utilization of the Marquardt-Levenberg iteration well serves this need.
- further improve the „physical“ solution by a technique which furnishes capability of escaping some local minima of the cost functional; this is best accomplished by temporarily allowing increased cost values during an iteration. Utilization of the (multiple) simulated annealing well serves this need.
- improve the solution by releasing its strong resemblance of the a priori profile; this is best accomplished by modifying the cost functional in terms of changing the weight of the a priori. Utilization of a Monte Carlo technique (like iterative improvement/a priori lightweighting) well serves this need.

The above described algorithms all together, are organized as follows into a hybrid algorithm. The algorithm starts with the search for a suitable a priori state out of a catalogue, containing representative states. The best states, judged in terms of their cost function value, are stored. Five different a priori profiles were found to be enough to ensure that the correctly retrieved state can be found, finally. The following steps are done for each of the five different a priori states.
After the a priori has been found, the Marquardt-Levenberg iteration algorithm begins (cf. section 3). Once the convergence criterion for this algorithm is met, the simulated annealing technique in its multiple version (cf. section 5) is applied, starting with the Marquardt-Levenberg solution. When the simulated annealing algorithm ends, a modified cost function (see Eq.(7) and its description below) is established, and an iterative improvement/a priori lightweighting Monte Carlo method begins (cf. sections 4 and 7). After examining the neighboring states for an improved cost (judged by the modified cost function value), a $\chi^2$ test for correct convergence is done. Hereafter, the algorithm is run starting with the next a priori.

Finally, all correct retrieved states are compared, and the best of the results of all correct states is selected as the best estimate $\mathbf{x}_{\text{ret}}$.

One might ask why we did not prefer to conclude the hybrid algorithm by a gradient method, since we can reasonably believe to be already near the minimum, and the cost functional surely could be approximated by a linearized version. The reason is the numerical insignificance of the gradient (rendering a gradient method useless), since the hypersurface often is experienced to be extremely flat near the minimum.

7. AN ALGORITHM FOR WATER VAPOR PROFILING

Profiling constituents of the atmosphere using spaceborne downlooking sounder data is a typical geophysical example of a nonlinear optimization problem (see, e.g., Wang et al. [1983], Ulaby et al. [1986]). We used the Special Sensor Microwave Humidity Sounder (SSM/T-2) data for the purpose of water vapor profiling (brightness temperatures around the 183.31 GHz water vapor absorption line, see Littlejohn [1995] for a description of the sensor). As described in Rieder [1998], the weighting functions of the sensor SSM/T-2 are considerably dependent on the actual state, thus we have to deal with a nonlinear mapping from measurement space into state space.

THE REPRESENTATION

The state vector has been chosen to represent relative humidities (in short: r.h.). We represent the state using eight height layers, spaced 2 km in vertical (2 km to 16 km height). The measurement vector $\mathbf{y}$ comprises the brightness temperatures and the total precipitable water content (TPW); thus we denote the measurements to be $\mathbf{y} = (T_{b1}, T_{b2}, ..., \text{TPW})^\top$. Subsequently, the forward operator $\mathbf{K}$ has to consist of the weighting functions and, in the last row, an integration over the actual state in order to obtain the total precipitable water.

CHOICE OF AN A PRIORI

Since the a priori contributes a considerable fraction to the value of the cost functional, a careful choice is of major importance for the profiling algorithm. We used a climatological sample of representative profiles and a simple library method (for details see Rieder and Kirchengast [1998]) to find an a priori that is climatological well-founded as well as statistically reasonable (in terms of a moderate cost value) and adapted to the specific case at hand. A subset of the TIGR data set, the TIGR/IASI data set [Chédin et al., 1985], which assembles profiles assigned to some geographical and temporal conditions, was taken to provide the a priori state $\mathbf{x}_{\text{ap}}$. This a priori state served as the initial state for the algorithm.

