EFFICIENT METHODS TO ASSIMILATE DATA FROM FUTURE REMOTE SOUNDING INSTRUMENTS

Joanna Joiner, Arlindo M. da Silva, and Richard Ménard
NASA/Goddard Data Assimilation Office (DAO)
Greenbelt, MD, USA

1. INTRODUCTION

Traditionally, methods for utilizing remotely sensed data in data assimilation systems (DAS) have involved assimilating retrieved products in the form of state variables (e.g., temperature and humidity). The direct assimilation of radiances has been shown to be a more optimal use of remotely-sensed data (e.g., Eyre et al., 1993) and this approach is currently operational at both NCEP (Derber and Wu, 1996) and ECMWF (Courtier et al., 1997; Rabier et al., 1997; Andersson et al., 1997). Although it is computationally feasible to implement direct radience assimilation with TOVS and ATOVS the cost of this approach for future high-spectral-resolution sounding instruments such as IASI and AIRS, which have orders of magnitude more spectral elements, may be prohibitive.

As an alternative to direct radience assimilation and traditional retrieval assimilation, we have developed two methods for assimilating remotely sensed data with an appropriate state-dependent error covariance derived from first principles. The framework combines a characterization of retrieval errors with statistical analysis equations leading to a consistent assimilation of the data. This implementation may be considerably less expensive than direct radience assimilation. An outline of the approach is given here, but more details will appear in Joiner and da Silva (1997).

2. CONSISTENT ASSIMILATION OF REMOTELY-SENSED DATA

2.1. Linear Statistical Analysis

In the general case of assimilating retrieved data, the statistical interpolation scheme that attempts to obtain an optimal estimate of the atmospheric state is linear and is given by

\[
    w^a = w^f + \left( P^f \mathcal{H}^T - X^T \right) \left( \mathcal{H} P^f \mathcal{H}^T + R^z - \mathcal{H} X^T X \mathcal{H}^T \right)^{-1} \left( z - \mathcal{H} w^f \right),
\]

where \( w^a \in \mathbb{R}^n \) is the analysis vector, \( w^f \) is the background (forecast) vector, \( P^f \) is the background error covariance, \( R^z \) is the retrieval error covariance, \( z \in \mathbb{R}^p \) denotes the observation vector which in this case is retrieved data, \( X \) is the background-retrieval error cross-covariance (i.e., \( X = \langle e^z (e^f)^T \rangle \)), and \( \mathcal{H} \in \mathbb{R}^n \times \mathbb{R}^p \) is the linearized observation operator which in the case of retrieval assimilation is just a linear interpolation operator. The Physical-space Statistical Analysis
System (PSAS), currently implemented at Goddard’s Data Assimilation Office (DAO) solves this system using a conjugate gradient algorithm (e.g., da Silva et al., 1995). A fully non-linear version of (1) can be written for radiance assimilation in which the observation operator is the radiative transfer equation.

The quantity to be inverted in (1) is of dimension \( p \times p \) where \( p \) is the dimension of the observation vector. In the design of PSAS it is assumed that the dimension of the observation vector \( p \) is smaller than the dimension of the state (analysis or background) vector \( n \). For high-resolution sounders, it is easy to see that the vector of radiance observations (of the order of thousands of radiances observations per grid box) can be quite a bit larger than \( n \) (of the order of 100 states per grid box). However, if the state is not completely observed, the observations can be compressed into pseudo-observations such that \( p < n \). There are several available methods of data compression. Perhaps the most simple method is to reduce the number of channels used for assimilation. However, this would amount to discarding some of the information content of the measurements. A second form of data compression for high-resolution sounders is a traditional 1D retrieval. However, as we will show below, this is also a suboptimal form of data compression for data assimilation. A third option is to perform a linear transformation on the radiances. This transformation can be from radiance space to physical space (essentially an unconstrained retrieval) or can be performed in channel-space (i.e., pseudo-channels). Two implementations of the third approach will be given below.

