USE OF TENDENCY-BALANCE TO DIAGNOSE THE PERFORMANCE OF MODELS AND DATA ASSIMILATION SYSTEMS

E. Klinker and L. Ferranti (ECMWF)

1. Introduction

The success of finding the origin of forecast errors is limited as error causes can be of different nature and errors are often the result of local and remote forcing problems. A separation of forecast errors into model errors and analysis errors is a particularly difficult task. Some progress is possible by averaging over a large ensemble of forecasts and verifying analyses. Assuming that analysis errors play a decreasing role with increasing forecast range, time mean of medium range forecast errors are most likely dominated by systematic model errors.

The approach by Klinker and Sardeshmukh (1992) to link initial tendency errors to model errors has been motivated by the fact that ensemble forecast errors in the short and in the medium range were largely similar. The results for one step model integrations suggested that deficiencies in the models mechanical and thermal forcing could be identified from the initial conditions. In particular gravity wave drag and vertical diffusion in the upper troposphere were found to give rise to errors in the momentum budget. Subsequent changes based on the budget calculations, advances in orographic drag parameterizations and additional sensitivity experiments led to the modifications of the parameterization of sub-grid-scale momentum forcing with noticeable improvements of the models systematic errors.

The same approach applied to the UK Meteorological Office forecasting and analysis system (Milton and Wilson, 1996) provided supporting evidence for the hypothesis that the vertical distribution of parameterized orographic drag in the operational model was wrong. From the budget studies and experiments with different drag formulations, it appeared that in the operational model more momentum drag should be applied in the lower troposphere and less in the stratosphere. Changes to the UK Meteorological Office forecast model in that direction improved the momentum budget and the models systematic errors of the large-scale circulation.

Similar budget studies like that for January 1987 as described in Klinker and Sardeshmukh (1992) have been carried out for January 1999. As the initial conditions for short range forecasts have been produced by the new 4D-Var data assimilation system, it could be argued that the initial conditions produced by the analysis represent a well balanced state from which it might be difficult to disentangle mechanical and thermal forcing errors. This is partly true as the 4D-Var data assimilation represents not only an initialization of the gravity modes but of the Rossby modes as well, at least over the period of the current assimilation time window of 6 hours.

However, the model trajectory that minimizes the distance to the available observations is still influenced by model errors in a way that the final trajectory deviates from the ideal one in the limits of observational error. In a weakly constrained assimilation system this limit should be relaxed to allow the trajectory to deviate from observations even more with increasing model integration time. With the differences between the optimal interpolation technique and 4D-var for data assimilation in mind we have to judge the usefulness of the budget approach in the 4D-Var assimilation environment from the results for January 1999.

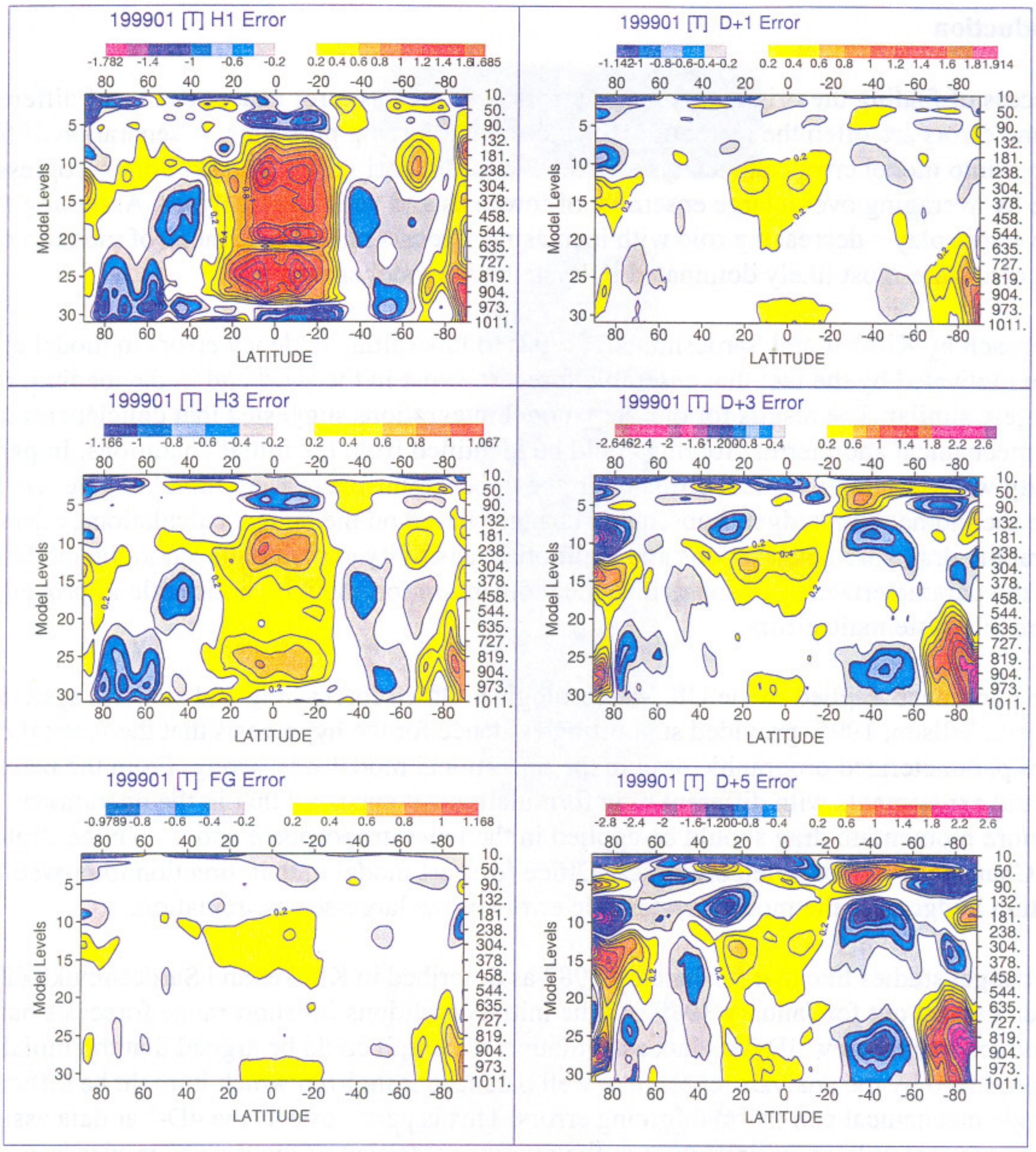

Figure 1: Zonal mean temperature errors for January 1999. Left hand panels from top: Initial tendency error, 3 hour tendency error and first guess (9 hour) error. Unit: K/day. right hand panels from top: day1, day-3 and day-5 errors. Units: K.

