A relationship between busted European forecasts, a Rockies trough, and storms over eastern North America suggests the importance of improving quality and use of observations, model depiction of convective systems, and representation of uncertainties.

In common with other weather centers, the European Centre for Medium-Range Weather Forecasts (ECMWF) produces a “single high-resolution forecast” and an “ensemble of lower-resolution forecasts.” The former represents, on average, the most accurate single forecast, while the latter provides a measure of the uncertainty in the outcome. A standard approach to assessing the performance of a single forecast is to calculate how well the height of the 500-hPa pressure field (Z500) agrees with the observed outcome. The agreement can be quantified by, for example, calculating the area-averaged root-mean-square error (RMSE) or the spatial anomaly correlation coefficient (ACC). Figure 1 shows ACC time series of day 6 forecast skill over Europe from some of the world’s NWP centers (within the TIGGE program): Met Office (UKMO), Japan Meteorological Agency (JMA), Canadian Meteorological Centre (CMC), and NCEP. The dates correspond to the start of the forecast. The score shown is the spatial ACC of Z500. Europe is defined in this article as the region 35°–75°N, 12.5°W–42.5°E.
series of Z500 over Europe at a lead time of 6 days for single forecasts started at 0000 and 1200 UTC by several of the world’s weather prediction centers. In general, scores fluctuate about the 80% level, but occasionally, as between 7 and 10 April 2011, there is a strong drop in performance. All centers suffered during this episode, with the Met Office recovering earliest.

While occasional poor forecasts also occur in other regions of the world, we focus on Europe and define a “forecast bust” to be an occasion when the day-6 high-resolution forecast of European Z500 has an RMSE greater than 60 m and an ACC less than 40%. These two conditions ensure that a bust is associated with errors of a sufficient magnitude and also involve a pattern or phase discrepancy. The forecast from 0000 UTC 10 April 2011 did represent such a bust event for ECMWF.

Over the years, significant progress has been made in reducing the frequency of busts. Figure 2a (red curve) shows that annual totals of busts for the ECMWF operational high-resolution forecast have decreased from around 70 per year in 1990 to around 5 in 2011. However, even this low level of busts causes problems for users of NWP products and has a significant impact on seasonal means of daily scores. Figure 2b (red curve) shows that busts occur throughout the annual cycle.

Here, we aim to understand the nature of these busts. Are they associated with a particular flow regime? Can we reduce errors in the initial conditions or forecast model? Do we need to improve uncertainty estimates in our ensemble predictions?

Previous studies (e.g., Grazzini and Isaksen 2002) have investigated ECMWF bust events on a case-by-case basis using single forecasts. Although these have been informative, it is difficult to draw clear conclusions from single cases. For example,

**Fig. 2.** The number of European forecast busts at day 6 from the ECMWF operational high-resolution forecast and forecasts made within the ERA-Interim project, for 1989–2011. (a) Time series of annual totals. (b) Mean annual cycle. Results are based on forecasts started at 1200 UTC since operational high-resolution forecasts were not made at 0000 UTC before 2001. Since operational annual totals have decreased, the operational mean annual cycle will emphasize the early part of the record. In addition to improvements to the forecast model and improvements in the processes used to derive its initial conditions, the resolution of ECMWF’s high-resolution forecast increased steadily over the period, from 190 km in 1989 to 16 km in 2011. ERA-Interim is based on a fixed IFS cycle that was operational in 2006 and has a fixed resolution of 80 km. The two curves in (a) should not necessarily intersect in 2006 since the high-resolution forecast had a resolution of 25 km at that time.
there may be other plausible explanations associated with teleconnections and instabilities of the flow. In addition, we have to accept some inherent unpredictability associated with the growth of chaos. Hence, the first approach taken here is to make a large composite of several hundred bust events with the hope of identifying common features. Composites of ensemble (rather than single) forecasts also allow us to estimate predictability and to relate this to forecast error. Alongside the composite investigation, we look at single and ensemble forecasts made from the initial conditions at 0000 UTC 10 April. Importantly, this is the last initial time for which the ECMWF high-resolution forecast displayed a bust in the episode highlighted in Fig. 1. Salient errors and uncertainties associated with the poor forecasts at day 6 are therefore confined to the first 12 hours of this forecast and its initial conditions (since the next forecast had fully recovered). We show how this “10 April event” relates to the large composite, before diagnosing it in more detail. It is recognized that, unless sensitivity experiments on a poor forecast are well motivated or justified, score improvements could simply reflect “regression to the mean” (Galton 1886) and not real improvements. However, there are aspects (e.g., associated with the use of observations in the analysis) that will be applicable to all such forecasts. Note that further results on this topic of occasional poor forecasts are presented in Rodwell et al. (2012a,b).

