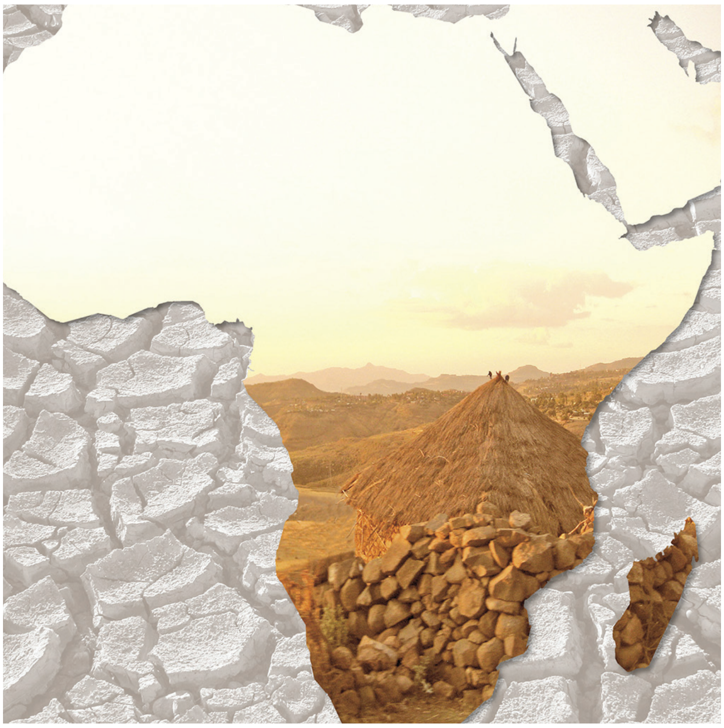
ECMWF Feature article

from Newsletter Number 131 – Spring 2012

METEOROLOGY

A case study of occasional poor medium-range forecasts for Europe



www.ecmwf.int/en/about/news-centre/media-resources

doi:10.21957/ocaqy91i

This article appeared in the Meteorology section of ECMWF Newsletter 131 – Spring 2012, pp. 16–24.

A case study of occasional poor medium-range forecasts for Europe

Mark J. Rodwell, Linus Magnusson, Peter Bauer, Peter Bechtold, Massimo Bonavita, Carla Cardinali, Michail Diamantakis, Paul Earnshaw, Antonio Garcia-Mendez, Lars Isaksen, Erland Källén, Daniel Klocke, Philippe Lopez, Tony McNally, Anders Persson, Fernando Prates, Nils Wedi

In the companion to this article ('*Characteristics of occasional poor medium-range forecasts for Europe*'), it was demonstrated that poor medium-range forecasts for Europe often occur when there is a high over northern Europe. During spring, the initial conditions for these poor forecasts tend to involve warm, moist, southerly flow and high convective available potential energy (CAPE) ahead of a trough over the Rockies. In these situations, the forecast is more sensitivity to the initial conditions – as demonstrated by increased ensemble spread. Hence it is likely that general improvements in the analyses used to initiate our forecasts will result in a reduced frequency of these forecast 'busts'. Using Potential Vorticity budgets, it was also shown that mesoscale convective systems (MCSs) – that accompany the high CAPE – act to slow-down the eastward propagation of the trough, thereby perpetuating the trough/CAPE regime.

Here, we complement the general characterisation of forecast 'busts' with a more detailed investigation of specific case studies. The aim is to identify key factors that could help reduce the frequency or severity of these forecast busts. For computational and technical reasons, sensitivity studies can generally only be made for a few cases. Here attention focuses primarily on just one poor forecast – that of 10 April 2011 (see Figure 1 of the companion paper).

An important issue arises if only poor forecasts are considered – namely the fact that any change to the system is more likely to improve the poor score than degrade it. The effect is an example of 'regression toward the mean'. In the present context, it is uncertainties associated with chaos (when small modifications are made to the model, observations etc.) that tend to improve the bad score, and so this spurious improvement effect is termed here 'chaotic improvement'. Because of this effect, it is the sensitivity studies that do not improve the scores that are most valuable – they allow us to focus other experiments and diagnostic tools on establishing whether the changes that did improve the scores did so for 'real reasons' or simply due to chaotic improvement.

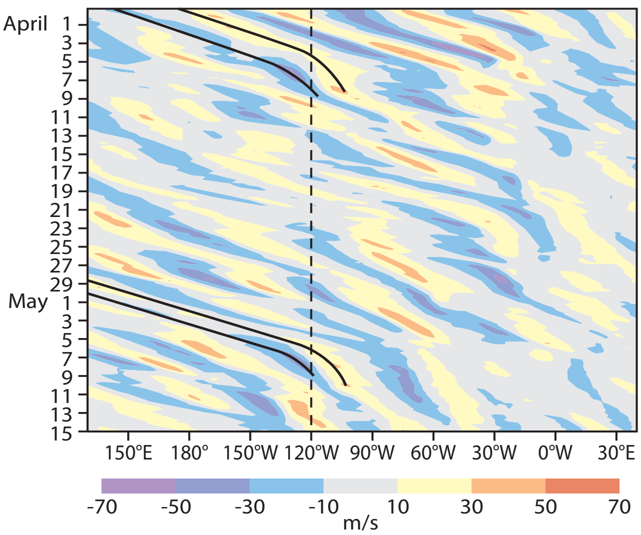


Figure 1 Meridional wind anomaly on the 330 K isentropic surface, averaged over 35°N – 50°N, and plotted as a function of longitude and time. The dashed line indicates the location of the Rockies. The solid lines highlight waves approaching the Rockies from the west. Data are the operational analyses at 00, 06, 12 and 18 UTC from 1 April to 15 May 2011.

 doi:10.21957/ocaqy91i

Correspondence between case study work and general characterisation

Before discussing our sensitivity and diagnostic results, we first look to see how well the bust of 10 April 2011 fits the general characterisation.

Forecast initial conditions preceding the bust

Figure 1 shows meridional wind on the 330 K isentropic surface (approximately at 250 hPa), averaged over the latitude band 35°N – 50°N, from 6-hourly operational analyses, as a function of longitude and time for April and the first half of May 2011. Diagonal stripes, in the left half of this figure, are indicative of (Rossby) waves travelling east across the Pacific. The dashed line indicates the approximate location of the Rockies. It can be seen that on two occasions these waves slow down as a trough crosses over the Rockies (meridional wind negative to the west and positive to the east, of the dashed line). These two events correspond exactly to the two European busts for forecasts starting between 8–10 April and 9–11 May. In this sense it would appear that these two busts fit well the general characterisation. In addition, they highlight the pre-cursor role of waves crossing the Pacific (although there is clearly no one-to-one relationship with the busts).