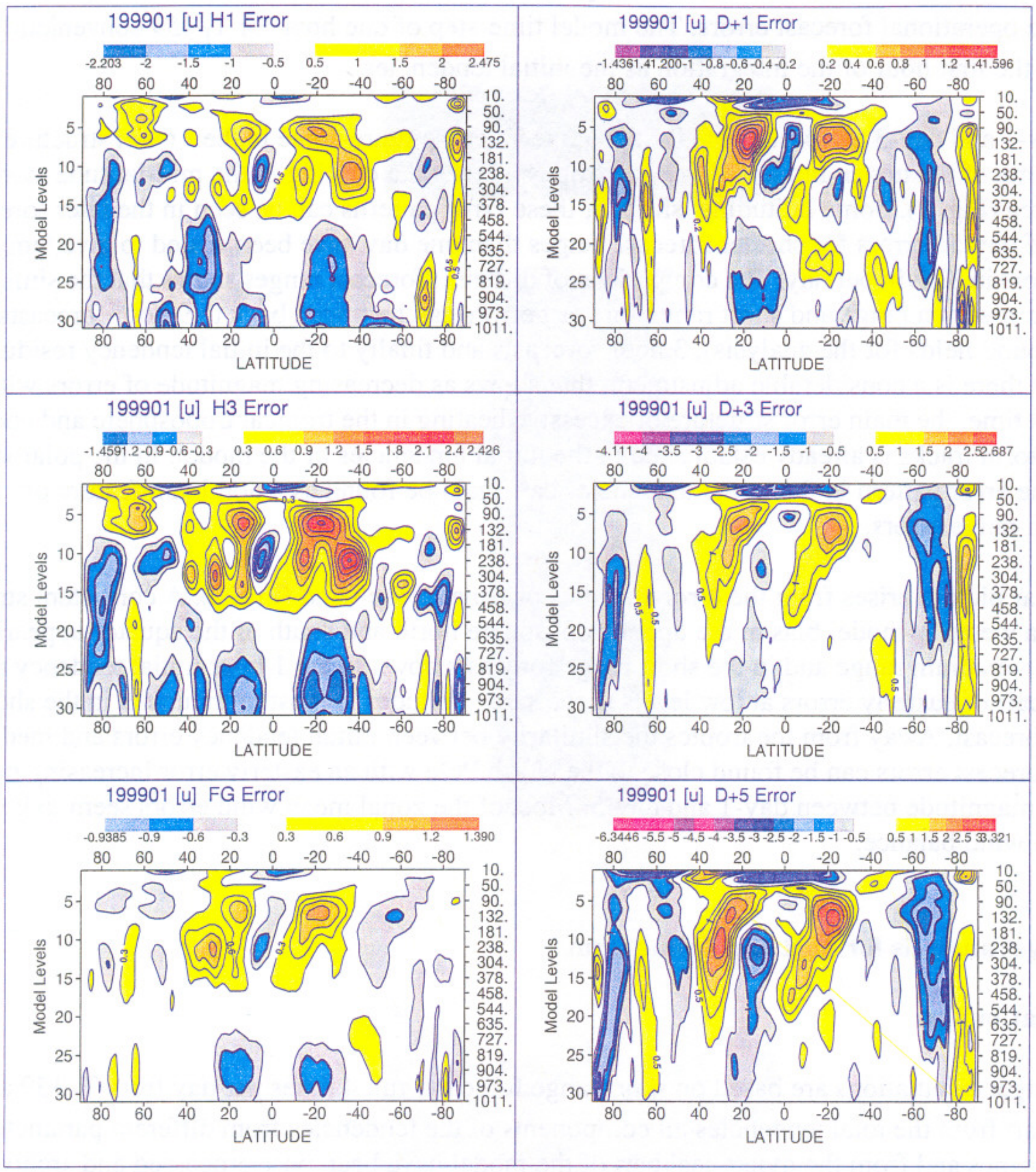

Figure 2: Zonal mean zonal wind errors for January 1999. Left hand panels from top: Initial tendency error, 3 hour tendency error and first guess (9 hour) error. Unit: m/s/day. right hand panels from top: day1, day-3 and day-5 errors. Units: m/s.

2. Time mean errors for January 1999

For reasons of economy the model runs for January 1999 have been performed with a horizontal resolution of T_L159 based on initial conditions produced by a 4D-Var data assimilation system with an outer loop of T_L159 and an inner loop of T63. For the operational forecast model and for the outer loop of the data assimilation a higher resolution of T_L319 is used. However, the monthly mean errors of the low resolution T_L159 system and the high resolution operational T_L319 system are almost identical. Therefore we can compare the initial tendency residual of the T_L159 system with the operational forecast errors. The model time-step of one hour for T_L159 conveniently defines the first hour of the integration as the initial tendencies.

The day-1 and day-3 forecast errors for zonal mean temperature agree in their main structure with a warming of the upper and lower tropical troposphere and a cooling in the mid-latitude storm track area (Fig. 1). Some latitudinal shifts of these error patterns can be seen in the later forecast range of day-5. Errors for shorter forecast ranges than one day have been scaled to give temperature errors in Kelvin per day. The comparison of different forecast ranges shows that the similarity between medium range and short range forecast errors can be traced back to 9 hour forecasts (the background fields for the analysis), 3 hour forecasts and finally to the initial tendency residual. Though there is a considerable adjustment, that shows as decreasing magnitude of errors with forecast time, the main error structure of excessive heating in the tropical troposphere and cooling in the storm tracks is already established in the initial tendencies of the model. In the polar latitudes the error pattern in the initial tendencies can again be found as a dominant feature of medium range errors.

