Estimates of short-range forecast-temperature error correlations and the implications for radiance-data assimilation

By A. P. MCNALLY*

European Centre for Medium-Range Weather Forecasts, UK

(Received 25 September 1998; revised 16 July 1999)

SUMMARY

The importance of an accurate specification of short-range forecast-error structures for the successful assimilation of satellite sounding data is discussed, with special emphasis on the role of vertical (inter-level) error correlations. Two different statistical approaches for estimating these correlations are examined. The first is based on comparing radiosonde profiles with the short-range forecast. The second compares forecasts of different lengths which verify at the same time. It will be shown that, in some aspects, the two approaches produce consistent results which suggest that significant geographic and seasonal variability exists in the vertical error correlations. The effect of not modelling this variability in the analysis of TOVS (TIROS-N (Television Infra-Red Observation Satellite) Operational Vertical Sounder) radiance data is simulated using a linear retrieval estimator, and it is shown that it may result in a serious misinterpretation of the information in the observations.

KEYWORDS: Data assimilation Satellite sounding

I. INTRODUCTION

The current network of meteorological observations is not sufficient to specify uniquely the initial conditions required for global numerical weather prediction (NWP). Analysis schemes such as optimal interpolation (OI) (Lorenc 1981) and variational methods (Lorenc 1988) combine observations with background information provided by a short-range forecast to produce an optimal estimate of the atmospheric state. However, this estimate is only optimal to the extent that the error characteristics of both the observations and the forecast background are correctly specified in the analysis. The general importance of an accurate specification of the background-error covariance is discussed in Daley (1991), but it has special relevance for the assimilation of satellite sounding data such as that obtained from the TIROS (Television Infra-Red Observation Satellite) Operational Vertical Sounder (TOVS) (Smith et al. 1979). The very poor vertical resolution of these data means that the analysis depends strongly on the vertical correlation of background temperature (and humidity) error to distribute the radiance information in the vertical. That is to say, there is a significant 'null space' in the radiance observations (vertical modes to which the data are not sensitive, Rodgers (1976)), such that a particular radiance signal suggests an infinite number of equally likely increments to the background field (in the absence of other data). It is the background and its corresponding error covariance which resolve this ambiguity. This is, of course, the case for any observation, but to a much smaller extent for data such as radiosonde measurements. These have a very high vertical resolution and the observed information is effectively given more weight in defining the vertical structure of increments in the analysis.

Radiosonde data have been used in a number of studies to estimate the error covariance of a short-range forecast background. A very simple approach described in Watts and McNally (1988) used the covariance of observation-minus-background departure statistics (sometimes called innovation statistics) as an estimate of the background-error covariance. While the radiosonde data were stringently quality controlled and

* Corresponding address: European Centre for Medium-Range Weather Forecasts, Shinfield Park, Reading, Berkshire RG2 9AX, UK.
local biases were removed, such an estimate inevitably contains a contribution from
the radiosonde error (as well as errors of representativeness). Attempts have been made
to partition observation error from background error in departure statistics by modelling
the contributions from each in terms of a small number of parameters (McNally 1990;
Dee and da Silva 1999; Dee et al. 1999). Values of these parameters that 'best fit' the
departure statistics are then estimated. The problem with such an approach is that the
choice of covariance parametrization is inevitably subjective, and limited to the num-
ber of parameters (or degrees of freedom) that can be independently estimated from
the data. A more successful partition of observation and background error in depart-
ure statistics is described in Hollingsworth and Lonnberg (1986) and Lonnberg and
Hollingsworth (1986) (subsequently HL86 and LH86). By extrapolating statistics of
horizontal correlations to zero separation for the North American radiosonde network,
they estimated (and subsequently subtracted) the contribution to the departure statistics
from observation error. The obvious limitation of such an approach is that it can only be
applied (with any confidence) in an extensive and homogeneously observed area such
as North America. It is unfortunate that it cannot be used to estimate background errors
in the regions sparse in conventional data where the correct use of satellite data is most
crucial.