THE MARQUARDT-LEVENBERG ITERATION ALGORITHM

We employed the Marquardt-Levenberg iteration to find the minimum of the cost functional in a mathematically rigorous way. The iteration procedure is demonstrated using synthetic data cases, where we know the true profile; and we want to estimate the retrieved profile based on the corresponding (simulated) measurements. The calculated measurements were perturbed by white Gaussian noise compliant with the variance specified within the measurement error covariance matrix $\mathbf{S}_e$. 

454
Figure 1 shows the a priori profile, the true profile, and the retrieved profile at each step of iteration (for 3 iterations). Furthermore, we present the actual differences between the a priori and the retrieved profile, respectively, and the true profile. Furthermore, the value of the $\chi^2$ cost function is presented in the panels, together with the actual step of iteration. The true profiles are some typical examples of U.S. Standard Atmosphere profiles (see Bilitza [1992] and the references therein), the a priori profiles are humidity profiles from the TIGR/IASI data set [Chédin et al., 1985]. The a priori profile in this example has not been chosen to fit the measurement that accurate, the a priori profile was the best estimate after employing a library method of scanning all stored profiles (with their corresponding measurement vector) to match the observed measurement. It thus represents a typical example of what happened when starting the algorithm.

Figure 1: Example for the Marquardt-Levenberg iteration.
MONTE CARLO ITERATIVE IMPROVEMENT/A PRIORI LIGHTWEIGHTING

The improvement of states is examined in terms of the cost functional. We defined a modified cost functional, in order to circumvent the problem of arriving at states that resemble the a priori state too much (cf. Fig. 1). The modified cost functional simply is a $\chi^2$ functional that gives less stress to the term that measures the difference between the estimated state and the a priori state. Thus we obtain

$$\chi^2 = (y - f(x_{\text{retr}}))^T S_e^{-1} (y - f(x_{\text{retr}})) + c_r (x_{\text{retr}} - x_{\text{ap}})^T S_{ap}^{-1} (x_{\text{retr}} - x_{\text{ap}}).$$  \hspace{1cm} (7)

The modified cost function value is a linear combination of two quadratic measures of deviations, and as we want to fit the measurement first and foremost, we set the regularization factor $c_r < 1$. A key idea for practically applying Eq. (7) was to alter $c_r$ throughout the iteration process. We started the Monte Carlo search with a regularizing factor of about 0.9, in order to avoid transitions to very improbable states. Subsequently, the factor has been decreased to approach 0.3 at the end of the algorithm. Note that one needs to normalize the cost function with respect to the regularization factor in order to obtain a comparable measure for each particular case of $\chi^2$ with different factors $c_r$. We termed this method also a priori lightweighting as it increasingly releases the resemblance of the a priori profile in the solution.

Figure 2: Example for the Monte Carlo iterative improvement/a priori lightweighting technique.
The new states are computed by a random perturbation of the best state so far, whereby the perturbation is applied within a Gaussian-weighted neighborhood region of the current state. The new state is accepted if the cost function value of the modified cost function is less than the one of its predecessor.

The above example, Fig.2, demonstrates the use of this Monte Carlo iterative improvement/a priori weighting technique. Figure 2 illustrates the profile at the beginning of the algorithm (rh_{wp}), and, from left to right, the panels give the subsequently improved (in terms of the cost function) profiles, together with the corresponding $\chi^2$ values of the cost function (multiplied by a factor of 100 for numerical convenience).

SIMULATED ANNEALING

We tested the simulated annealing algorithm in order to demonstrate its usefulness within a cost function topology in state space which is rather flat. We give one example of the performance of the simulated annealing algorithm in Fig.3.