A similar picture arises from the error statistics of zonal mean wind (Fig. 2). A consistent structure of a westerly model bias in the upper troposphere north and south of the equator appears in both, the medium range and in the short range forecasts down to the 1 hour initial tendency residual. A pair of easterly errors at low levels at the same latitudes is most pronounced in the short range forecast. Away from the tropics the similarity between initial tendency errors and medium range forecast errors can be found close to the North Pole with an easterly error increasing noticeably in magnitude between day-1 and day-5. Most of the zonal mean wind errors seem to be in thermal wind balance.

3. Budget analysis for heat and momentum

3. 1 Heat budget

The budget calculations are based on short range forecasts run 4 times per day from T_L159 analyses. Apart from the total tendencies all components of the tendencies from different parameterization schemes and from the dynamical part of the model have been post-processed and archived. Fig. 3 shows a comparison between the monthly mean of all parameterized temperature tendencies and the dynamical tendencies with the sign reversed. The large imbalance in tropical troposphere as shown in Fig. 1 (top left panel) coincides with strong diabatic heating arising from cumulus convection. A negative budget residual in the region of evaporative cooling on top of the convective towers indicates that the convective processes are not only too strong but penetrate too high into the upper troposphere as well. The negative residual in the storm track region corresponds roughly to the latitude band of maximum condensational heating from large scale processes. The indication of excessive low level cooling at high latitudes in the Northern Hemisphere indicates an underestimation of turbulent heating by boundary layer processes. From maps of budget residuals and short range forecast errors (not shown) it seems that for cold air outbreaks over coastal waters the heating of the boundary layer is insufficient.

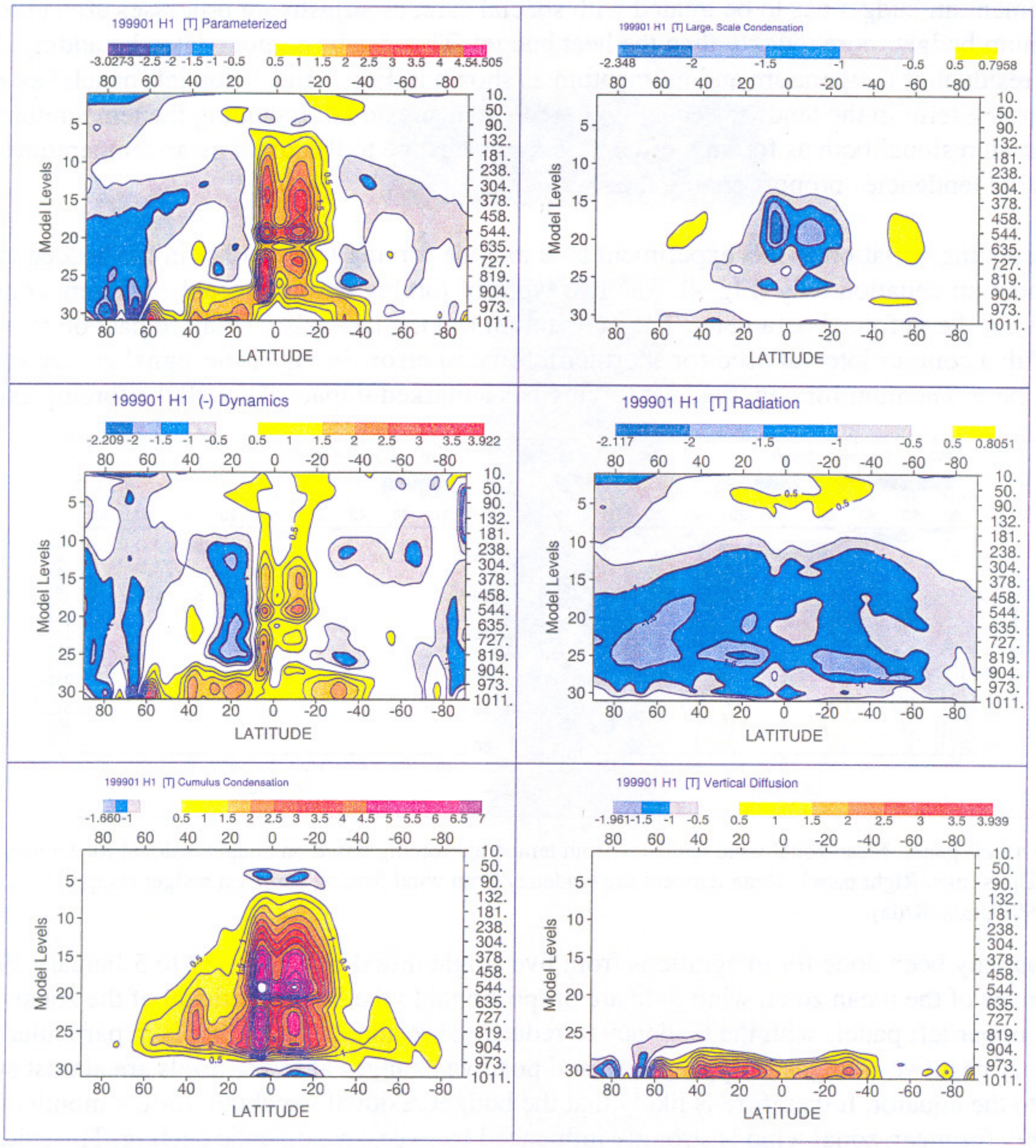

Figure 3: Monthly mean of temperature budget components. Left panels. Top: sum of all parameterized tendencies, middle: dynamical tendencies (sigh reversed), bottom: condensational heating due do cumulus convection. Right panels. Top: condensational heating due do large scale processes, middle: radiative cooling, bottom: heating due to vertical diffusion. Units: K/day.

3.2 Momentum budget

The momentum budget has to be treated with special care, as adjustment processes affect the momentum budget more quickly than the heat budget. This can be demonstrated by adding the budget residual of temperature and momentum as shown in Fig. 1 and 2 (top left panels) as an extra forcing term in the tendency equation to reduce this residual. Including the temperature and momentum residual both as forcing terms produces a change to the zonal mean temperature and zonal wind tendencies proportional to the residual.