**THE BUST SITUATION.** It would be most beneficial to understand busts in recent cycles of the ECMWF Integrated Forecasting System (IFS; see, e.g., Jung et al. 2012), but this means that there are very few busts to investigate. The approach taken here is to first use forecasts made within the Interim ECMWF Re-Analysis (ERA-Interim; Dee et al. 2011). Figure 2a (blue curve) shows that bust frequency for the ERA-Interim forecast only decreases slightly over the last 22 years, owing to the use of a fixed IFS cycle. Using all 22 years worth of data from this stable forecasting system allows us to characterize the busts. Later, we check whether these characterizations are still valid for more recent IFS cycles.

To characterize the situation most clearly associated with busts, a “bust composite” is made using all 584 dates for which the ERA-Interim single forecast had a European bust during the period 1 January 1989–24 June 2010 (more recent dates are left for analysis in the next section). Figure 3 shows the Z500-mean verifying analysis for the bust composite (an “analysis” is our estimate of the true state of the atmosphere; see the figure caption for more discussion). Bold colors indicate mean values that are statistically different from zero at the 5% significance level. Despite only defining busts by their gross scores, it appears that there is a particular verifying analysis associated with many busts (including a dipole pattern with high-pressure over northern Europe and low-pressure over western Europe).
and a low over the Mediterranean) somewhat like the blocking situation discussed by Rex (1950). This flow situation is also reminiscent of the action of anticyclonic wave breaking (as described by Thorncroft et al. 1993) within a Rossby wave train stretching back across the Atlantic.

We can use the same bust composite to search for the key features in the initial conditions of these poor forecasts. Figure 4a shows that the statistically significant features in the Z500 mean initial conditions of the bust composite are not over Europe but include a “Rockies trough” embedded in an apparent Rossby wave covering the United States, and a northern “Canada high.” Results discussed later suggest that the Canada high does not play a major role in the busts, and so we concentrate here on the Rockies trough. Over northern Europe, there is a weak and statistically insignificant low center in the mean initial conditions. This might indicate that busts are often associated with a particular difficulty in developing the European block 6 days later.

Grazzini and Isaksen (2002) linked busts to mesoscale convective systems (MCSs) over the United States, particularly around the Great Lakes region. Figure 4b shows the convective available potential energy (CAPE) in the mean initial conditions of the bust composite. While each grid point is not individually statistically significant, there is a coherent region of increased CAPE over the United States that stretches north to the Great Lakes and into southern Canada.

When the bust initial conditions are stratified by season, we find that the Rockies trough extends farthest south in autumn, while the strongest and...
most widespread positive CAPE anomaly occurs in summer (90 J kg\(^{-1}\)). In spring, both aspects are clearly present, with the strongest Rockies trough (35 m; which remains statistically significant) and with positive CAPE anomalies penetrating north to beyond 40°N along the eastern side of the Rockies.

The Rockies trough and positive CAPE anomaly also occur in the 10 April event of spring 2011. Figure 5a shows Z500, CAPE, and 850-hPa wind anomalies from ERA-Interim for the April 10 forecast initial conditions. The similarities with the composite initial conditions (Figs. 4a,b) are striking. Ahead of a Rockies trough anomaly, there are low-level southerlies transporting heat and moisture from the southern United States and Gulf of Mexico, providing the environmental conditions (CAPE) for the development of storms. Indeed the precipitation observed over the United States (Fig. 6b) highlights two MCSs that coincide with the CAPE anomalies. One of these is to the west of the Great Lakes (up to 30 mm of observed precipitation and, incidentally, associated with numerous tornado reports), and the other is near the east coast of the United States (up to 50 mm of observed precipitation). The other panels in Fig. 6 are discussed later. Note that the Canada high anomaly identified in Fig. 4a is most clear in the winter and summer composites, and does not appear in the 10 April event.