![Simulated Annealing Graph](image)

*Figure 3: Escape from a spurious minimum. All panels show the a priori profile (rh_{wp}), the true profile (rh_{ref}), and the difference between the a priori and the true profile. The solution after a certain step (step 5 and 6) is the dashed-dotted profile rh_{ref}.*

Figure 3 demonstrates that the simulated annealing algorithm is capable of escaping from a local minimum. The simulated annealing algorithm takes a step leading to an increased cost value, followed by a step of an improved fit, which significantly surpasses the fit of the previous steps. The middle panel shows the case where a profile clearly is less well fitted to the measurements after some iteration step (step 5), as the cost function value indicates. Nevertheless, the transition was accepted and we end at a better state (after step 6), finally. This is exactly what is meant when stating: escape from a local minimum. Of course, also in the Monte Carlo iterative improvement technique described above, the transition to this final state could have happened; in a case, where the random perturbation happened to change the profile from the state at an earlier iteration step directly to the state of iteration step 6. In practice, however, the probability of such an occurrence is very low. It has been experienced, that the property of accepting also steps of increasing cost function values very much increases the probability of finding final states at lower costs (like the one in Fig.3, right panel) more efficiently.
8. THE HYBRID ALGORITHM – EXEMPLARY RESULT

The below example, Fig. 4, shows a typical case as it occurs during the routine application of the hybrid algorithm outlined in section 6 above (see Rieder [1998] for details). The pair of plots illustrates a retrieval over sea, where we used all 5 channels of SSM/T-2 to estimate the humidity profile.

![Figure 4: Example of a retrieval employing the hybrid algorithm. The value for the χ² cost function is denoted in the middle of the panels. The left panel shows the setup of the problem after the library search (selection of an a priori), the right panel shows the finally achieved solution (best estimate). We note that the retrieved profile resembles the a priori profile, especially in those regions, where the contribution of the measurements are weak (which can be seen, when a characterization of the measurement system, following the guideline of Rodgers [1990] is performed, see Rieder [1998]).](image)

9. VALIDATION OF RETRIEVALS WITH ECMWF DATA

The European Centre for Medium-Range Weather Forecasts (ECMWF) provides a global data set ("atmospheric analyses") containing the most important atmospheric parameters as quantities on a (temporal and) spatial grid. The parameters are joint estimates based on both measurements and weather forecast model results, obtained by an assimilation model (cf., e.g., Technical Attachment of ECMWF [1994]). The (currently) latest version of the data set is organized on a grid of 31 height levels (at "model resolution"), corresponding to pressure levels from surface pressure to about 10 mbar, and a latitudinal and longitudinal grid of about 1 x 1 deg reasonably matches the horizontal resolution of the model. The ECMWF analyses are established every 6 hours (beginning at midnight, four times a day), using various types of measurements (radiosonde observations, satellite data, ship and aircraft reports) consistently together with a global forecast model.

We used the ECMWF data set for the purpose of having a global reference state of the atmospheric parameters temperature (used as auxiliary input) and humidity. The humidity data set served as a "ground truth" state to which the SSM/T-2 retrieved profiles were compared for validation. Regarding coincidence in time, we selected one SSM/T-2 orbit arc from 15:15 UT to 17:05 UT on Oct 20, 1995; the 18:00 hrs ECMWF analysis of that day was then used. Regarding collocation, we searched for the nearest neighbor point of the ECMWF grid to the instantaneous field-of-view of SSM/T-2, which is sufficient due to the horizontal sampling density of the ECMWF data.
We checked the suitability of using ECMWF data for the purpose of a validation of SSM/T-2 retrievals by computing the ECMWF-derived brightness temperatures in the forward direction first, and comparing these brightness temperatures to the measured SSM/T-2 quantities. In general, the forward model, started with the atmospheric parameters given by the ECMWF data set as an input, yields good agreements of the measured SSM/T-2 brightness temperatures and the calculated brightness temperatures, at least in the cases over sea. Over land and over ice we have to exclude channels 4 and 5, as we have no proper knowledge of the surface parameters. Besides, in both cases, we have to adjust the cloud and precipitation parameters of the utilized forward model. MWMOD [Simmer, 1993], employing (well known) estimations based on the measured brightness temperatures, according to the rules described in the papers of Fleming et al. [1991] and Muller et al. [1994], respectively.