The interesting variation to this experiment is to include forcing terms either in the heat equation or momentum equation only (Fig. 4). The two types of forcing have noticeably different consequences for the unforced parameter. The momentum forcing has only a small impact on temperature. With a contour interval used for short term forecast errors in Fig. 1 the signal is very small. Unlike the momentum forcing, the heat forcing has a marked impact. Though the forcing experiment has only been done for integrations from five single initial conditions (1 to 5 January 1999), the changes of the mean zonal wind field are proportional to the budget residual of the zonal wind (Fig. 2, upper left panel) with the tendency to reduce the residual. Interesting is in particular that the pair of negative changes at upper levels and positive changes at lower levels are almost symmetric to the equator. It therefore is likely that the budget residual we obtain for the month of January 1999 for mean zonal wind is strongly influenced by excessive diabatic heating. The evidence for an adjustment process is not surprising as there is apparently not enough data in the tropics to constrain the analysis sufficiently.

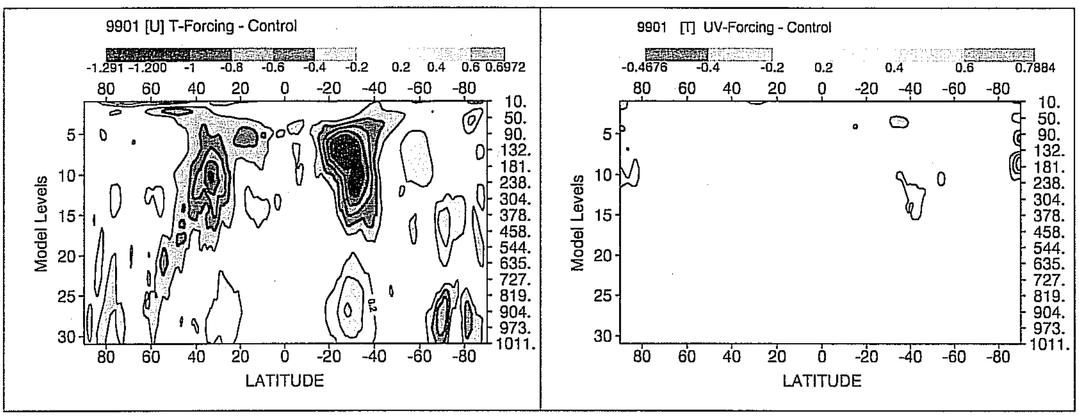

Figure 4:Left panel: Mean zonal wind tendency from temperate forcing based on budget residual for January 1999. Units: m/s. Right panel: Mean temperature tendency from wind forcing based on budget residual for January 1999. Units: K/day.

The major components of the momentum budget are displayed in Fig. 5. The largely unbalanced dynamical tendencies in the tropics have to be seen in the light of the adjustment experiments described in the previous section. Adjustment to temperature errors like those diagnosed for this month could account for a large part of the monthly mean upper and lower level momentum tendencies from the dynamics of the model. Due to the sparse data coverage the results in the tropics are therefore less clear than in the extra-tropics.