**PREDICTABILITY AND FORECAST UNCERTAINTY.** The ECMWF ensemble of forecasts (ENS) represents a natural framework for investigating predictability. It includes a set of 50 forecasts, each with perturbed initial conditions to represent the effects of deterministic chaos (Lorenz 1963) and perturbations to the forecast that reflect sub-model grid-scale uncertainty. The perturbed initial conditions are derived from a 10-member ensemble of data assimilations (EDA; Isaksen et al. 2010), with additional “singular vector” initial perturbations to represent fast-growing errors. To represent model uncertainty throughout the forecast, stochastic (i.e., “random”) perturbations are applied to the combined physical tendencies (SPPT; Shutts et al. 2011) and the kinetic energy that is dissipated at the subgrid-scale is “backscattered” onto the resolved flow (Berner et al. 2009).

Lower predictability should be reflected with increased spread between ensemble members. Indeed, the average distance of ensemble members from the ensemble mean (e.g., the ensemble standard deviation) is a prediction of the distance of the truth from the ensemble mean (i.e., a prediction of the ensemble-mean error). When averaged over many forecast start dates (and taking account of the finite size of the ensemble), the mean spread should match the mean error (Leutbecher and Palmer 2008). Here, we assess the spread–error relationship for forecasts started from initial conditions that include the trough/CAPE regime. Do we find a situation of large mean spread correctly predicting large mean error? Alternative possibilities include a situation where the error is not
particularly large or where the spread is too small to match the error. For a single start date, this latter situation might be classified as an “ensemble bust.”

To identify trough/CAPE situations, we project the 0000 and 1200 UTC operational high-resolution analysis anomalies of Z500 and CAPE onto the patterns within the highlighted regions shown in Fig. 4. We focus on more recent IFS cycles by using the period 10 November 2010–20 March 2012 and select dates for which the trough has a projection coefficient greater than 3 and the CAPE has a projection coefficient greater than 1. This means that, if a Z500 analysis anomaly had exactly the same spatial pattern as that shown in the Rockies trough box in Fig. 4a, it would need to have 3 times the magnitude. The CAPE threshold was set lower because Fig. 4b indicates small-scale uncertainties in the pattern. Results are not sensitive to the precise choices of these projection thresholds. Using this approach, 84 date/times are selected. These tend to be concentrated in spring and include, incidentally, the 10 April event.

Figures 7a,c show the mean ensemble spread and mean ensemble-mean error for the 84 dates/times in the “trough/CAPE composite.” Figures 7b,d show corresponding background spread and error, respectively, for days when at least one of these projection thresholds was not exceeded. For the background composite, mean spread does match well the mean error. Indeed, the ENS is somewhat tuned to ensure this. Consistent with the bust composite, errors for the trough/CAPE composite are enhanced relative to the background (by around 30% over the North Atlantic and Western Europe). This increase is predominantly associated with random errors (rather than mean errors), suggesting reduced predictability. Spread is also enhanced a little, particularly around Iceland, indicating that the ensemble is forecasting some (but not all) of the reduced predictability associated with the trough/CAPE situation. (Using a similar approach, it was found that the Canada high feature in Fig. 4a does not lead to increased forecast error.)

Note that the day-6 mean flow of the forecasts in the trough/CAPE composite does include the

![Fig. 7. Spread and error results for Z500 at day 6 from the ensemble predictions.](image-url)

Here, spread is actually the ensemble standard deviation, scaled slightly to account for finite ensemble size. To ensure a fair comparison of composites, spread and error, for each month of the year, are given the same weight in the background composite as they are given in the trough/CAPE composite. The use of dates after 10 Nov 2010 ensures independence from the ERA-Interim composites and avoids a major change to the initialization of the ENS on 24 Jun 2010 (with the commencement of the EDA) and subsequent changes on 9 Nov 2010 (which affected our representation of uncertainty estimates).
north European blocking high. The difficulty with predicting the north European high is well supported by dynamical studies of flow mechanisms in connection with blocking situations. For example, Mauritsen and Källén (2004) demonstrated that ECMWF ensemble spread systematically increases a few days ahead of the onset of a European block, thus indicating a bifurcation point with increased sensitivity to the initial conditions. One can perhaps think of this situation as being close to the body of the “butterfly” in the diagram of Lorenz (1963). Since the Rockies trough is an integral component of the negative Pacific–North American (PNA) pattern, the present findings are also consistent with the results of Corti and Palmer (1997), who demonstrated that the amplitude of the PNA pattern over North America influences the onset of European blocking.