More recently, the method developed at the US National Centers for Environmental
Prediction (NCEP) and described in Parrish and Derber (1992) has been used to estimate
background-error structures. Fields obtained from a 48-hour forecast are essentially
'verified' with the corresponding 24-hour forecast valid at the same time, and statistics
accumulated over an ensemble of different cases. Under certain (rather stringent) condi-
tions it can be argued that the covariance obtained is a reasonable approximation
to the required error covariance of the short-range forecast (see Bourtier 1994). The
full mathematical justification will not be reproduced here, but it relies upon the 24-
hour forecast being run from an analysis that has benefited from an extra 24 hours of
data assimilation (compared with that used for the 48-hour forecast). Thus, the most
serious limitation of the method is that it is invalid in areas where there is little or
no data updating the assimilation. It is often argued that the global use of satellite
radiance data leaves few areas completely unobserved, but it must be remembered
that the updates or increments due to satellite data are strongly influenced by the
prescribed background-error covariance used in the analysis. Thus, the results obtained
from the 48–24-hour method will not be independent of the existing estimate of the
background-error covariance. While there are substantial doubts regarding the validity
of such an approach, it is an extremely attractive option since it allows the simultaneous
estimation of global background-error correlations for all variables. The method has
been applied operationally at NCEP and many other NWP centres. At the European
Centre for Medium-Range Weather Forecasts (ECMWF) the method has been used to
derive separable background-error covariances in spectral space (Rabier and McNally
1993; Rabier et al. 1998) for use in the three-dimensional variational analysis scheme
(3D-Var).

In section 1 of this paper, the 48–24-hour forecast method has been applied in
grid point (physical) space to investigate the geographic (and temporal) variability of
the estimated background-error structures. These are compared with estimates obtained
from (unpartitioned) radiosonde departure statistics in section 2. Section 3 demonstrates
the sensitivity of radiance-data assimilation to the specification of background error, and
section 4 concludes with a discussion of the implications of the previously estimated
correlations.
Figure 1. Zonal-mean cross sections of the covariance of 48–24-hour forecast-temperature differences at model level 23 (approximately 700 hPa) for (a) July 1995 and (b) January 1996. Contours are at 0.2, 0.5, 1.0, 2.0, 4.0 and 6.0 Kelvin. Note for a surface pressure of 1013 hPa the model levels shown approximately correspond as follows: level 5: 90 hPa, level 10: 208 hPa, level 15: 378 hPa, level 20: 588 hPa, level 25: 817 hPa, level 30: 995 hPa.

2. BACKGROUND ERRORS FROM FORECAST COMPARISONS

The 48–24-hour forecast method was applied in grid-point space to an ensemble of 30 forecast pairs from January 1996 and 30 forecast pairs from July 1995. These were generated from the ECMWF forecasting system that was operational in February 1996 (i.e. the 3D-Var analysis scheme as described in Courtier et al. (1998) and a forecast-model spectral resolution of T213 with 31 sigma levels in the vertical). Figure 1 shows a zonal-mean cross section of the estimated covariance of forecast-temperature error at model level 23 (typically near 700 hPa) during both periods. Figure 2 shows the mean vertical error correlations implied by the covariances of Fig. 1 (it should be noted that only the results at 700 hPa are shown since other levels were found to display very similar behaviour). The covariance cross sections clearly show errors to be largest in the southern hemisphere and smallest in the tropics (as might be expected). The errors are also inflated in the winter hemispheres relative to the summer. Two, very significant features can be seen in the correlation cross sections for both periods. The first is the much sharper error correlations in the tropics compared with those of the extratropical regions. The second is the strong vertical coupling between the tropospheric
Figure 2. Zonal-mean cross sections of the correlation of 48–24-hour forecast-temperature differences at model level 23 (approximately 700 hPa) for (a) July 1995 and (b) January 1996. Contours are at 0.2, 0.4, 0.6, 0.8 and 1.0 K.

and stratospheric errors in the extra tropics, with little or no such coupling in the tropics. The transition between the two regimes is very sharp, with a strong seasonal shift in position.