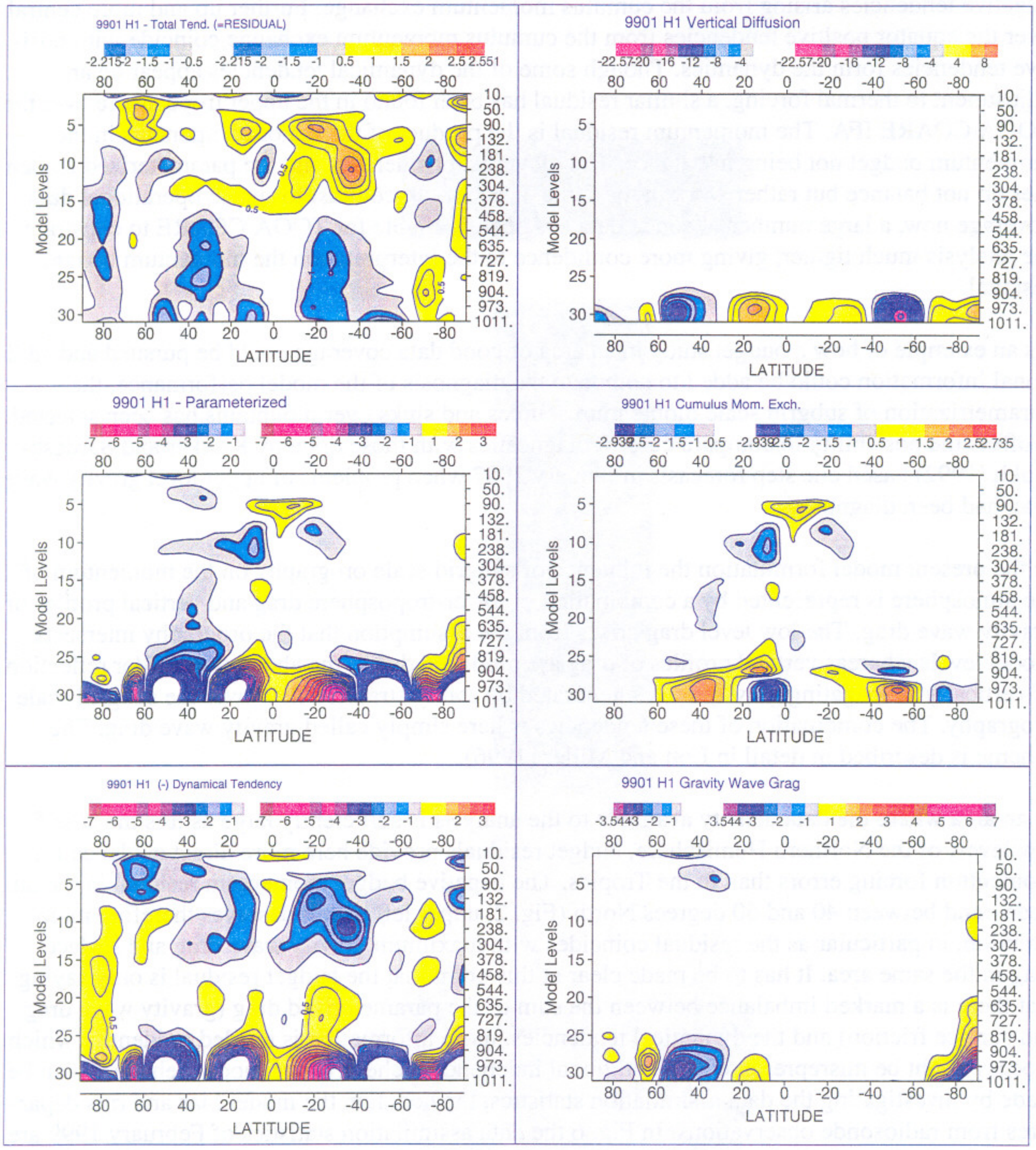

Figure 5: Monthly mean of zonal mean wind budget components. Left panels. Top: Budget residual, middle: sum of all parameterized tendencies, bottom: dynamical tendencies (sigh reversed), Right panels. Top: vertical diffusion, middle: cumulus momentum exchange, bottom: vertical diffusion. Units: m/s/day.

However, where cumulus momentum exchange plays a role, we may see problems in the models mechanical forcing. A negative residual in the upper tropical troposphere coincides with large negative tendencies arising from the cumulus momentum exchange. Further up and more central over the equator positive tendencies from the cumulus momentum exchange coincide with positive tendencies form the dynamics. Though some of the dynamical tendencies appear be an adjustment to thermal forcing, a similar residual has been found in the upper troposphere over the TOGA COARE IFA. The momentum residual is the product of the major components in the momentum budget not being in balance. The dynamical tendencies and the parameterized tendencies do not balance but rather work in the same direction. In contrast to sparse operational data coverage now, a large number of sonde data has been available for TOGA COARE to constrain the analysis much tighter, giving more confidence to the interpretation the momentum budget residual.

As an example of how a budget study in an area of good data coverage could be pursued and additional information could be added to complete the diagnosis of the model performance, the parametrization of subgrid scale momentum sources and sinks over mountains has been selected. It offers the possibility to compare current diagnostics with those done by Klinker and Sardeshmukh (1992) based one step forecasts in January 1987 when problems of upper level gravity wave drag had been diagnosed.

In the present model formulation the influence of subgrid scale orography on the momentum of the atmosphere is represented by a combination of lower-troposphere drag and vertical profiles of gravity wave drag. The low level drag arises from the assumption that the orography intersects model levels whereas vertical profiles of drag are produced due to the absorption and/or reflection of vertically propagating gravity waves generated by stably stratified flow over the subgrid scale orography. The combination of these tendencies is here simply called gravity wave drag. The scheme is described in detail in Lott and Miller (1996).

In an area where the data density available to the analysis is sufficiently large as it is the case for low levels of the Northern Hemisphere, budget residuals provide here more direct evidence for momentum forcing errors than in the Tropics. The negative budget momentum residual in the latitude band between 40 and 60 degrees North (Fig. 2, upper left panel) deserves therefore more attention, in particular as the residual coincides with maximum gravity wave drag and surface drag in the same area. It has to be made clear at this point that the budget residual is only saying that there is a marked imbalance between the sum of the parameterized drag (gravity wave drag and surface friction) and the dynamical tendencies, more information is needed to identify which process might be misrepresented.An additional independent check of the model behaviour can be made by investigating the data assimilation statistics, they contain the model bias and rms departures from radiosonde observations. In Fig. 6 the data assimilation statistics of February 1999 are shown for the two wind components covering an area of the Rocky Mountains from 30 to 50 degrees north and from 120 to 100 degrees west. A marked negative bias for the north-south component of the wind dominates the troposphere. The fact that the model winds have a biased southerly component (observation minus background negative) is consistent with an adjustment of a predominant westerly flow to excessive west-east drag in the area.

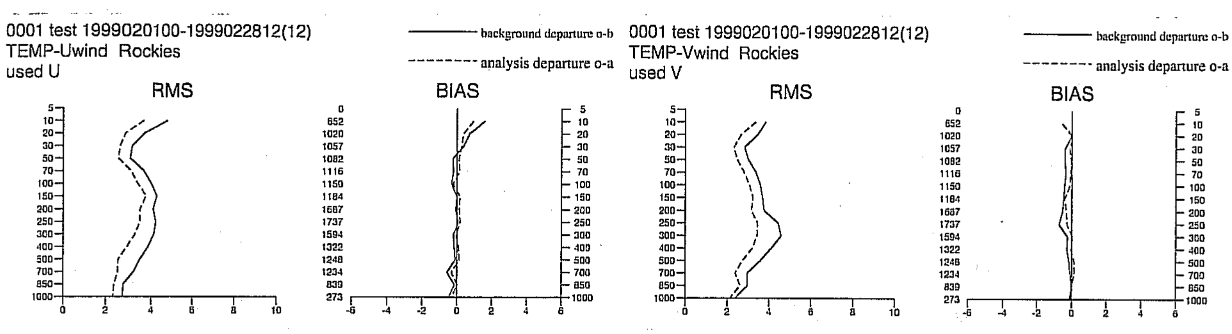

Figure 6: Fit of the two wind components to radiosonde measurements over the Rocky Mountains for February 1999. RMS values and bias are shown for the analysis and background. Left panel: zonal wind component, right panel: north-south wind component.Units: m.s

The geographical extent of the surface drag problem can be seen by integrating the momentum tendencies in the vertical (Fig. 7). Along the major mountain ridge over the west coast of North America the vertically integrated momentum tendencies show a spatially consistent picture of a westerly residual. Integrating the 3-hour tendencies in the vertical the picture is different, the flow has adjusted to excessive drag by an increased southerly component.