The 10 April case is also associated with increased forecast uncertainty. Figure 8 shows, for the first 12 days of April 2011, the Z500 European spatial ACC for the operational high-resolution forecasts from ECMWF and from the Met Office’s Unified Model (see, e.g., Tang et al. 2012), along with those for each member of the ECMWF ENS. 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 ENS and there is even an ensemble member score matching that of the Met Office high-resolution forecast, which for this case recovered earliest. Since both high-resolution 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. (Indeed, there are cases when the IFS suffers least.)

In the real world, reduced predictability indicates enhanced probability divergence in state space. Similarly, in the model world, increased forecast errors and ensemble spread could be associated with initial condition errors that grow quickly, but they could also be associated with model errors. Below, we investigate these possibilities.

**KEY INITIAL CONDITION UNCERTAINTIES.** In this section, we consider the possibility that the errors are associated with uncertainties in the initial conditions. Figure 9a shows the day-6 Z500 errors for the ECMWF high-resolution forecast from 10 April. Consistent with the European bust, there are large errors over the eastern North Atlantic and Europe. 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. A clear progression is found with errors tracing back toward North America, as shown for day 2 in Fig. 9b.

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, we look at the 50 ensemble members of the ENS. 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 day-6 Z500 RMSE over Europe 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.

We have further isolated the key initial condition perturbations of the best ensemble member by turning off the model uncertainty aspect and progressively confining the initial condition perturbations to ever-smaller domains. Figure 9c shows the result of this process. The key perturbations have been confined to the North American/eastern North Pacific region. They highlight a strengthening (on the order of 5%–10%) of the Rockies trough and the downstream ridge (and presumably increased CAPE).
The day-6 errors for the forecast initiated with this perturbation (Fig. 9d) are indeed reduced over the eastern North Atlantic and Europe compared to the control (Fig. 9a). The above results are consistent with comparisons with the Met Office forecast for the 10 April event. For example, a strong sensitivity to the initial conditions was evident when the initial conditions of the two forecasts were swapped. The day-6 Z500 error pattern of the ECMWF forecast was reproduced when the Met Office model was initiated with the ECMWF analysis and vice versa. Further, the importance of the central North American region was evident from the day-1 differences in the two operational high-resolution forecasts. This difference, which highlights the fast-growing modes, was strikingly similar to the difference at day 1 between the “best perturbation” forecast and the control.

By identifying a key perturbation structure (Fig. 9c), we have been able to go further, for the 10 April case, than the general characterization. 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. Nevertheless, we do obtain an idea of the spatial scale and magnitude of the key initial condition uncertainties associated with springtime busts.

The EDA is designed to estimate flow-dependent uncertainty in our knowledge of the true state of the atmosphere. Consistent with the best ensemble member results, the standard deviation of the EDA analyses highlights Z500 uncertainties in the same mid–North American region. Over the next few hours, the spread of EDA short-range “first guess” forecasts show how uncertainties in temperatures at 200 hPa (T200) develop from this region along the jet so that, by 0600 UTC 10 April, they predict enhanced uncertainties associated with the MCS event to the west of the Great Lakes (Fig. 10a). Both initial condition and model uncertainties (SPPT only) are represented in this ensemble of forecasts. These large uncertainties are consistent with the results of Zhang et al. (2002), who highlight the rapid growth of mesoscale forecast errors when moist processes are active. Figure 10b shows the equivalent analysis spread of the subsequently generated EDA analyses for 0600 UTC 10 April. The fact that the spread of the analyses is less than that of the short forecasts is consistent with the desire for the assimilation of new observations to improve our knowledge of the true state. However, the general underrepresentation
of ENS spread at day 6 in trough/CAPE situations (Fig. 7) raises the possibility of further optimization of flow-dependent uncertainty in the EDA and subsequent uncertainty representations within the ENS (singular vectors or model uncertainty).

While we have to accept some enhanced forecast uncertainty in these trough/CAPE situations, there may still be scope to reduce forecast error if we can produce more accurate initial conditions or improve the model. These possibilities are discussed in the next two sections.