The spatial distribution of the covariance of 48–24-hour differences is shown in Figs. 3 and 4 where the variance of temperature near 700 hPa is plotted for the two periods. It would appear that the covariances generated by forecast comparisons may, to some extent, reflect where there is conventional data available to the analysis. In other words, it shows places where the 24-hour forecast gains most benefit from 24 hours of data assimilation such that it then differs significantly from the 48-hour forecast (e.g. to the east of the USA and Japan in February). It is important to note that contributions from (and immediately downstream of) these well observed areas (where we have most confidence in the method) will dominate the zonally averaged covariance due to their size (note for example, an error of 2 K has four times the covariance contribution of a 1 K error). The very different error estimates obtained in the areas sparse in conventional data (e.g. the South Pacific in July) appear unrealistically small (compared with those from the data-dense regions) and probably reflect that the 24-hour forecast is not being significantly updated during the extra 24 hours of assimilation. This
suggested that it would be unwise to use the 48–24-hour forecast method to derive a fully three-dimensional model of the background errors (i.e. one that included longitudinal variations) if results from data-sparse regions are unreliable.

3. BACKGROUND ERRORS FROM RADIOSONDE COMPARISONS

Forecast background-minus-observation departure statistics were accumulated for the July 95 and January 96 periods from the ECMWF radiosonde monitoring system. The background is interpolated to the radiosonde location and interpolated in
time between 3-, 6- and 9-hour forecasts made from the previous analysis. Before co-
variances were computed, the local mean (or bias) was computed at each station and
subsequently removed (in general these biases were small compared with the random
departures). Data from suspect stations routinely black-listed by ECMWF and gross
departures (greater than three standard deviations at any level) were also removed.
Figure 5 shows the resulting standard deviations of temperature differences and the
associated sample sizes for three latitude bands. Note that the sample sizes are re-
stricted to cases when a complete observation from 850 hPa to 100 hPa is available
(i.e. a complete profile is necessary to construct a covariance). It can be seen that the
differences are generally comparable to those found in the 48–24-hour forecast for the
northern and southern hemisphere extra tropics. In tropical areas the radiosonde dif-
fferences (typically greater than 1 K) are larger than those suggested by the forecast
comparisons (typically less than 1 K). As expected, the radiosonde differences reach
a maximum near the tropopause and have a marked seasonal change in the northern
and southern hemispheres. The correlations implied by the covariance of background-
minus-radiosonde differences are shown in Fig. 6. The data coverage does not allow
sensible cross sections to be generated so zonally averaged values are shown. It is clear
that there is some agreement between these and the previous (48–24-hour forecast) es-
timates for the southern hemisphere (i.e. broad coupled error structures), but for the
tropics it is less and almost non-existent for the northern hemisphere. In particular, the
latter does not show any significant coupling between tropospheric and stratospheric
levels.

In the southern hemisphere (particularly during the winter) the departure statistics
will be more representative of the background-error structures (since they are larger and
the relative contribution of the radiosonde error will be less). In the tropics and northern
hemisphere the relative contribution of the radiosonde error will be more significant
(and may even dominate the departure statistics). For the North American region we
may refer back to the results from HL86 and LH86 that tend to support the idea of broad
and coupled correlations, but for the tropics it is difficult to say that the radiosonde
comparisons offer any strong confirmation of the results from the 48–24-hour forecast
method.

4. IMPLICATIONS FOR RADIANCE ASSIMILATION

Given that there may be some considerable uncertainty in the true background-error
correlations it is relevant to test the sensitivity of a radiance assimilation scheme to an
incorrect specification of the background-error covariance. This is done using a similar
approach to that described in McNally (1990), where a one-dimensional analysis (or
retrieval) is performed using a linear maximum-likelihood retrieval estimator of the form

\[ X_a = X_b + W(Y_m - Y(X_b)) \]  \hspace{1cm} (1)

where \( X_a \) is the analysed (or retrieved) profile, \( X_b \) is the background profile, \( Y_m \) is the
vector of observed radiances and \( Y(X_b) \) is a vector of radiances computed from the
background profile. The maximum-likelihood weights (where all errors are assumed to
be unbiased and normally distributed) are given by the expression

\[ W = (HB)^T(HBH^T + (O + F))^{-1} \]  \hspace{1cm} (2)