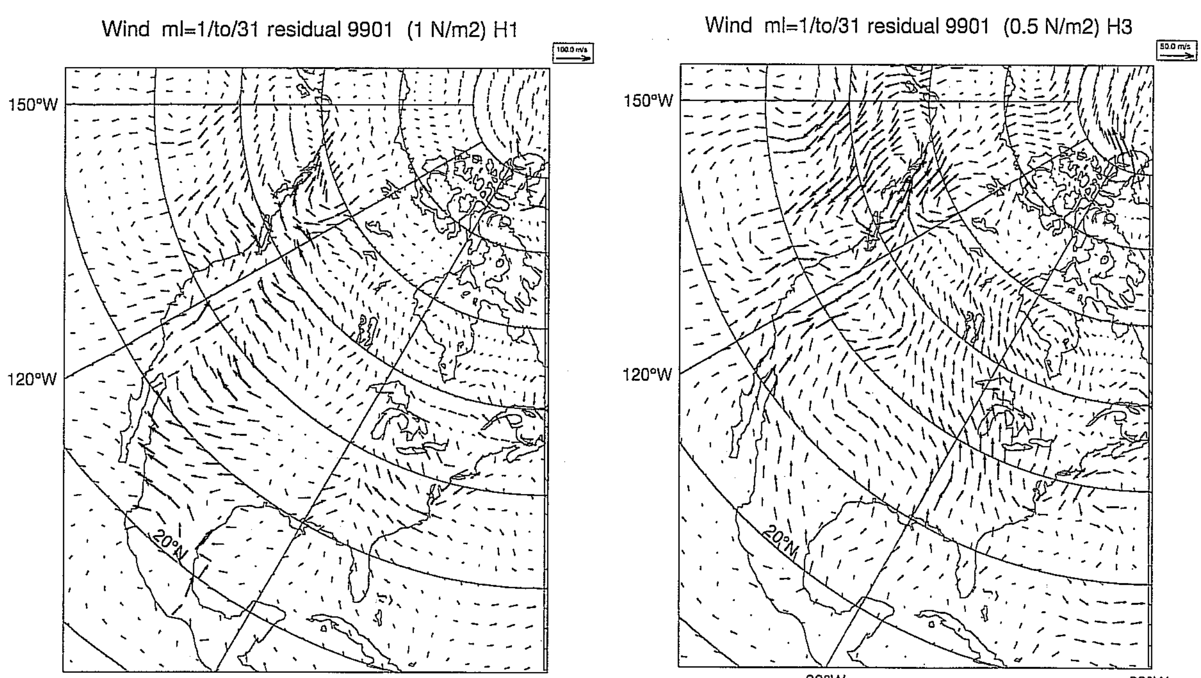

Figure 7: Vertically integrated momentum tendencies. Left panel: 1-hour tendencies, right panel: 3-hour tendencies. Units: N/m**2

4. Flow dependent errors and diabatic forcing

4.1 Simple ensemble technique

So far the focus has been on time mean errors that can explain to a certain extent systematic errors. However, systematic errors can be very much flow dependent which means that errors may be partly compensated by averaging over different flow regimes. Strongly flow dependent errors

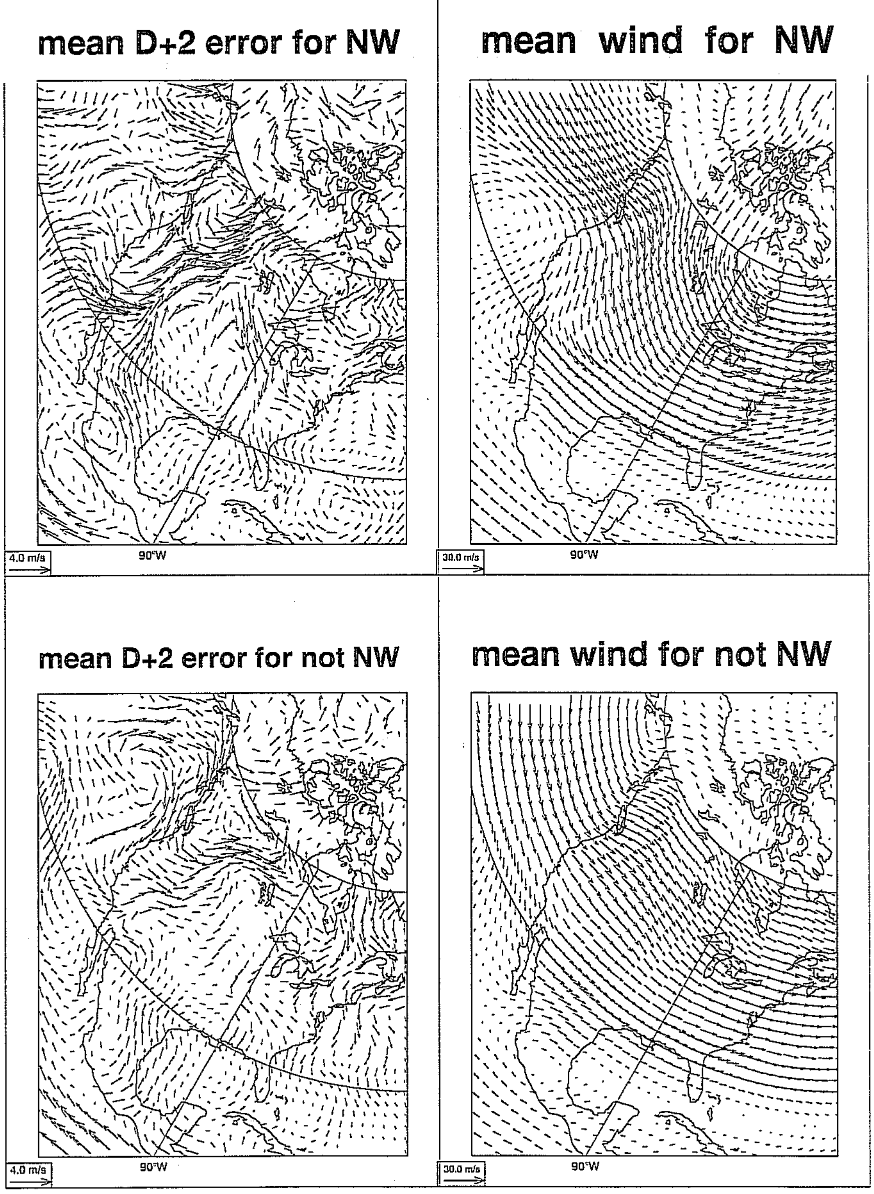

Figure 8: Day-2 forecast errors for strong north-westerly flow (errors upper left and flow upper right panel, 31 cases) and alternative weaker westerly flow (errors lower left and flow lower right panel, 59 cases). Units: m/.s