**IMPROVING THE INITIAL CONDITIONS.**

A reviewer’s analysis links the strength of the Rockies trough on 10 April to a previous cyclogenesis event over the Pacific. While it is interesting to understand the processes by which a strong Rockies trough can come about, our main aim is to produce a better analysis of it (and of the associated MCSs). Uncertainties in upstream processes could, of course, lead to increased background uncertainties in the Rockies trough for the data assimilation. Although we would wish to decrease these background uncertainties, trying to do so here would simply shift the focus of the study farther upstream and to longer forecast lead times, where statistical significance is harder to obtain. Instead, we focus on the effectiveness of observations (particularly observations over North America) to correct such background errors.

Different observation types have different impacts on the analyses. Diagnostic techniques have been used to assess which observation types are most useful and which, if any, degrade the forecast. One technique uses the adjoint of the tangent-linear model. When applied to the week including the 10 April event, only the observation type comprising U.S. radiosonde data at significant levels (e.g., inversions) led to a general degradation of day-1 high-resolution forecast error. (The U.S. radiosonde network is indicated by open circles in Fig. 11.) However, a subsequent nonlinear “observation system experiment” demonstrated that removing this significant-level data did not improve the 10 April bust. While our results suggest that all observation types tend to help rather than hinder the forecast, it is still possible that they could help even more.

With this possibility in mind, we have experimented with increasing the weight given to observations in the data assimilation by increasing the specified background errors. (These errors are probably near optimal, as judged by average forecast performance, but may be less optimal in trough/CAPE situations.) However, results showed that the 10 April event was little changed. (A bust on 10 May 2011 was reduced but, without further experimentation, we cannot discount the role of regression to the mean.)

Another relevant aspect is that the meteorology of the day has a strong impact on the availability of observations and, where present, their utility within the data assimilation. For example, aircraft tend to avoid MCS events, and it is generally not possible at present to assimilate satellite observations in cloudy conditions. Figure 11 demonstrates the lack of
assimilated midtropospheric microwave temperature observations within the two MCS events (indicated by green contours). The result is that there are far fewer in situ observations available to correct MCS errors in the short-range first-guess forecast used in the data assimilation. This is consistent with the large standard deviation seen in the EDA analyses in Fig. 10b. Work at ECMWF is focused on the better assimilation of cloud-affected satellite observations, although it may be difficult to persuade pilots to routinely fly through strong convective events.

Another important issue is whether the observations we do have are accurate enough to constrain structures like that of the best initial perturbation (Fig. 9c). For example, it appears that the standard deviation of Z500 errors from radiosonde observations over the United States have a similar magnitude to that of the best initial perturbation.

One curious result requires further investigation; not using any data over the land for a single assimilation cycle did improve the 10 April event. The effect of this change on the initial conditions, which could be by chance, had some similarities to the best initial condition perturbation (Fig. 9c).

**IMPROVING THE MODEL.** We have seen that MCSs are linked, one way or another, to European forecast busts. Whether they are just symptoms of a developing (adiabatic) bust situation or whether the diabatic physics (convection, etc.) embodied within the MCS plays and active role in the busts needs to be determined. Hence, before considering model improvements per se, we first investigate whether there is likely to be a significant role for model physics in the development of bust forecasts.

For the reasons discussed in the introduction, we look to the trough/CAPE composite (rather than individual bust cases) and assess whether diabatic processes have the potential to significantly modify the large-scale flow. To do this, the composite potential vorticity (PV) budget on the 330-K isentropic surface (approximately at 250 hPa) is calculated. The budget is based on analyses so that it is as model independent as possible. The aim is to assess whether the time evolution of the Rockies trough (and the Rossby wave feature in which it is embedded) is consistent with simple adiabatic advection or whether diabatic/frictional processes (convection, radiation, diffusion, etc.) are essential.