Figures 2a and 2b show the full fields and anomalies from ERA-Interim of 500 hPa geopotential height (Z500), CAPE and 850 hPa wind at this time. The similarities between Figure 2b and the composite initial conditions (Figures 4a and 4b in the companion article) are striking. One sees the upper-level trough over the Rockies. Ahead of the trough low-level southerlies advect heat and moisture from the Southern USA and Gulf of Mexico – providing the environmental conditions (CAPE) for the development of storms

Forecast uncertainty

The 10 April case is also associated with increased forecast uncertainty. Figure 3 shows, for the first 12 days of April 2011, the Z500 European spatial anomaly correlation coefficient (ACC) for the operational high-resolution forecasts made by ECMWF and the UK Met. Office, along with those for each member of the ECMWF Ensemble Prediction System (EPS). The increased spread in ensemble scores near 10 April is consistent with increased forecast uncertainty. Note that the ECMWF high-resolution forecast score lies well within the spread of the EPS and there is even an ensemble member score matching that of the UK Met. Office deterministic forecast which, for this case, recovered earliest from the bust. Since both deterministic forecasts lie within the ensemble spread, it would be difficult to conclude anything from this single bust case about the underlying relative performance of the two systems. Nevertheless, a comparison of the two systems has proved useful to gauge the relative importance of initial conditions relative to the forecast model; these results might apply more generally.

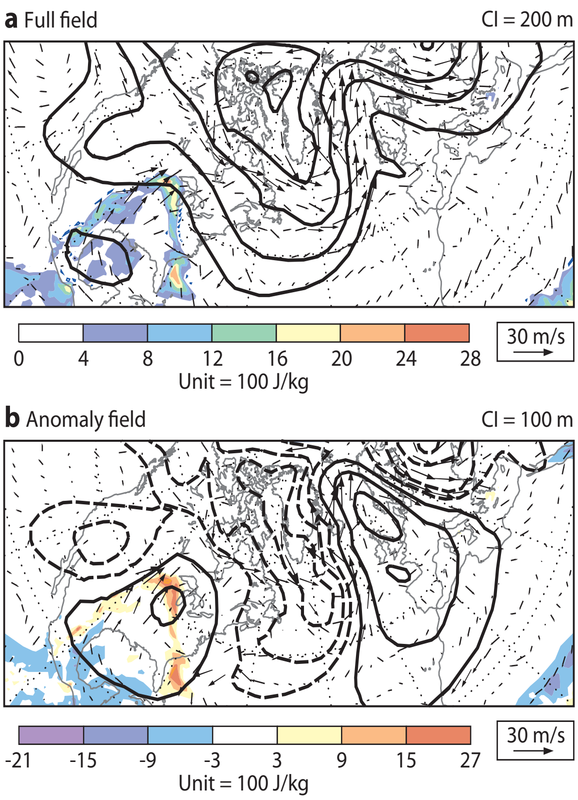


Figure 2 Operational analyses of Z500 (contours), CAPE (shading) and 500 hPa wind (vectors) at 00 UTC on 10 April 2011: (a) full fields and (b) anomalies from ERA-Interim climatology 1989–2008.

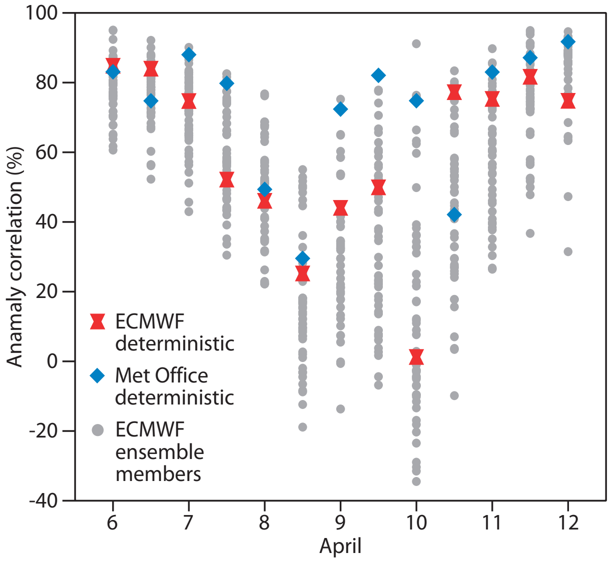


Figure 3 ACC of Z500 forecasts at day 6 for Europe initiated between 6 and 12 April 2011. Grey: for each ensemble member of the ECMWF ensemble prediction system. Red: the ECMWF deterministic forecast. Blue: the UK Met. Office deterministic forecast system. Each centre's forecasts are verified against their own operational analyses.

The importance of the initial conditions

Figure 4a shows the Z500 error at day 6 for the ECMWF operational forecast initiated at 00 UTC on 10 April 2011. The largest errors associated with this European bust occur over the eastern North Atlantic and into Europe (somewhat consistent with the trough/CAPE composite results in Figure 5 of the companion paper). The corresponding errors for the UK Met. Office operational forecast, Figure 4d, are generally smaller – consistent with the UK Met. Office recovering more quickly from the bust. Notice, however, that scores for each forecast are sensitive to the precise region chosen. For example, on this occasion, the ECMWF score for Europe is sensitive to how much of the strong positive height error is included in the north-western corner of the domain (between Scandinavia and Greenland).

When the UK Met. Office forecast is initiated from the ECMWF analysis (Figure 4c), it appears to reproduce the larger ECMWF operational errors. Similarly, when the ECMWF forecast is initiated with the UK Met. Office analysis (Figure 4b), it reproduces the smaller UK Met. Office operational errors. The correspondence between error and initial conditions would appear to indicate that the initial conditions are more important than the model used to make the forecast in this particular case. This tends to reinforce the interpretation of the composite results.