2.2. General Assimilation of Retrievals

The retrieval process can be written as a nonlinear estimation, i.e.,

\[
z = D(y, b, z^p) = D(f[z^t, b^t] + e^y, z^p),
\]

where \( D \) is the inverse function, \( y \) is a vector of radiance measurements, \( b \) is a vector of parameters needed in a forward model, the superscripts \( p \) and \( t \) refer to the prior estimate of the state (background) used in the retrieval process and the true state, respectively, \( e^y \) will be referred to as the observation error and includes contributions from forward model error, detector noise, and linearization error, and \( f \) is the forward model function or observation operator. The retrieval error, \( e^z = z - z^t \), (neglecting bias in the inverse problem) can be written to a linear approximation as

\[
e^z = (I - A) e^p + D_y e^y
\]

(Eyre, 1989; Rodgers, 1990) where \( e^p = z^p - z^t \) is the prior error, and

\[
D_y = \frac{\partial D}{\partial y}\bigg|_{z=z^t}.
\]

\( A = D_y F_z \) is often referred to as the averaging kernel, where

\[
F_z = \frac{\partial f}{\partial z}\bigg|_{z=z^t},
\]
where $i$ refers to the iteration of the retrieval. The first term on the RHS of (3) may be thought of as a smoothing error that results from the use of prior data in components of the state space that the observing system is not able to accurately measure. Although in some instances it may be possible to model $e^p$, practical implementations will likely rely on statistical modeling from innovation sequences (e.g., da Silva et al., 1996) or some form of online parameter estimation. For simplicity, we will assume that the prior and observation errors are unbiased, i.e., $<e^p>=<e^y>=0$.

For over-determined retrieval problems where the effect of prior data is negligible, $A \approx I$ so that $X = (I - A) \langle e^p(e^f)^T \rangle \approx 0$. In this case, the implementation of (1) is greatly simplified. One example of a potentially overdetermined inverse problem occurs in the retrieval of total ozone from backscatter ultraviolet instruments.

For the under-determined case in which the prior information does not come from the data assimilation system forecast, one is left with the almost impossible task of modeling the forecast-retrieval error cross-covariance. Even if the cross-covariance can be modeled, it can be shown that (1) cannot be implemented in this case owing to numerical instability. Next we present two practical methods that attempt to eliminate the effect of prior data in order to simplify the implementation of (1).

2.3. Assimilation of Null-space Filtered (NSF) Retrievals

In order to simplify the error characterization for data assimilation, we seek a linear transformation of the retrieval that eliminates or renders negligibly small the contribution from the prior state $z^p$. This can be accomplished by projecting the retrievals $z$ on the null-space of the operator $I - A$. Specifically, consider the singular-value decomposition

$$I - A = U S V^T,$$

where $U/V$ are matrices of left/right singular vectors of $I - A$, and $S$ is a diagonal matrix of singular values sorted in descending order. Let $U = (U_L | U_T)$, where the columns of $U_L/U_T$ are composed of leading/trailing left singular-vectors. We take $U_T$ to correspond to negligibly small singular-values $S_T$, so that

$$U_T^T(I - A)z^p = S_T V_T^T z^p \approx 0$$

and $X \approx 0$. We will refer to the subspace spanned by the leading modes $U_L$ as the null-space of the observations. Although not strictly a null-space in the mathematical sense, the null-space can be thought of as components of the state-space that the instrument cannot measure accurately.

It follows that the error of the new data type $\beta = U_T^T z$ has a negligible contribution from the prior error, i.e.,

$$e^\beta = U_T^T e^z \approx U_T^T D_p e^y.$$
The data type \( \beta \) can be assimilated using
\[
w^a = w^l + P^l H^T U_T (U_T^T H P^l H^T U_T + U_T^T D_y R^y D_y^T U_T)^{-1} (\beta - U_T^T H w^l),
\]
where \( R^y \) is the radiance error covariance. One potential advantage of this approach is that it does not explicitly require the radiances. It does however optimally require knowledge about how the retrieval is performed. The disadvantage of this approach is that it may not completely eliminate all of the prior information. In addition, it may actually filter some of the useful information contained in the radiances. A more optimal approach requiring radiances is given below.