As the measured data in most cases is somewhat contaminated with radiation modulated by precipitation and clouds, we have to conclude that the differences in the forward direction could be reduced by a cloud clearing algorithm or a proper estimation of the most influencing parameters, which could be provided by other sensors. Under clear sky conditions (which were assessed when looking at the ECMWF data), the calculated brightness temperatures met the SSM/T-2 measurements very accurately. Over sea, we calculated unbiased differences with a standard deviation of 1.6 K brightness temperature for the first three channels and of 2.1 K for the window channels 4 and 5, respectively.

Above ice and above land further input on the surface radiation were needed in order to perform retrievals routinely. For the purpose of demonstrating some examples we estimated the necessary parameters corresponding to the average ice/land surface properties at the particular locations. Nevertheless, because of the weak estimate, we got some bias in the channels 4 and 5 of 1.4 and 2.2 K brightness temperature, respectively, together with rather high standard deviations. The (atmospheric) channels 1, 2, and 3 remained unbiased and showed a variance of 2 K brightness temperature. The fact of unknown surface properties and the differences in the forward calculated brightness temperatures in channels 4 and 5 led us to the decision to drop the window channels in the ice/land retrievals for this initial validation. We want to stress that an adequate land surface as well as a suitable ice model, together with some measurements of the actual ice/land parameters based on satellite data from other sensors, should be capable of providing the needed additional information, which then in fact would allow proper inclusion of channels 4 and 5 over ice/land, too.

For the purpose of the retrieval, also the total precipitable water content, TPW, has to be known. We assumed that the TPW might be known to an accuracy of 10 % of its absolute value from the retrieval based on another data source (e.g., another sensor). As we did not acquire such other data for this work, we calculated the TPW based on the ECMWF data, and applied a random perturbation to the values corresponding to a variance of 20 %, which is a rather conservative estimation.

**Comparison of Individual Humidity Profiles**

We performed a series of specific individual retrievals in order to demonstrate the properties of the hybrid algorithm used. Note that there is always some contamination of the measured data by clouds, precipitation and surface effects. In cases where we knew these parameters exactly, we could obtain more accurate retrievals, as the synthetic data examples indicate.

The two cases presented (Fig.5) are retrievals over sea, whereby case no. 1 is a retrieval under rather dry atmospheric conditions. Case no. 2 is a retrieval under moist atmospheric conditions (which was checked by looking at the ECMWF data plots). The particular locations are denoted in the middle of the panels, „lat“ denoting the latitude and „lon“ the longitude. All retrievals have been done with samples of the selected SSM/T-2 orbit arc segment of Oct 20, 1995.
Figure 5: Retrievals over sea; dry atmospheric conditions (left), and moist atmospheric conditions (right), respectively.

The figures display the SSM/T-2 retrieved profile, the corresponding ECMWF profile, and the difference of both. Furthermore, we plotted the standard deviations of the retrieval as error bars at the associated height level. The standard deviations of the retrieval have been calculated by taking the square root of the diagonal elements of the retrieval covariance matrix. Note that the standard deviations, in general, are not the same at each height level; especially the region above 12 km is expected to give enlarged differences, since the sensor is not capable of providing useful information in this region (as one learns when performing the characterization and error analysis following Rodgers [1990] using the actual state as an input; see Rieder [1998] for details).

The retrieval in „dry“ air is a typical example for a clear sky retrieval over sea. The maximum difference between the SSM/T-2 retrieval and the collocated ECMWF profile is about 10 % r.h., the error bars show that there is no significant deviation from the expected differences.

The example of a profile retrieval over sea in rather moist conditions indicates the problems involved in cloud contamination. We had to introduce some estimations of the cloud coverage, which might be the reason for a moderate deterioration of the retrieval. Although the differences between the retrieval and the ECMWF profile are not significant, the maximum difference is at about 15 % r.h. for this particular case.