Identifying the salient errors in the initial conditions

Although the above results suggest the importance of the initial conditions for the 10 April bust, we have not yet made a link to the trough/CAPE situation over the USA. As a first approach to identifying the key aspects of the initial conditions, we continue the comparison between the UK Met. Office and ECMWF

Figure 5a shows the difference in operational analyses (UK Met. Office minus ECMWF) at 00 UTC on 10 April 2011. Although there are differences over the USA, there are actually differences everywhere, and these reflect random and systematic aspects. However, it is those differences that project onto fast growing modes that will play the most important role in the error differences that develop by day 6. Figure 5b shows the difference in operational forecasts at day 1 (the contour interval is double that of Figure 5a). The differences over the USA have developed to show a strengthened Rockies trough and downstream ridge in the UK Met. Office forecast. However, differences in other regions remain. In order to better isolate the salient aspects, we now focus on just one forecast system, the ECMWF Integrated Forecasting System (IFS).

If forecast errors are not dominated by model problems, then it is useful to trace these errors back to shorter lead-times in order to highlight the salient errors in the initial conditions. Figure 6a shows the Z500 day-2 error in the ECMWF operational forecast from 00 UTC on 10 April 2011, and this more clearly highlights North America as a key region.

It becomes difficult to trace errors back to even shorter lead-times, as uncertainties in the verifying analysis begin to affect the calculation of forecast error. Instead, one can look at the 50 ensemble members of the EPS. Each ensemble member is started from a slightly perturbed set of initial conditions. Results show a strong correspondence between the initial condition perturbation of a given member and its eventual error over Europe. For example, the two ensemble members that had the smallest root-mean-square error (RMSE) over Europe at day-6 shared essentially the same initial perturbations. Furthermore, another two ensemble members shared the negative of these initial perturbations, and they produced the worst and sixth-worst European scores at day-6.

doi:10.21957/ocaqy91i

We have further isolated the key initial condition perturbations of the best ensemble member by progressively confining the perturbations to ever smaller domains. Figure 6b shows the result of this process. The key perturbations have been confined to the North American/eastern North Pacific region. They highlight a strengthening (of order 5 – 10%) of the Rockies trough and the downstream ridge (and presumably increased CAPE). This is consistent with the comparison with the UK Met. Office. Indeed, over North America the difference at day 1 between the forecast initiated with this perturbation and the control is almost identical to that shown in Figure 5b. At day 6, errors for the perturbed forecast (Figure 6c) are, indeed, reduced over the eastern Atlantic and western Europe compared to the control (Figure 4a).

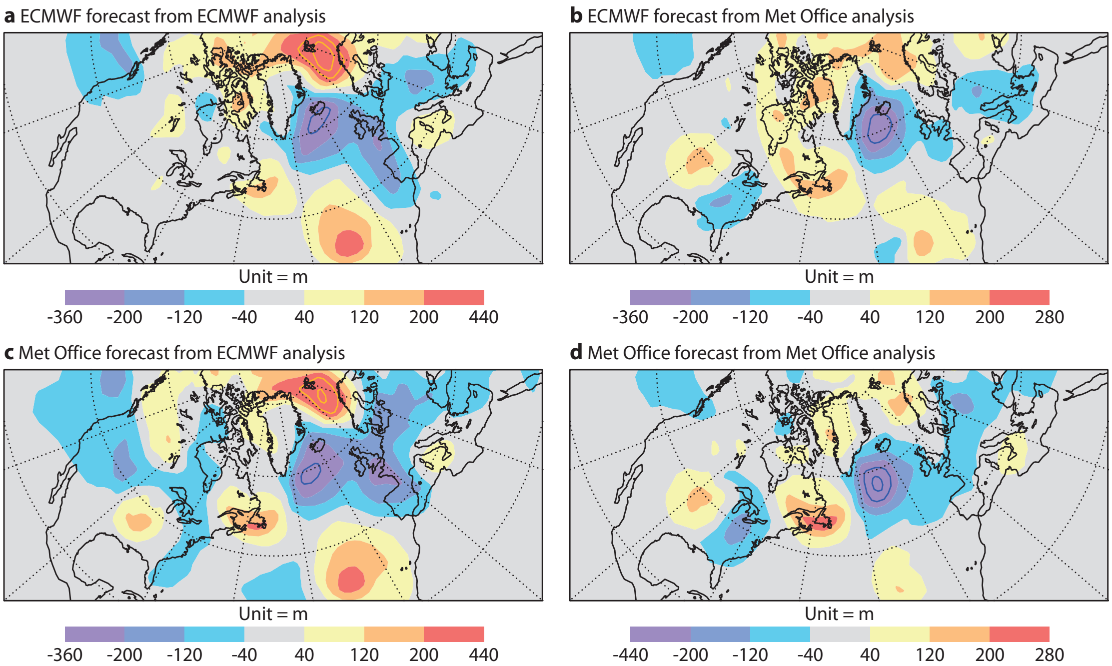


Figure 4 Errors in day-6 forecasts of Z500 from initial conditions at 00 UTC on 10 April 2011. (a) ECMWF forecast started from ECMWF analysis. (b) ECMWF forecast started from UK Met. Office analysis. (c) UK Met. Office forecast started from ECMWF analysis. (d) UK Met. Office forecast started from UK Met. Office analysis. Verification data is ECMWF analysis at 00 UTC on 16 April 2011.

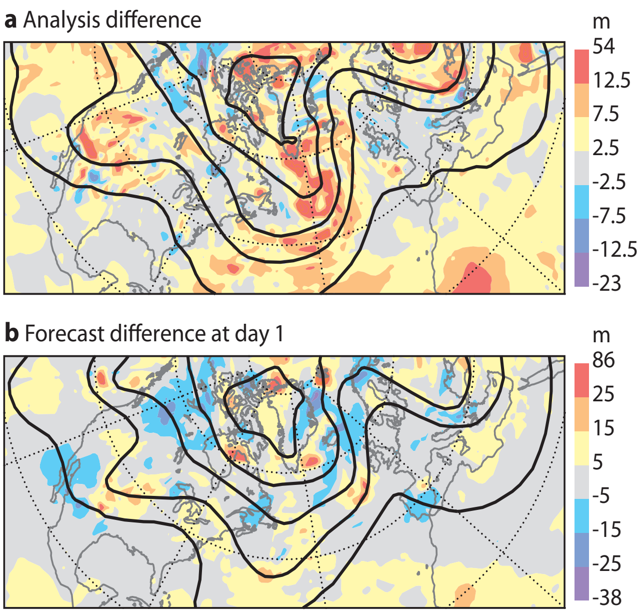


Figure 5 Difference in Z500 between operational forecasts (UK Met. Office minus ECMWF) at (a) day 0 (i.e. the analysis difference) and (b) day 1. Contours show the full field for the UK Met. Office at the same lead-times with interval 200 m.