2.4. Transformation of Radiances using Partial Eigendecomposition (PED)

In order to completely remove the effects of prior data, the radiances may be transformed to physical space such that \( A = I \) and therefore \( X = 0 \). Consider an unconstrained weighted least-squares retrieval linearized about an initial state \( z^p \), i.e.,
\[
z = z^p + (F_z^T (R^y)^{-1} F_z)^{-1} F_z^T (R^y)^{-1} (y - f(z^p, b)).
\]
In this case, \( D_y \) is given by
\[
D_y = (F_z^T (R^y)^{-1} F_z)^{-1} F_z^T (R^y)^{-1},
\]
and \( A = D_y F_z = I \). Substituting (11) into (3), the retrieval error covariance, \( R^z \), is given by
\[
R^z = D_y R^y D_y^T = (F_z^T (R^y)^{-1} F_z)^{-1}.
\]
If \( [F_z^T (R^y)^{-1} F_z] = (R^z)^{-1} \) is singular or nearly singular, it is possible to redefine the retrieval vector redefined in terms of the well-determined eigenvectors of \( (R^z)^{-1} \).

Because \( (R^z)^{-1} \) is a symmetric, positive semi-definite real matrix, we may write
\[
[(F_z^T (R^y)^{-1} F_z)] = (R^z)^{-1} = U \Lambda U^T,
\]
where \( U \) is unitary and orthogonal \( (U U^T = I) \), and \( \Lambda \) is a diagonal matrix with eigenvalues that are real and non-negative sorted in descending order by convention. The columns of \( U \) (eigenvectors of \( (R^z)^{-1} \)) can be written as \( U = [U_L \mid U_T] \), where \( U_L / U_T \) are the leading/trailing modes. Eigenvectors in the null-space of \( (R^z)^{-1} \) will have zero eigenvalues, and modes that are not well-determined by the measurements \( (i.e., \ \text{the trailing modes}) \) will have small eigenvalues.

The new retrieval vector, \( \alpha_L \), defined in terms of the leading modes of \( (R^z)^{-1} \), can be computed by linearizing about an initial state \( z^p \), i.e.,
\[
\alpha_L = \alpha_L^p + \Lambda_L^{-1} U_L^T (F_z)^T (R^y)^{-1} [y - f(U \alpha_L^p, b)].
\]
The error covariance of \( \alpha_L \) is equal to \( \Lambda_L^{-1} \) and has no contribution from the prior state. The partial eigen-decomposition (PED) transformation minimizes the mean-square difference between
the exact and truncated-mode solution. The data type or pseudo-observation $\alpha_L$ can be assimilated in as in (9) with $U_L$ in place of $U_T$ and $\alpha$ in place of $\beta$. For practical implementation, it may be possible to define a stationary or limited set of fixed transformation operators $U$ for PED and NSF retrieval assimilation. In this case, the pseudo-observation error covariance may not be diagonal. The transformation to produce pseudo-channels (or linear combination of channels corresponding to the physical-space eigen-vectors in $U$) can be computed via the singular value decomposition of $(R^\alpha)^{-1/2}F_Z$.

In order to implement this approach, a prior estimate of the state is needed for linearization. This prior estimate could come from the background or alternatively from a 1D traditional retrieval. The use of a 1D retrieval will reduce linearization errors and is therefore more desirable. Because the subsequent assimilation of the PED retrieval is linear, this amounts to performing linearization in 1D which could lead to a significant reduction in computational cost as compared with the 3D linearization that is performed in radience assimilation, especially for PSAS.

It is also possible to perform an unconstrained retrieval without the transformation and truncation steps provided of course that the number of channels is larger than the number of states and that there are no zero eigen-values in (12). The unconstrained retrieval can be assimilated using the appropriate error covariance (even though the retrieval will look very strange owing to the ill-posed nature of the inverse problem). This approach would ensure that none of the information content is lost and also provides a significant data compression for AIRS and IASI.