The retrieval over land turned out to be less accurate than those over sea. The reason of why it seems that the deviation of the retrieval is not statistically significant is the high expected variance of the differences. The maximum differences are at about 15 % r.h., where there is particularly some trouble at heights up to about 6 km due to the rather poorly known (i.e., only roughly estimated) surface parameters.

The retrieval over sea ice, compared to the ECMWF data, differs up to 20 % r.h., generally because of the same reasons as it was the case for the retrieval over land (in fact, it is even more difficult to correctly account for the influence of ice). Note that there is much effort to improve the radiative transfer models to cope with the problem of surface emission of ice, and that there already exists a possibility to retrieve the total precipitable water content (TPW) over the typically „dry“ polar regions [cf. Wang et al., 1992]. Such improvements could stabilize retrievals like the above one effectively.
10. STATISTICAL VALIDATION WITH ECMWF DATA

When trying to compare retrievals with ECMWF measurements, one has to include an error analysis of the ECMWF measurements. In general, the error of the ECMWF humidity profiles can be estimated to be about 15 % r.h. (cf., e.g., Technical Attachment of ECMWF [1994]). We restricted ourselves to plot the mean, the standard deviation, and the standard deviation of the mean of the vectors \((x_{\text{ret}} - x_{\text{ECMWF}})\).

Because of the different treatment of retrievals over ice/land and over sea we have to divide into these both regimes of retrievals when statistically evaluating the accuracy of the retrievals.

Retrievals Over Sea

Along a track of the orbit arc, ranging from 11 to -64 degrees in latitude, we used a sample of 63 retrievals available in order to calculate the mean difference, the standard deviation of the mean (heavy short bars) and the standard deviation (light longer bars) of the differences. We used all 5 channels of SSM/T-2 to perform the retrievals over sea and weighted the window channels 4 and 5 to be less informative than the three atmospheric channels, i.e., we set the corresponding standard deviations in the measurement covariance matrix \(S_x\) to 1.5 K (the standard deviations of the first three channels are 10 K).

The result gives significant biases at some heights. Similar to the cases with synthetic data, rather large standard deviations (mainly because of a typically large retrieval covariance matrix \(S_{\text{ret}}\)) can occur, too. At those heights, where the measurements are not capable of supplying much information to the retrieval, the standard deviations are larger [cf. Rieder, 1998].

![Graph showing retrieval comparison statistics over sea.]

In general, the retrieval appears accurate, according to Fig.6, at heights up to about 8 km. Besides the favorable contribution functions at these heights, this is because of the TPW input, which helps most at the moist lower height layers. This is also the reason why the error bars at these heights are smaller indicating more accuracy.

Figure 6: Statistical mean (dashed profile), standard deviation of mean (heavy short bars), and standard deviation (light longer bars) for the retrievals over sea (SSM/T-2 retrieved profiles minus ECMWF humidity profiles).

11. ACKNOWLEDGMENTS

Thanks are due to Prof. Dr. C. Simmer, Institut für Meereskunde, Kiel, who provided his radiative transfer model MWMOD utilized for this study. We also acknowledge the support of Noelle Scott and Marco Matriculari, LMD, France, for providing the TIGR/IASI atmospheric data set. Furthermore, we want to thank for the satellite data, assembled by the sensor SSM/T-2, used in this work; these data sets have been kindly provided by the National Oceanic and Atmospheric Administration (NOAA). The first author (M.J. Rieder) received financial support for this work from IMG/UoG’s “ATFERN“ discretionary funds.
12. REFERENCES

5. Laarhoven, P.J.M. van, Theoretical and computational aspects of simulated annealing, CWI Tract, CWI, 1996.
TECHNICAL PROCEEDINGS OF THE TENTH
INTERNATIONAL ATOVS STUDY CONFERENCE

Boulder, Colorado
27 January - 2 February 1999

Edited by
J. Le Marshall and J.D. Jasper
Bureau of Meteorology Research Centre, Melbourne, Australia

Published by
Bureau of Meteorology Research Centre
PO Box 1289K, GPO Melbourne, Vic., 3001, Australia

December 1999