doi:10.21957/ocaqy91i

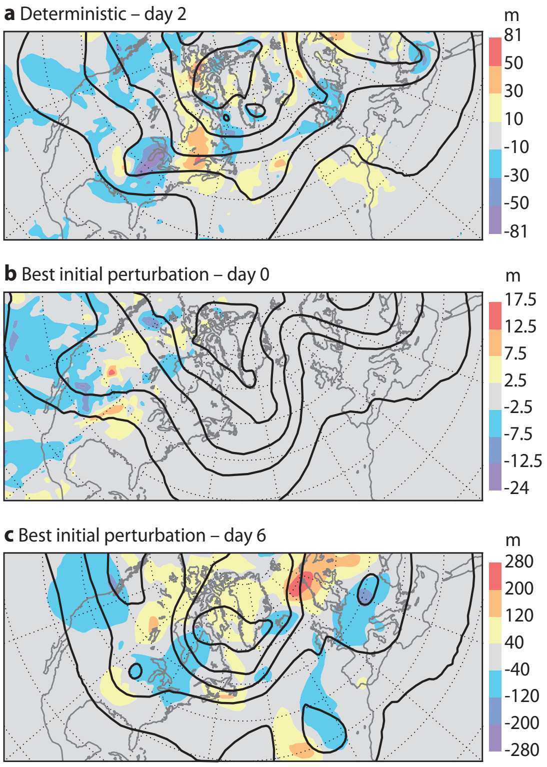


Figure 6 Z500 forecast error (shaded) for ECMWF forecasts initiated at 00 UTC on 10 April 2011. (a) Operational forecast at day 2. (b) and (c) Forecasts with initial condition perturbation equivalent to that of the 'best' operational EPS member, but confined to the region [180°W–90°W, 0°N–90°N] at day 0 and day 6. In all panels, the full forecast field is contoured with contour interval 200 m.

Summary of the correspondence between the case study and the general characterisation

For the bust of 10 April 2011, there appears to be a strong similarity with the general characterisation of spring busts discussed in the companion article. A wave packet crossing the Pacific slows-down when a trough is over the Rockies and a ridge is over the eastern USA. We have shown that small perturbations to this trough/ridge structure lead to large differences in day-6 errors (and ensemble spread) over the eastern North Atlantic and into Europe. In this case, the best initial perturbation strengthens the trough and ridge. By identifying a key perturbation structure, we have been able to go further, for this particular case, than the general characterisation. However, there is no reason to assume that the sign of the best perturbation is the same for all busts. Indeed, it is unclear from this one case how common the best perturbation structure is to all busts.

Mesoscale convective systems in the data assimilation system and forecast model

The composite results demonstrated that MCSs (associated with increased CAPE) are active components in the propagation of the trough. Hence it is worth understanding how these systems are represented in the forecast model and corrected by the data assimilation. Several cases have been considered – both for MCSs that occur during busts and for those that occur during 'no-bust' conditions. Conclusions are similar in all cases but, for consistency we continue to focus on the 00 UTC 10 April 2011 analysis and forecast

MSC events in the first-guess forecast

Figure 7a shows 'first-guess' precipitation accumulated over the 12-hour data assimilation window that was used to make the analysis. The area plotted has been limited to that reliably 'observed' by NEXRAD ground-based radar – shown in Figure 7b. By eye, the first-guess and observed fields display good correspondence – both showing two MCSs over the USA: one to the west of the Great Lakes (up to 30 mm and, incidentally, associated with numerous tornado reports) and the other near the east coast of the USA (up to 50 mm). However, there are considerable differences between the first-guess and the observations (Figure 7c) of over 25 mm – associated with location and intensity errors.

One of the strongest MCS events during spring 2011 was on 24 April and centred over Cleveland, Ohio (up to 100 mm). In this case, the first-guess forecast managed to predict the location reasonably well, but the intensity was underestimated (even at large scales) by as much as 60 mm.

Grazzini & Isaksen (2002) highlighted a tendency for the model at that time to erroneously resolve the fluxes associated with convection and thus produce 'large-scale' precipitation. Our investigations have revealed that the precipitation is now largely associated with parametrized convection – which is thought to be a significant improvement since present model resolution is still too course to resolve real convective fluxes.

doi:10.21957/ocaqy91i

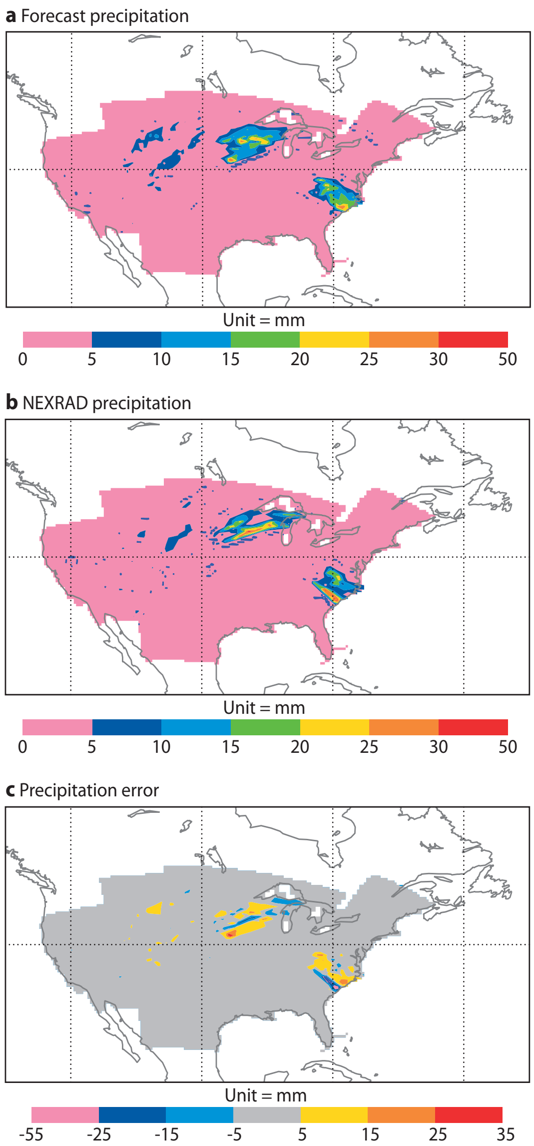


Figure 7 Accumulated precipitation over the 12-hour period from 21 UTC on 9 April 2011 to 09 UTC on 10 April 2011. (a) Precipitation from the first-guess forecast started at 18 UTC on 9 April 2011. (b) Precipitation as retrieved from NEXRAD radar data. (c) Precipitation error: (a) minus (b).