3. 1-D SIMULATION FOR COMPARISON OF DATA ASSIMILATION METHODS

In order to compare the above methods with direct radience assimilation, we have performed a set of Monte Carlo simulations in 1 dimension. The details of the simulation are described in Joiner and da Silva (1997). The general results will be shown here. In this experiment, we simulated radiances for two different infrared sounders: the Atmospheric Infrared Sounder (AIRS) and HIRS. AIRS, scheduled to fly on the NASA Earth Observing System (EOS) PM platform in the year 2000, is a high-spectral resolution ($\nu/\Delta \nu = 1200$) grating spectrometer with approximately 2600 contiguous (with some gaps) spectral elements that cover wavelength regions from 3.4 to 15.4$\mu$m (650 to 2670 cm$^{-1}$). The spatial resolution of AIRS is about 13 km at nadir in the 705 km orbit of EOS PM. The specified single spot equivalent noise temperature for a given spectral resolution element is 0.2$^\circ$C at a 250K scene for channels $\nu > 750$ cm$^{-1}$ and 0.35$^\circ$C at a 250K scene for channels at lower frequencies. The AIRS channels set used here includes a total of 550 channels consisting of all channels between 650 and 742 cm$^{-1}$, between 2160 and 2270 cm$^{-1}$, and between 2379 and 2407 cm$^{-1}$. The HIRS channel set used here includes a total of 10 channels covering the first 2 AIRS wavelength regions (channels 1-7 and 13-15). We could have used all the AIRS or HIRS channels and included other state variables (i.e., humidity, surface temperature, and ozone) in the analysis. However, given the biases that have been observed in HIRS channels 11 and 12,
Figure 1: Expected thickness errors for the forecast and direct radian analysis for a low latitude profile.

we are not confident that the humidity channels can be modeled accurately enough to provide additional information about temperature in a multivariate analysis. Therefore, we have chosen this simplified and idealized example (where we have assumed perfect knowledge of the humidity, ozone, and surface temperature) for demonstration purposes only.

Figure 1 shows expected thickness errors for the forecast and the analysis using direct radian assimilation for a model-generated low latitude profile computed with the linear error estimation theory described above. The forecast error covariance ($P_f$) used here is the error covariance for a 6 hour forecast from the 18 level (0.4, 1, 2, 5, 10, 30, 50, 70, 100, 150, 200, 250, 300, 400, 500, 700, 850, and 1000 mb) Goddard Earth Observing System (GEOS) general circulation model. $P_f$ is generated statistically from innovation (observed minus forecast) height time series over North America. Horizontal components of $P_f$ are neglected here. The largest impact from the sounding data is in the stratosphere and lower troposphere, and significantly greater impact is shown for AIRS than for HIRS. $P_f$ is generated in the northern hemisphere where conventional data is dense. Over the southern hemisphere, where forecast errors are typically larger, the impact of AIRS and HIRS data will in general be greater than that shown here.

Figure 2 shows the first 10 eigenvectors and eigenvalues of $F_x^T (R_y)^{-1} F_x$ for AIRS temperature sensing channels using two different profiles: a tropical profile and an arctic profile. The modes for the low latitude profile have more vertical resolution in the lower troposphere than the modes of the high latitude profile. The error associated with a particular mode (inverse eigenvalue) is also lower for the low latitude profile than for the high latitude profile.

Figure 3 shows $J = \text{trace}(P_a)/\text{trace}(P_f)$ plotted as a function of number of modes included in the analysis for AIRS and HIRS for the low and high latitude profiles, where $P_a$ is the analysis error covariance. The trace is a measure of the vertically-averaged error variance. $J$ indicates the
Figure 2: Leading eigenvectors and eigenvalues of $F_z^T (Ry)^{-1} F_z$ for AIRS channels for a low latitude profile (solid curve, $\lambda_1$) and high latitude profile (dashed curve, $\lambda_2$).

reduction in the vertically-averaged error variance from the assimilation of a particular mode or set of modes that are assimilated. Approximately 8-10 modes provide most of the AIRS impact, while approximately 4 modes provide most of the HIRS impact. Again, the total AIRS impact on the analysis is seen to be significantly greater than the HIRS impact. For HIRS, approximately 4 modes are needed to provide most of the impact for both the low and high latitude profiles, but the impact is slightly greater for the high latitude profile as a result of the warmer stratospheric temperatures that provide slightly enhanced sensitivity. For AIRS, the total impact for both the low and high latitude profiles is about the same, but 2 fewer modes are required for the high latitude profile. It should be noted that additional information about humidity, surface temperature, clouds, emissivity, etc. will be provided by AIRS and IASI. If the humidity sensitive channels can be modeled accurately, they will provide additional information about temperature as well. Again, our purpose here is not to show exactly how many pieces of information are contained in the radiances, but to show that significant data compression of the radiances is possible.