where \( B \) is the background error covariance and \( (O + F) \) is the combined observation
and forward model error respectively (the actual values used for this study and the
Figure 5. Standard deviation of radiosonde-minus-background temperature statistics at all levels in three latitude bands (northern hemisphere 60°N–40°N, tropics 30°N–30°S, southern hemisphere 40°S–60°S). (a) for the July 1995 period and (b) for the January 1996 period.
Figure 6. Vertical correlation of radiosonde-minus-background temperature statistics at 700 hPa in three latitude bands (northern hemisphere 60°N-40°N, tropics 30°N-30°S, southern hemisphere 40°S-60°S). (a) for the July 1995 period and (b) for the January 1996 period.
channels used are those quoted in Table B.3 of Courtier et al. 1998). H is the forward operator mapping departures from the background in profile space to radiance space and is derived using the RTTOV radiative transfer model described by Eyre (1991). For this study H was computed using a mean mid-latitude profile, but the mapping from temperature to radiance space is not found to be strongly sensitive to this choice. The same is not true for humidity (where H is very sensitive to the atmospheric state), but in this study humidity background errors are set to very small values (which are uncorrelated with temperature) essentially removing humidity from the problem.

It is argued that this simplified one-dimensional analysis is a reasonable approximation of the result of analysing a radiance observation in a 3D-Var scheme far away from any other observations and in the absence of any other constraints imposed upon the solution.

If all the parameters (i.e. H, B, O and F) are specified correctly in the estimator W, the resulting analysis, Xa, is optimal and has error covariance

\[ S_{\text{opt}} = (I - WH)B(I - WH)^T + W(O + F)W^T \] (3)

where I is the unit matrix. It can be shown that if the true background error is actually described by B, but an incorrect estimate \( B^* \) is assumed in the estimator \( W_{\text{sub-opt}} \), the resulting sub-optimal error covariance of the analysis \( X_{\text{a}} \) is given by

\[ S_{\text{sub-opt}} = (I - W_{\text{sub-opt}}H)B(I - W_{\text{sub-opt}}H)^T + W_{\text{sub-opt}}(O + F)W_{\text{sub-opt}}^T. \] (4)

Thus, we have a mechanism to investigate the sensitivity of analysis error to an incorrect specification of the background-error covariance.

The 48–24-hour forecast method identifies two, very different correlation structures for background errors. The broad and coupled structure appropriate in the northern and southern extra tropics and the very sharp and decoupled errors of the tropics. Two background-temperature error-covariance matrices, \( B_{\text{ET}} \) (with broad and coupled correlations) and \( B_{\text{TR}} \) (with sharp and decoupled correlations), were constructed that have the same diagonal elements (i.e. variance). Figure 7 shows the analysis and background-temperature errors (expressed in terms of standard deviation) when the true background errors are described by \( B_{\text{ET}} \). Two, different TOVS channel configurations were used to simulate clear conditions (where all infrared and microwave channels can be used) and cloudy conditions (where only microwave channels can be used with a significant loss of vertical information). It can be seen that the incorrect use of \( B_{\text{TR}} \) in the linear estimator (sub-optimal analysis) causes a degradation of the analysis error (larger in the cloudy case), but the sub-optimal analyses are still an improvement over the background. Figure 8 shows the analysis and background-temperature errors when the true background errors are described by \( B_{\text{TR}} \). The incorrect use of \( B_{\text{ET}} \) in the estimator (sub-optimal analysis) again causes a degradation of the analysis error. However, the sub-optimal analysis, using only microwave channels, now has a larger error than the background. Put another way, the assimilation of the microwave-only radiances (with an incorrect description of the background-error correlations) has actually degraded the background. This result reflects the general difficulty in extracting useful information from low-resolution radiances when the background errors are sharp and uncorrelated (as our estimate suggests they are in the tropics). It can be seen in Fig. 8 that, even using the correct covariance in the analysis operator, the improvement over the background is marginal.
Figure 7. Simulated standard deviation of background and analysis-temperature error (optimal and sub-optimal) when the true background-error correlations are broad and strongly coupled. The pair of lines labelled ‘CLEAR’ correspond to the TIROS (Television Infra-Red Observation Satellite) Operational Vertical Sounder (TOVS) channel configuration that includes infrared and microwave channels. The pair of lines labelled ‘CLOUDY’ correspond to the TOVS channel configuration that includes only microwave channels.