Use and impact of observations in the correction of MCS errors

In order for the data assimilation to correct MCS errors (or any other errors) in the first-guess forecast, it requires relevant observations. To compare the first-guess field with the observations, the first guess is interpolated to the observation locations. The data assimilation system then acts to draw the analysis away from the first-guess and closer (in general) to the observations in a way that is consistent with estimated observation and model errors. The difference between the final analysis and the first-guess is known as the ‘analysis increment’. Figure 8 shows analysis increments for two representative observation types during the production of the 00 UTC analysis on 10 April 2011.

Aircraft data (Figure 8a – known as AIREP data) are particularly important for the upper-tropospheric analysis over the USA (lower-down, flights converge at airports and the horizontal data coverage becomes poorer). Although other observation types will have an influence on these increments, it is likely that the AIREP data plays a major role in correcting upper-tropospheric winds and temperatures in the region of the MCS over the east coast of the USA. However, there was no AIREP data assimilated in the region of the other MCS, to the west of Lake Michigan. Comparison with AIREP data on the same day of the week, seven days later, suggests that flights were avoiding the extreme weather associated with the MCS.

If signficant cloud is present, then satellite observations are also difficult to assimilate. The coloured squares in Figure 8b show the analysis increments for the AMSUA microwave channel 5, which measures mid-tropospheric temperatures. While AMSUA data is generally very powerful within the data assimilation system, ‘holes’ can be seen over the MCS regions where cloud has led to the rejection of data. Similar holes occur for the AIRS infrared channel 156 and the AMSUA microwave channel 7, both of which measure upper-tropospheric temperatures.

While AMSUA data is generally very powerful within the data assimilation system, ‘holes’ can be seen over the MCS regions where cloud has led to the rejection of data. Similar holes occur for the AIRS infrared channel 156 and the AMSUA microwave channel 7, both of which measure upper-trop-ospheric temperatures.

The black circles in Figure 8b show the radiosonde network. While thought to be quite accurate, these data tend to be too sparse to resolve features of the scale of MCSs. Other data can be rejected if the difference with the first-guess is too large. In addition, we are only able to use some satellite observations over the ocean.

These results indicate that, for the variety of reasons discussed above, there are fewer in-situ observations available to the data assimilation within MCSs. Similar conclusions can be drawn from the other MCS events investigated.

Note that, since the busts of spring 2011, the assimilation of NEXRAD precipitation rates has been implemented in the IFS. Results show that the analysis does draw towards to these observations within MCS events, but the actual impact of the radar data on the forecast (relative to the impacts of other surface observations) remains to be quantified.

The four-dimensional variational data assimilation system (4D-Var) optimally fits a model trajectory through all the available observations and this means that observations outside an MCS can correct the first-guess within the MCS. Above, we interpolated to individual observation locations to highlight the reduced availability of in-situ data but, to assess the aggregate impact of all observations, we now look at the model fields.

Figure 9c shows these analysis increments for temperatures and winds at 500 hPa. The question we would like to answer is whether the magnitudes of the increments in the MCS regions are consistent with the first-guess precipitation errors shown in Figure 6c.

We can decompose the evolution of the first-guess forecast into the contributions from the dynamics and each of the physical processes. Within the MCS events, there is strong convective heating due to latent heat release (Figure 9b). The precipitation data suggest that this heating is in error by at least 50%. However, much of the convective heating is balanced by dynamical cooling associated with ascent (Figure 9a), and so it is not appropriate to compare the magnitude of the increment with that of the convective heating error.

Other processes (not shown) involving clouds, radiation and vertical diffusion are also important but smaller in magnitude. The sum of the impacts of the dynamics, physical processes and the analysis increment represents the analysed evolution of the flow (Figure 9d). It is the magnitude of this evolution that is most appropriate to compare with the increments in the MCS regions. Comparison of Figure 8d with Figure 8c (which is plotted with a much smaller contour) suggests that the increments are typically about ⅓ those of the evolution. This ratio is probably too small when we consider the magnitudes of the precipitation errors. Similar results apply to the mid-tropospheric specific humidity budget.

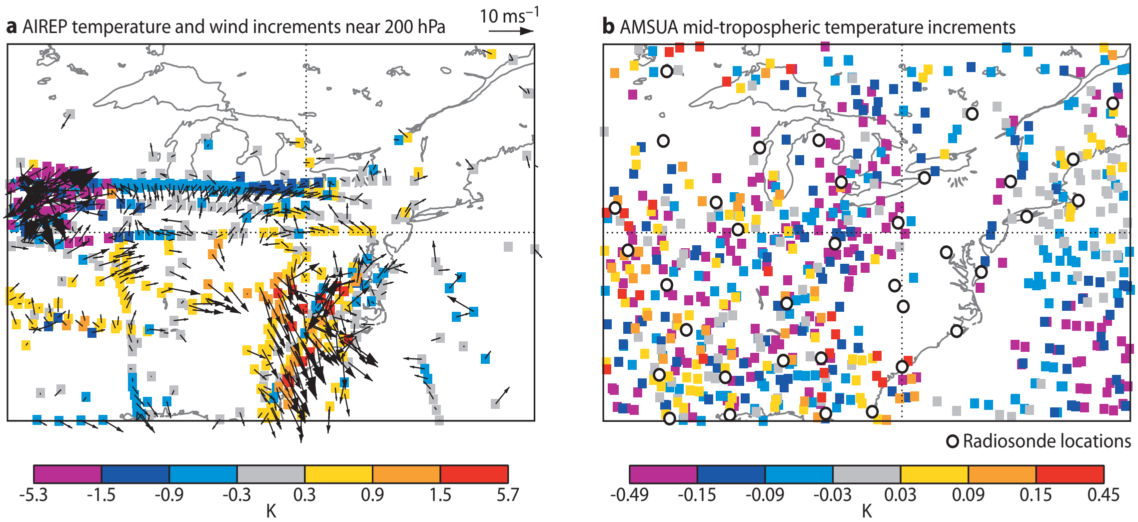


Figure 8 Analysis increments of temperatures and winds (interpolated to observation locations) during the production of the operational analysis at 00 UTC on 10 April 2011. (a) Aircraft observations near 200 hPa (185–215 hPa). (b) AMSUA microwave channel 5 (which measures mid-tropospheric temperatures). The radiosonde network is also indicated.