Figure 12a shows, contoured, the anomalous PV calculated using all 95 trough/CAPE events in the period 25 June 2010–20 March 2012. Vectors show the anomalous horizontal wind. The trough over the Rockies is clearly evident as a positive PV anomaly. The ridge over the eastern United States is also evident. Shading in Fig. 12a shows the anomalous local time tendency of PV. The ridge is strengthening to its northeast, while little tendency is evident on the leading (eastern) edge of the trough. Shading in Fig. 12b shows the anomalous adiabatic advection of PV and vectors show the full horizontal winds of the trough/CAPE composite. Since the advection anomalies occur to the east of PV anomalies of the same sign, this advection clearly acts to propagate the Rossby wave eastward and accounts for much of the local time tendencies found in Fig. 12a. Nevertheless, there are nonnegligible differences, and these must be attributed to the combined effects of anomalous diabatic and frictional processes.

The inferred diabatic plus frictional PV tendency (i.e., the difference between Figs. 12a, b) is shown in Fig. 12c. This term is seen to oppose the adiabatic advection term and thus slows down the eastward propagation of the wave and, indeed, virtually halts the eastward propagation of the leading edge of the trough. The term includes diabatic advection, diabatic changes in stratification, diabatic tilting, surface friction, and turbulent mixing. The negative values seen to the east of the Rockies in Fig. 12c (where Fig. 4b shows strong positive CAPE anomalies) may well be
due to the “destruction” of PV above a maximum in (mesoscale) convective heating, associated with stratification changes. Frictional effects on southerly flow along the eastern flanks of the Rockies would be more likely to lead to positive vorticity forcing and so are probably of secondary importance on this level. Based on these results, it seems likely that MCSs (and the diabatic physics in general) do play an active role in the evolution of the synoptic-scale trough: they slow it down. Interestingly, the 10 April case study is consistent with this result. The Rockies trough is embedded in a Rossby wave that has crossed the Pacific and which slows down over North America.

How quickly errors develop in the MCSs will affect the time scale at which they start to affect the larger-scale forecast. The precipitation produced by the short-range first-guess forecast used in the data assimilation (Fig. 6a) appears to be in qualitative agreement with the observations (Fig. 6b). However, there are large differences (Fig. 6c), even at this short range, of over 25 mm, associated with location and intensity errors. The relative contributions to this precipitation error associated with initial condition error, model physics error, and the limited predictability of the MCSs themselves remains unclear. Indeed, it is possible that our best initial condition perturbation is best because it “pre-corrects” for errors in the modeling of the MCS. Nevertheless, these results do confirm that the physics is, in general, an active (nonnegligible) component in the evolution of the bust forecasts.

What is the scope for model improvement? In the past, precipitation associated with MCSs in the ECMWF model was “stratiform” in nature (Grazzini and Isaksen 2002). Now it is largely associated with parameterized convection. This change is thought to be an improvement since present model resolution (16 km) is still too coarse to resolve real convective fluxes. A challenge for the future, as model resolution improves.

**Fig. 12.** Mean PV anomaly (contoured) and anomalous PV budget terms (shaded) for the trough/CAPE composite on the 330-K isentropic surface in operational high-resolution analyses. (a) The local time tendency calculated as the central difference of analyses displaced by ±6 h, together with anomalous horizontal winds. (b) The adiabatic advection of PV calculated using IFS spectral transforms, together with full horizontal winds from the composite. (c) The diabatic plus frictional PV tendency, deduced as the difference (a) minus (b). The composite is the mean of all 95 trough/CAPE situations at 0000 or 1200 UTC during the period 25 Jun 2010–20 Mar 2012. (We extend the period back further than for Fig. 7 since the budget is based on analyses, which were not affected by the changes to the ENS). Anomalies are relative to a climatology made using the same days of the year, from the preceding three years. All components were calculated at the full analysis resolution before a spatial filter on total wavenumbers greater than 10 was applied. PV anomalies are contoured with interval 0.4 PVU and with contours smaller or equal to −0.2 PVU dashed. Statistical significance at the 5% level is indicated with bold colors, black vectors, and contours.
increases further, will be to move the model back to a state where more of the precipitation is legitimately associated with resolved fluxes.

Further diagnosis could focus on the realism of the best ensemble member throughout the first 6 days of the forecast. Does this forecast predict the observed MCS precipitation particularly well or does it display errors, the day-6 effects of which were precorrected by the perturbation to the initial conditions?