Figure 4 shows the thickness errors obtained in the Monte Carlo simulation for the low latitude PED assimilation with 8, 9, and 10 modes retained. The direct radiance assimilation (same as 17 mode PED assimilation) is the same as performing a 1-D minimum variance interactive retrieval and is the optimal solution (provided that the linearization error has been correctly specified). The difference between the direct radiance assimilation and the PED retrieval assimilation is very small.
Figure 3: trace($P_f$)/trace($P_a$) as a function of number of modes included in the analysis.

Slight degradation is shown in the upper stratosphere when only 8-10 modes are assimilated, and the degradation increases slightly when the number of retained modes decreases. Similar results were obtained for the high latitude case with a somewhat different error structure. The errors shown here are slightly greater than those from expected from figure 1. It was verified that this effect results from incorrect specification of the linearization error.

Figure 5 is similar to figure 4, but shows the results for the NSF retrieval assimilation using different numbers of modes. Also shown for comparison are results from the direct radiance assimilation. The difference between the NSF retrieval assimilation and the direct radiance assimilation is insignificant, especially when 10 modes are retained. The slight degradation of the direct radiance solution in the stratosphere is the result of incorrect specification of the linearization error.

The above results show that 550 AIRS radiance observations can be compressed into approximately 10 pieces of information about the atmospheric temperature profile. This compression will result in a significant reduction in computation for the full 3-D PSAS implementation. The data compression for HIRS, although much smaller (approximately 10 channels to 4 pieces of information), still leads to a significant reduction in computation.

4. CONCLUSION

We have provided cost-effective methods for assimilating retrieved data from future high-spectral-resolution sounders such as AIRS and IASI. These methods attempt to limit the amount of prior information contained in retrieved data and provide a compact storage mechanism for
Figure 4: Thickness errors vs. pressure for analysis using direct radiance and PED retrieval assimilation with AIRS channels for the low latitude case.

Figure 5: Same as Figure 4 for NSF retrieval assimilation.
the information content of the measurements. This data compression will significantly reduce
the computation cost of assimilating data from AIRS and IASI, especially for the DAO’s PSAS
data assimilation system. In addition, computation is reduced by performing the linearization
in one dimension rather than in three dimensions. Finally, there are other potential advantages
of this approach. The radiative transfer calculations can be performed off-line. Therefore, with
an adequate first guess, the psuedo-observations can be computed once and for all. This would
significantly reduce costs when running several parallel experiment such as Observing Systems
Experiments (OSE) in which one or more data types are withheld from an analysis or when running
multiple-year re-analyses.

5. Acknowledgments

The authors are grateful to S. E. Cohn, D. P. Dee, and C. D. Rodgers for enlightening discus-
sions and helpful comments.

6. REFERENCES

Andersson, E., Haseler, J., Undén, P., Courtier, P., Kelly, G., Vasiljević, D., Branković, C., Cardinali, C.,
Gaffard, C., Hollingsworth, A., Jakob, C., Janssen, P., Klinker, E., Lanzinger, A., Miller, M., Rabier,
F., Simmons, A., Strauss, B., Thérault, J.-N., and P. Viterbo, 1997: The ECMWF implementation
of three dimensional variational assimilation (3D-Var). Part I: Experimental results, Submitted to

Courtier, P., Andersson, E., Heckley, W., Pailleux, J., Vasiljević, D., Hamrud, M., Hollingsworth, A.,
Rabier, F., and M. Fisher, 1997: The ECMWF implementation of three dimensional variational

Derber, J., and W.-S. Wu, 1996: The use of cloud-cleared radiances in the NCEP’s SSI analysis system,


119, 1427-1463.


ECMWF implementation of three dimensional variational assimilation (3D-Var). Part I: Structure

Rodgers, C. D., 1990: Characterization and error analysis of profiles retrieved from remote sounding mea-
surements. J. Geophys. Res. 95, 5587-5595.

da Silva, A. M., Pfendtner, J., Guo, J., Siennkiewicz, M., and S. E. Cohn, 1995: Assessing the effects of
data selection with DAO’s Physical-space Statistical Analysis System. International symposium on
assimilation of observations in meteorology and oceanography, Tokyo, Japan, 273-278.

TECHNICAL PROCEEDINGS OF
THE NINTH INTERNATIONAL TOVS STUDY CONFERENCE

Igls, Austria

20-26 February 1997

Edited by
J R Eyre
Meteorological Office, Bracknell, U.K.

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

May 1997