 doi:10.21957/ocaqy91i

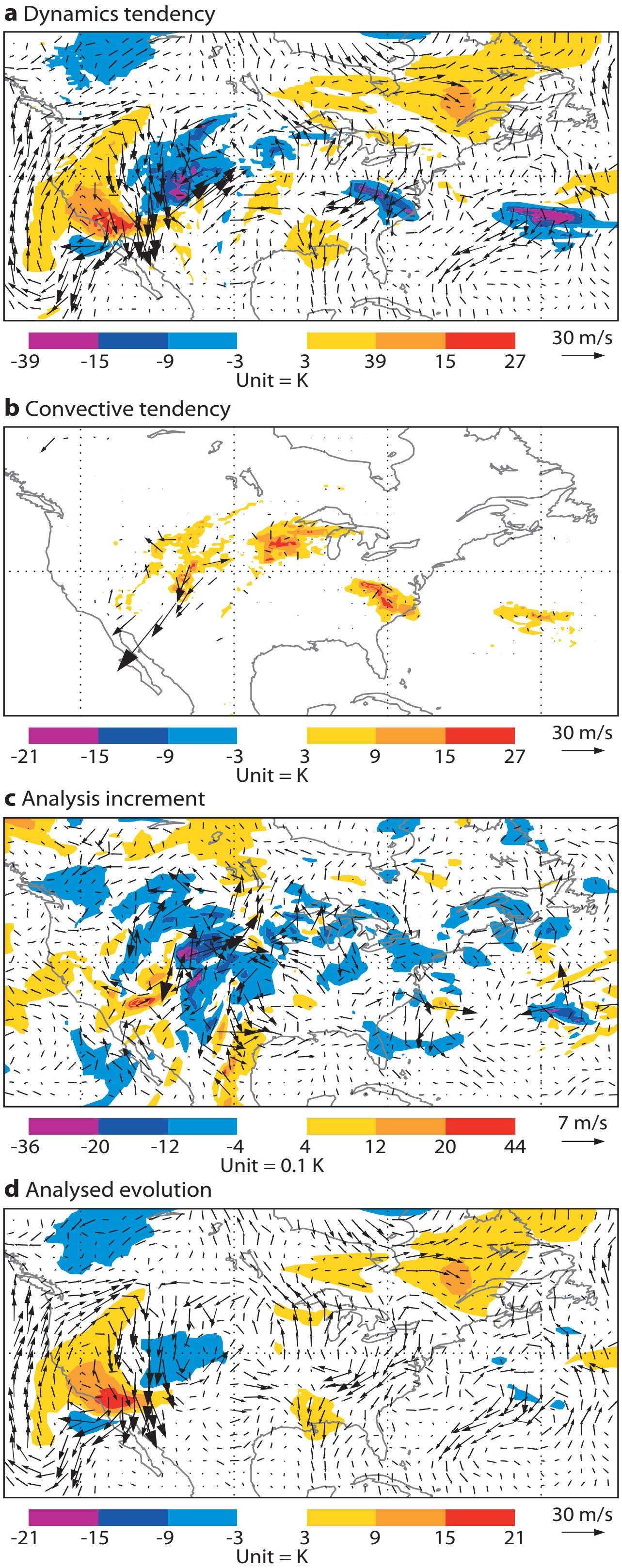


Figure 9 Diagnostics of 500 hPa temperature (shaded) and horizontal wind (vectors). (a) Dynamics tendencies integrated over the first 12 hours of the first-guess forecast started at 18 UTC on 9 April 2011. (b) Similar integrated tendencies from the convection scheme. (c) The analysis increment valid at the end of the 12-hour period. (d) The analysed evolution of the flow (the difference between the analysis at the start and end of the 12-hour period). See individual panels for contour intervals and reference vectors.

Summary of the investigation into mesoscale convective systems

Mesoscale convective systems are generally well predicted by the first-guess forecast, but precipitation accumulations can be in error by over 50% and locations can be offset. It is unclear at present whether these errors are adequately represented in the ensemble data assimilation system, and this could be a future area for research. The MCS events reduce the quantity of observational data available to the assimilation system. While 4D-Var does still produce analysis increments in these regions, the magnitudes of these increments in the mid-troposphere may be somewhat too small. Hence MCS events may act to degrade the analysis in addition to playing an active role in the evolution and chaos of the flow.

Sensitivity studies

The final part of the investigation into busts was to assess the sensitivity of busts to types of observations, flow-dependent background errors and other factors.

Sensitivity to types of observations

Here we diagnose the 'usefulness' of the observations when they are assimilated. To do this, a technique is used that quantifies the contribution to (global) 24-hour forecast error associated with each individual observation type (the so-called 'forecast error contribution', FEC). FECs were averaged over one week (6 April 2011 to 13 April 2011 – approximately 50% of this period is within the bust, and 50% outside the bust event). Hence the effect of chaotic improvement should be small. With the exception of PILOT data, mean FECs over the USA and north eastern Pacific were found to be negative – i.e. using the observations is reducing forecast error.

The mean forecast error contribution for AIREP data above 400 hPa is shown in Figure 10a. Over the USA – where the data density is high and thus the mean contribution is well quantified – these aircraft data are seen to decrease 24-hour forecast error. On the other hand, PILOT data over the USA (Figure 10b) often increases the 24-hour forecast error. For the USA, PILOT data are actually re-labelled radiosonde observations that provide additional information at 'significant levels', such as temperature inversions, and might be expected to be particularly difficult to reconcile with the first-guess forecast. In general PILOT data show weaker mean westerlies over North America than the other observation types. These results suggest that the impact of PILOT data should be the first candidate for further investigation.

An 'observation system experiment' was performed whereby PILOT data was denied (globally above 400 hPa) from the data assimilation system run at operational resolution from 1 April 2011. However, no reduction in the 10 April bust was found in the full non-linear forecast. Hence, while PILOT data could have a detrimental impact on our analyses in general, it does not appear to have a specific impact on the analysis that leads to the bust.