can be expected close to orographic features of large size like the Rocky Mountains. Therefore a large number of maps of short range forecasts have been processed to investigate the performance of the operational system with respect to the low level flow over the Rocky Mountains. For the winter 1998/99 the daily maps of the near 700 hPa wind fields have been composed into two sets of cases: strong north-westerly flow and alternative cases of mostly rather weak westerly flow in an area centred over the Rocky Mountains. From the result shown in Fig. 8 it is obvious that the forecast errors are rather different for the two flow types with a large southerly component for the cases of strong north-westerlies and much less pronounced for the weak flow. The signature of short range forecast errors is in agreement with the findings from the budget residual that for a strong flow over the mountains a veering of the flow in the forecast suggests again excessive orographic drag in those situations.

4.2 Diagnosis of flow dependent errors using the SVD technique

The question now arises whether the flow dependent error in the low level wind field can be more directly related to the low level drag. Here the technique of singular vector decomposition is used like in Ferranti et al (2000) to relate forecast errors to subgrid scale forcing. The SVD technique selects a linear combination of data from the two input streams that maximizes the covariance. The results are made more stable by decreasing the number of degrees of freedom in each data set. Following the procedure of Barnett and Preisendorfer (1987), this is achieved by performing an EOF analysis independently on the two data-sets, either the pair of analysis and forecast errors or the pair of forecast errors and diabatic forcing. The two components are projected onto a subset of the EOFs to form the two input streams for the singular vector decomposition.

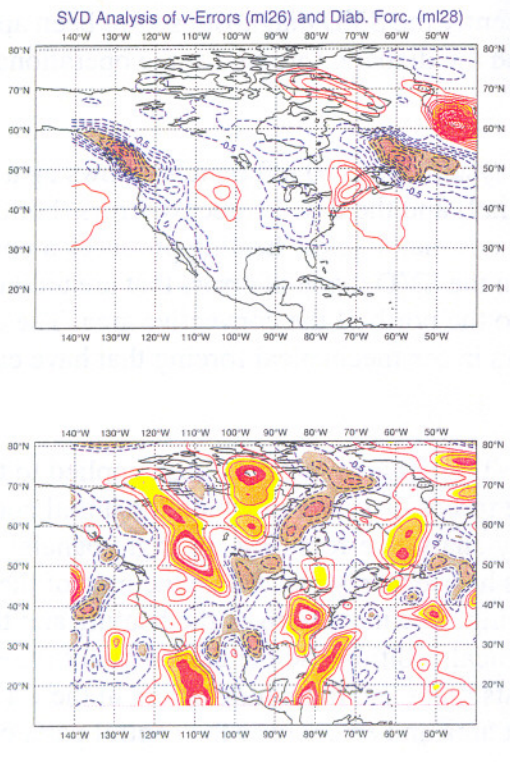

Figure 9: The dominant pair of modes from an SVD analysis applied to diabatic momentum forcing at model level 28 (~940 hPa) and the day-1 forecast error of the north-south component of the wind at model level 26 (~860 hPa). Top panel: momentum forcing, bottom panel: v-errors.

Here we follow a similar procedure by projecting forecast errors of wind and the parameterized momentum forcing onto a subset of leading EOFs before the two data sets are analyzed using the SVD technique. Fig. 9 shows the horizontal structure of the leading SVD pairs: the east-west momentum forcing in the boundary layer and the day-1 error of the north-south component of the wind just above the boundary layer. The correlation of the time series of the two components is 0.6. Though one might argue that the amount of explained variances with 4% for the forcing and 10% for the errors are rather small, it is noticeable that in the region of the Rocky mountains the result is in agreement with the study of flow dependent errors in the previous section. For flow configurations with an east-west drag against the flow over the Rocky Mountains between 50 and 60 degrees north (negative contours in Fig. 9, upper panel) a clear signal of positive errors in the north-south component of the wind (too strong southerlies) can be seen in the lee of the mountains (positive contours in Fig. 9, bottom panel).

A similar diagnosis of diabatic forcing errors using the SVD technique has been carried out by Ferranti et al (2000) to investigate problems related to cumulus convection in the Caribbean region. A high correlation was found between convective heating and day-1 forecast errors in the upper troposphere.

In trying to find a relation between model errors and diabatic forcing a constraining factor is the presence of analysis errors. We have, however, an objective method of reducing analysis errors via the estimate of the sensitivity of forecast errors to initial conditions. Using adjoint integrations of the tangent linear model in an iterative procedure to minimize the day-2 forecast error it is possible to identify the fast growing analysis errors, the so called key analysis errors (Klinker et al, 1998). For the winter 1998/99 we have therefore run short range integrations from two sets of initial conditions, from the standard operational analyses and from analyses that have been corrected by compensating for the key analysis errors. Subsequently the SVD diagnostics have been applied to the two pairs of data-sets, the temperature error and the diabatic forcing for the operational system and for the system with reduced analysis errors.

The leading SVD pairs of modes for the operational system (Fig. 10 left panels) shows again a high correlation between the diabatic forcing (top panel) and the day-1 forecast errors (bottom panel). The relative sign between the two fields indicates that diabatic heating produces a positive error in the Caribbean. There are, however, features in the SVD mode of errors that are unlikely to be directly related to the heating as they occur well to the north of the convective area. The cause of these errors may have been analysis errors or errors in the mechanical forcing that have caused an adjustment in the temperature.