5. Discussion of results and conclusions

Some features of the zonally averaged correlation structures derived from 48–24-hour comparisons appear reasonable in that they are consistent with synoptic investigations of a number of individual cases during the January and July periods. It is evident that large forecast errors (which dominate the covariance) tend to be associated with a poor location or phase problems in the prediction of extra-tropical features such as ridges and troughs (particularly visible in the southern hemisphere winter). We expect such errors to be strongly correlated in the vertical, but also negatively coupled between the troposphere and stratosphere (considering the normally reversed pole–equator temperature gradients). The radiosonde comparisons of past studies and the unpartitioned results described here (at least for the southern hemisphere) also support a broad coupled structure outside the tropics, and we may have some confidence in this model as a good estimate of the true background error.

The spatial distribution of differences obtained from 48–24-hour forecast comparisons demonstrate the weakness of the derived background-error estimates when the 24-hour forecast is not significantly updated by conventional data. This suggests that the 48–24-hour method should not be used to determine fully three-dimensional variations in forecast-error structures.

For the tropics there is some agreement between the 48–24-hour results and the radiosonde comparisons (i.e. neither suggest a strong coupling). However, the small size of the background error in the tropics indicates that it is meaningless to ignore the
observation-error contribution in the departure statistics. Thus, the radiosonde comparisons should not be regarded as a strong confirmation (or otherwise) of the 48–24-hour results. In other words, we cannot definitely say at this stage if there is a problem using strongly coupled forecast errors in the tropics or not. All we can say is that if there remains considerable uncertainty in our estimate of the background-error correlations for the tropics, the simulation of section 4 gives us some confidence to use the sharp and decoupled correlations suggested by the 48–24-hour method. It was shown that even if the true tropical errors are strongly coupled (there is no evidence to suggest that they are), the use of a sharp and decoupled covariance matrix in the analysis should still allow the extraction of useful information from the radiances (i.e. improve over the background). However, the simulations also suggest that the use of a broad coupled correlation model in the tropics could lead to the analysis of radiances degrading the background in cloudy conditions if the true tropical errors are sharp and decoupled. It is important to note that a globally evaluated correlation model (as is currently used in many NWP centres) would inevitably reflect the structure of large error contributions from the extra tropics (i.e. have broad and coupled correlations), and could therefore have such a detrimental effect in the tropics.

The question of how best to model the variations in background errors (suggested by these results) in an operational global-analysis system is beyond the scope of this paper. It is certainly not a trivial problem considering the sharp transition between the tropical and extra-tropical regimes and the seasonal migration of the transition zone. Also, the problem of modelling the geographic variations may be further complicated if the analysis is formulated in spectral space (as it is at ECMWF and NCEP). However, it is clear that the issue must be addressed.
The simulations presented here are intended to approximate the use of radiance observations in a three-dimensional analysis scheme. With the current development of four-dimensional variational analysis approaches (4D-Var) it is anticipated that the implied forecast-error structures will evolve dynamically during the assimilation period. However, the prescribed (or static) covariance model will still be important and in many situations still dominate the interpretation of radiance data.

It was noted in the introduction that the sensitivity to background-error correlations is particularly high for TOVS radiance assimilation due to the poor vertical resolution of the data. Best estimates of the information content for the next generation of satellite sounding instruments do not suggest a radical increase in the number of independently observed pieces of information (e.g. Prunet et al. (1998) estimate that the Infrared Atmospheric Sounding Interferometer, with many thousands of channels, has about 19 pieces of temperature and moisture information). If we also consider the expected future enhancements in forecast-model vertical resolution it is certain that the accurate specification of background errors will remain a major concern for the successful exploitation of satellite sounding data in NWP.

REFERENCES


Daley, R.


Dee, D. P., Gaspari, G., Redder, C., Rukhovets, L. and da Silva, A. M.

Eyre, J. R.

Hollingsworth, A. and Lonnberg, P.

Lonnberg, P. and Hollingsworth, A.


Parrish, D. F. and Derber, J. C.

Prunet, P., Thépaut, J. N. and Casse, V.

Rabier, F. and McNally, A. P.