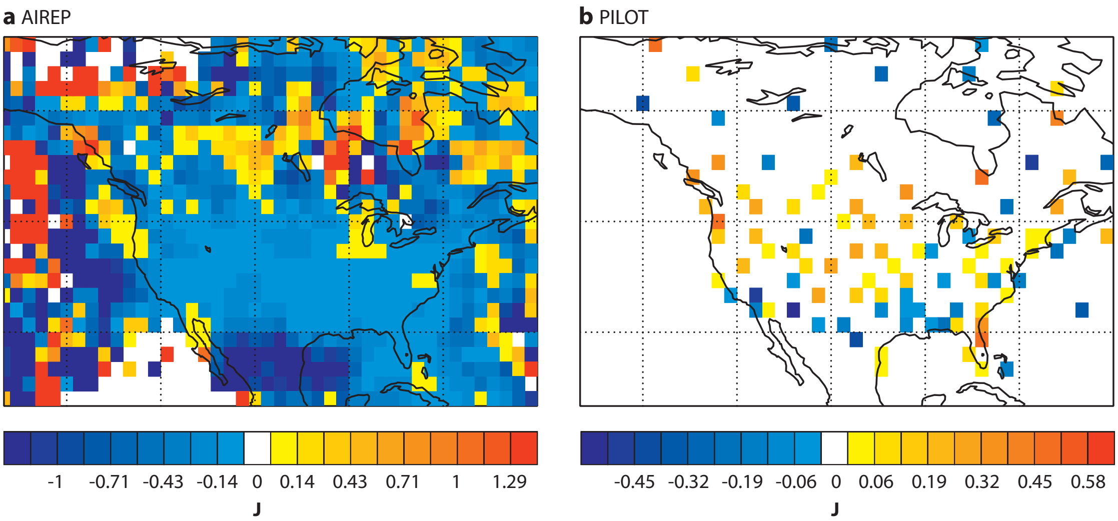


Figure 10 Contribution to 24-hour forecast error above 400hPa, averaged over the week 21 UTC on 6 April 2011 to 21 UTC on 13 April 2011 from (a) AIREP and (b) PILOT data types.

doi:10.21957/ocaqy91i

Sensitivity to flow-dependant background errors

When producing the analysis, the 4D-Var assimilation system requires knowledge of likely errors in observations and the background model. By altering the error covariances, the extent to which the analysis is drawn away from the background and towards the observations can be changed. In general, these error covariances are probably near optimal – as judged by average forecast performance. However, since it is possible that a poor trough/CAPE analysis is being perpetuated through the background (by systematic errors in the model for example), a sensitivity study was conducted whereby background error covariances were trebled for a few days leading up to the busts on 10 April and 10 May.

Although the effect of chaotic improvement cannot be discounted, Figure 11 shows that the Z500 European ACC at day 6 is moderately improved around the times of both busts – especially for the 10 May case. Since 18 May 2011, the Ensemble of Data Assimilations (EDA) has been used to incorporate flow-dependence into the background errors. Although the use of the EDA (in its default configuration) did not improve the 10 April bust either, further investigation of the EDA-generated background errors in these trough/CAPE situations over the USA would be beneficial.

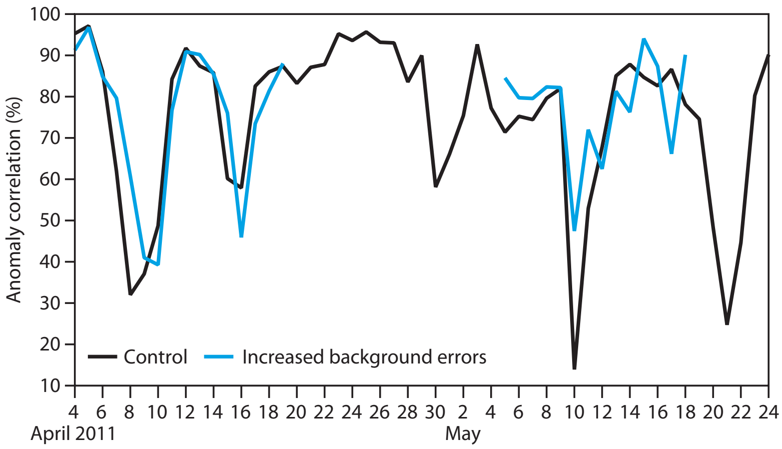


Figure 11 ACC for day-6 forecasts of Z500 for Europe for T511 forecasts initiated from T511 analyses generated with control, and trebled, background error covariances. A 12-hour data assimilation window was used.

Sensitivity to other factors

The previous result suggests that, if the observations are given more weight, this improves the forecast (particularly for the second bust case around 10 May). However, the complexity of this situation, and the extreme sensitivity to the initial conditions is illustrated by a somewhat contrary result. The denial of all data over land within a single assimilation cycle led to a markedly reduced bust on the 10 April (by more than a factor 2 in European ACC and RMSE at both day 5 and day 6). Figure 12 shows the change in initial conditions when the data over the land was denied. It shows a strengthened trough (somewhat consistent with the best ensemble member, Figure 6b) although it does not strengthen the ridge. Further investigation will ascertain if this is a case of chaotic improvement or otherwise.

We have also considered the possibility of sensitivities to the formulation of the model's dynamical core. The IFS and the UK Met. Office's Unified Model (UM) differ substantially in several aspects of their dynamical cores. The IFS version used operationally at ECMWF is a spectral model that solves the hydrostatic primitive equations, whereas the UM is a latitude-longitude grid-point model that solves the non-hydrostatic, deep atmosphere equations. Moreover, there are differences in the numerics associated with the vertical discretisation and the time marching scheme as well as in the coupling to the physical parametrizations. Hence an extensive series of experiments eliminating some of the differences in the dynamics between the UM and the IFS was conducted. The spring busts were found to be insensitive to any of these changes except that the April 10 bust was less severe with a change to the 'implicitness' of the vertical diffusion scheme in the boundary layer. Again, this one example could be associated with chaotic improvement. Based on these results, it is concluded that the dynamical core formulation is unlikely to be the key factor for the occurrence of the forecast busts.

Based on coincidences with the 10 April bust case, other work has investigated possible links to the El-Niño-Southern-Oscillation, the Madden-Julian Oscillation, and to Rossby wave forcing by the Tibetan Plateau. While all of these features undoubtably contribute to forecast error in general, no strong correlation was found with the incidences of European bust forecasts over the last decade or more. It has also been noticed that the Rockies trough leads to strong winds over the Sierra Madre mountains of Mexico, and the generation of gravity waves. Improved diagnostic techniques, and their application to other flow regimes, would be required to quantify the impact of these waves on the busts.