In order to reduce the effect of analysis errors, the SVD technique has then been applied to the forecast error and diabatic heating time series for the forecast system in which the initial conditions were modified on the basis of the diagnosed key analysis errors (Fig. 10, right panels). The effect of deleting some errors in the initial conditions has certainly cleared up the map of the forecast errors that correlates most with the diabatic heating. The maximum of the heating and the largest errors coincide more closely in their geographical position and the error signal further north has been noticeably reduced. The result suggests that the temperature errors in the Caribbean is strongly related to the diabatic heating in that area and errors over the central parts of

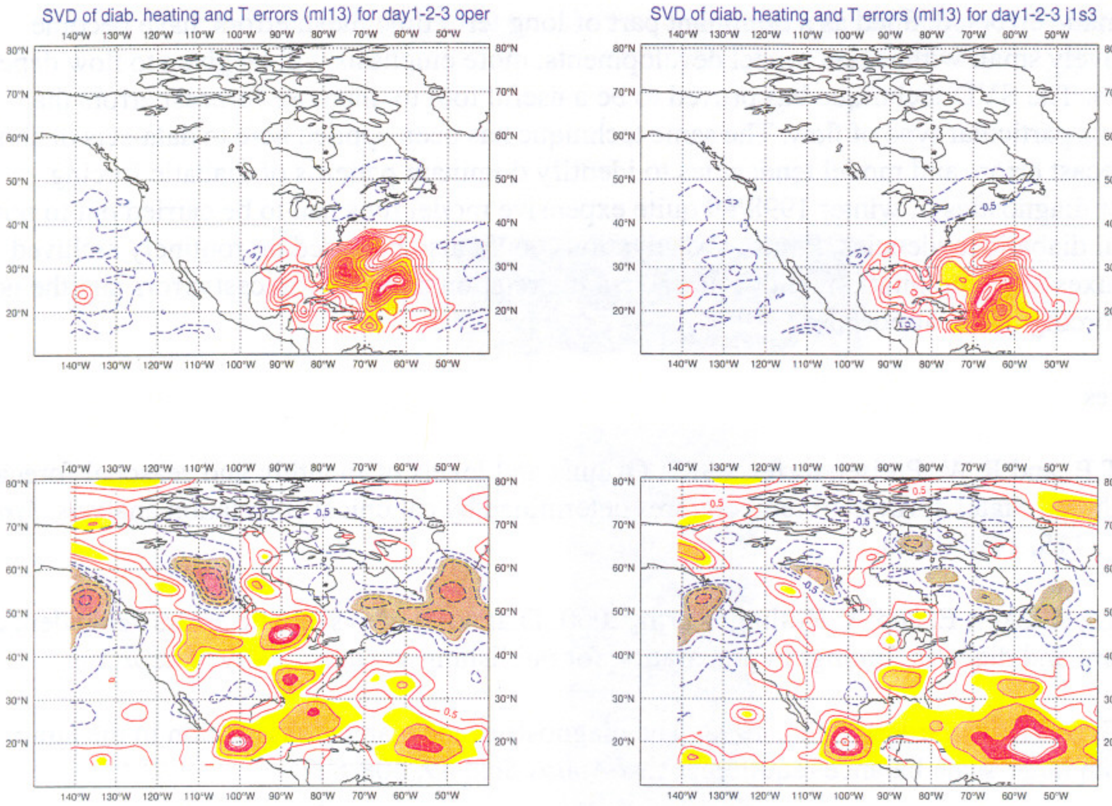

Figure 10: Dominant pair of modes from an SVD analysis applied to diabatic heating and temperature errors at model level 13 (~304hPa). Upper panels: diabatic heating, lower panels: temperature errors. Left row panels: based on operational forecasts, right row panels: based on forecasts from initial conditions with reduced analysis errors

North America are mostly arising from errors in the initial conditions. The result is reassuring for the adjoint calculations to identify analysis errors as well. It has been speculated that errors in the diabatic forcing may be partly compensated by changes in the initial conditions when the day-2 forecast errors are minimized. However, the clear decrease of errors signals away from the forcing suggests that the adjoint calculation are targeting mainly analysis errors.

5. Final comments

The investigation of forecast errors using the initial tendency approach has shown that useful information regarding errors in the thermal and mechanical forcing can still be extracted. Though the 4D-Var data assimilation represents to a certain extend an initialization of the Gravity Waves and Rossby Waves as well, the obtained residual in data rich areas suggests that model errors cause a measurable departure of the trajectory from observations. If future versions of variational analysis would take the evolution of model errors into account, the budget analysis based on initial model trajectories should give an even cleaner picture of those model errors. However, the extension of the data assimilation time window to 12 or 24 hours would make it increasingly difficult to exclude the tidal signal from the long term initial tendencies.

As systematic errors, defined as a dominant part of long term time mean errors, have become comparatively small with recent model developments, more emphasis will be given to flow dependent errors. The SVD-technique has proved to be a useful tool to diagnose forecast errors that depend on a particular type of flow. The same technique has been applied here to data-sets of short range forecast errors and model tendencies to identify dominant patterns of diabatic forcing errors. For diagnostics of winter 1998/99 quite expensive model runs had to be carried out to post-process all diabatic tendencies. Future investigations could also be based on routinely archived surface fluxes and precipitation fields to establish the relation between forecast errors and the net diabatic forcing of the atmosphere.

References

Barnett, T.P. and R.W. Preisendorfer, 1987: Origins and levels of monthly and seasonal forecast skill for United States surface air temperatures determined by canonical correlation analysis. *Mon. Wea. Rev.*, **115,** 1825-1850.

Ferranti, L., Klinker, E. and A. Hollingsworth, 2000: Diagnosis of systematic errors dependent on flow pattern and diabatic forcing. In preparation for publishing in *Quart.J.Roy.Meteor.Soc.*

Klinker, E. and P. D. Sardeshmukh, 1992: The diagnosis of mechanical dissipation in the atmosphere from large-scale balance requirements. *J.Atmos.Sci.*, **49**, 608-627

Klinker, E, F. Rabier and R. Gelaro: Estimation of key analysis errors using the adjoint technique. *Quart.J.Roy.Meteor.Soc.*, **124**, 1909-1933

Lott, F. and M.J. Miller, 1997: A new subgrid-scale orographic drag parametrization: its formulation and testing. *Quart. J. Roy. Meteor. Soc.*, **123**, 101-127.

Milton, S. F. and C. A. Wilson: The impact of parameterized subgrid-scale orographic forcing on systematic errors in a global NWP model. *Quart.J.Roy.Meteor.Soc.*, **124,** 2023-2045