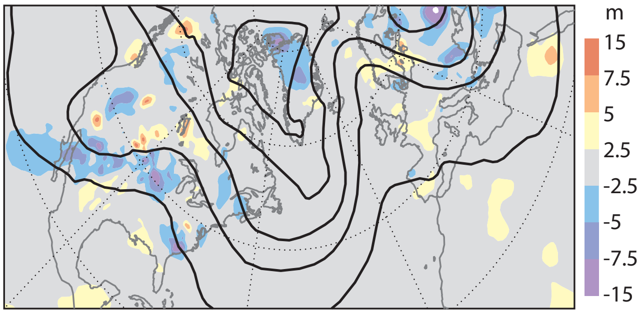


Figure 12 Z500 analysis at 00 UTC on 10 April 2011 when observations over land are denied from the data assimilation system for a single analysis cycle. Shaded: difference from control analysis. Contoured: full field with contour interval 200 m.

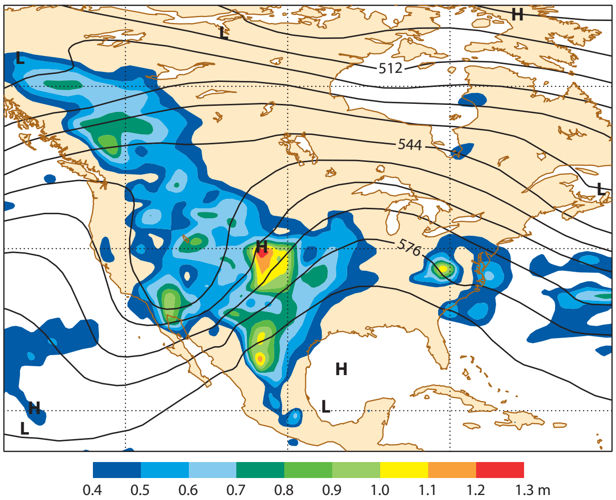


Figure 13 Standard deviation of Z500 for the EDA 12-hour forecasts from 12 UTC on 9 April 2011.

Future directions

Poor medium-range forecasts for Europe, as defined in the companion article, are an order of magnitude rarer now than they were 20 years ago. Nevertheless, even a single 'bust' is not good for our users, and has a significant impact on our seasonal-mean scores.

If past trends also predict future improvements, then general development of the IFS, with no focus on forecast busts per se, may continue to reduce bust frequency. However, if we wish to specifically target the bust issue then the above results can help inform decisions about future work.

The results suggest that the initial conditions in trough/CAPE situations need to be more accurate to reduce bust frequency or severity in our forecasts.

- The primary reasons for first-guess errors in the trough need to be identified. Are they associated with Tibetan wave forcing errors (passed through several assimilation cycles as the waves propagate across the Pacific) , or do the available observations over the Pacific and baroclinic instability in the Pacific storm-track obliterate this effect? Are there systematic errors in (Rossby) wave speeds or magnitudes over the Pacific?
- Do we have accurate enough, and abundant enough, observations to constrain circulation structures over the USA like those of the best ensemble member perturbation? (Note that the magnitude of the best ensemble member perturbation is similar to the standard deviation of Z500 errors from radiosonde observations over the USA.) Special focus could be placed on observations that are particularly important for the analysis over the USA– aircraft data for example. The role of PILOT data (for the USA, this is radiosonde observations at significant levels) on analyses in general could be investigated further. The analysis experiment that denied all data over land will be refined and extended to larger samples.
- There is an indication that background errors, which are optimal for general forecast performance, might not be optimal in this particular trough/CAPE flow regime. The IFS now uses the standard deviation of EDA first-guess forecasts to estimate the flow-dependence of background errors. Figure 13 shows this standard deviation in 12-hour forecasts started from 12 UTC on 9 April. Enhanced uncertainties do occur in the central USA region highlighted by the best ensemble member (Figure 6b), although they are weaker. Note, however, the lack of significant uncertainty in the region of the MCS to the west of the Great Lakes. One reason for this may be because sub-gridscale uncertainty associated with the

doi:10.21957/ocaqy91i

triggering of convection is not represented by the stochastic scaling of tendencies at present. Investigation of more cases of trough/CAPE situations will confirm whether the new EDA-generated background errors adequately represent local uncertainty in these convective situations.

The roles played by convection suggest that further investigation of MCSs would be beneficial.

- Can we improve the (deterministic) prediction of convection in the first-guess forecast? The representation of MCS convection has already improved a lot over the last decade. Further improvements in predicting the location and intensity of convection in the first guess forecast (let-alone at longer lead-times) will continue to be a challenge. Future increases in computing power will present new opportunities as we start to resolve the convection within MCS events.
- Can we improve the data assimilation in MCSs – present in the observations and/or first-guess? Do we have the necessary observations to constrain the convection? How representative are observations of model grid-box-mean values in these convective situations? How is the MCS represented at each iteration within the data assimilation? How well does the linear physics represent such extreme precipitation events? How can we reduce data rejection and improve the impact, throughout the troposphere, of non-rejected data?

For flow situations where forecast error is substantially different from its mean value, as in the case of the trough/CAPE regime, it is important to assess whether EPS spread in the medium-range adequately reflects the change in likely error.

- Can we develop diagnostics that better assess the flow-dependent spread-error relationship in the face of short datasets (due to frequent system updates) and the annual cycle in predictability?

This study has highlighted reasons for, primarily springtime, European forecast busts.

- Can autumn busts be explained by a similar trough/CAPE situation that arises over the North Atlantic when a tropical cyclone transitions into the extratropics and happens to encounter an upper-level trough (*Jones et al*., 2003)?

The conclusion of this study is that more accurate initial states around the Rocky mountains, and improvements in the assimilation and forecasting of mesoscale convective systems over North America, will be necessary to decrease the frequency of European medium-range forecast busts, particularly in spring. Indeed, it is also likely that much of the strong reduction in the frequency of these busts over the past decades must be attributed to improvements in these two aspects. However, due to the chaotic nature of the atmosphere, with flow states whose evolution is highly sensitive to the accuracy of the initial state, we may never be able to completely eliminate busts in the future.

Further reading

Grazzini, F. & **L. Isaksen,** 2002: North America Increments – a problem in 2002. *ECMWF Tech. Memo. 674*

Jones, S.C. & **Coauthors,** 2003: The extratropical transition of tropical cyclones: Forecast challenges, current understanding, and future directions. *Weather and Forecasting*, **18**, 1052–1092.

doi:10.21957/ocaqy91i