The tangent-linear model and its adjoint are used within four-dimensional variational data assimilation (4DVar) to link model state perturbations at the start of the assimilation window with observation departures throughout the window. To obtain the tangent-linear model, simplifications are made to the nonlinear physics, discontinuities are smoothed, and processes are linearized. While the tangent-linear model is extensively tested (Janísková and Lopez 2012), do the assumptions of linearity break down in fast-maturing instability events such as MCSs? In the data assimilation, the tangent-linear model is also run at a lower resolution than the nonlinear model. These issues could have implications for the structures and magnitudes of the analysis increments applied.

The sensitivity of the 10 April forecast to many aspects of the model, such as its dynamical core, has also been investigated. However, nothing led to an improvement in the bust except for when a change was made to the vertical diffusion in the boundary layer. Again, further tests would be required to discount the role of regression to the mean for this single improvement.

**FUTURE DIRECTIONS.** Poor ECMWF medium-range forecasts for Europe are much rarer than they used to be, but even a single poor forecast is problematic for our users. All forecast centers suffer from these forecast “busts” or “dropouts.” At the National Centers for Environmental Prediction (NCEP), a group exists to investigate these forecasts (K. Krishna Kumar 2012, personal communication). The aim here has been to identify the nature of these forecast busts, the key processes involved, and ways to reduce their frequency or severity. By doing so, it is likely that we will also reduce the frequency of less severe errors.

Forecast busts are often associated with a difficulty to predict the onset of European blocking. The key initial conditions for these poor forecasts, particularly in spring, comprise a trough over the Rockies and mesoscale convection over eastern North America associated with high convective available potential energy (CAPE). Forecast errors from this “trough/CAPE” situation are predominantly random, possibly indicating reduced predictability. The ensemble predictions show some increase in spread and thus also indicate a reduction in predictability. It is possible that our method of compositing favored this spring-time situation. For other flow situations (e.g., the extratropical transition of hurricanes; Jones et al. 2003), key uncertainties may be more variable in location and lead to low predictability at more variable lead times.

Two issues are raised for future work. Can we reduce errors (through better initial conditions, models, etc.) and, at the same time, improve our representation of forecast uncertainty (through increased spread etc.)?

In terms of reducing errors, our results suggest that the biggest benefits may come from improving the availability, accuracy, and usage of relevant observations (particularly in cloudy conditions) and the representation of mesoscale convective systems (MCSs) in the nonlinear and tangent-linear models. Both aspects present significant challenges for the future.

In terms of improving the representation of forecast uncertainty, there is a need for better diagnosis of each aspect of flow-dependent uncertainty within the ensemble of data assimilations and the ensemble prediction system. The fact that, for the busts, initial errors appear to be mediated through and possibly magnified by MCS events, which involve subgrid-scale structures, highlights the importance of the representation of model uncertainty. Does the stochastic scaling of physical tendencies represent well the structures and magnitudes of true model uncertainty for MCSs? Would model uncertainty be improved by including the energy backscattering scheme into the EDA? Do the ensemble tangent-linear forecasts reproduce the spread of the nonlinear forecasts in these convective situations? How does the representation of observation error (Isaksen et al. 2010), background error (Bonavita et al. 2011), and model error (Trémolet 2007; Fisher et al. 2011) affect the flow dependence of spread?

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 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. Because of the chaotic nature of the atmosphere, with flow states whose evolution is highly sensitive to the accuracy of the initial state
and the model, we may never be able to completely eliminate busts in the future. However, with developments in our ensemble data assimilation and prediction systems, we should be better able to predict the increase in forecast uncertainty, something that is of great utility in itself.

ACKNOWLEDGMENTS. The authors thank Anton Beljaars, Mike Fisher, Elias Holm, Marta Janisková, Tim Palmer, and the anonymous reviewers of this article for their helpful and insightful comments.

REFERENCES


Rodwell, M. J., and Coauthors, 2012a: Characteristics of occasional poor medium-range forecasts for Europe. ECMWF Newsletter, No. 131, ECMWF, Reading, United Kingdom, 11–15.

—, and Coauthors, 2012b: A case study of occasional poor medium-range forecasts for Europe. ECMWF Newsletter, No. 131, ECMWF, Reading, United Kingdom, 16–24.


AMS MEMBERS

GIVE A GREAT GIFT AT A GREAT PRICE

Looking for the perfect present for the weather enthusiast in your life? Want to make a valuable contribution to your local library or community college? Send a subscription to Weatherwise magazine (6 issues) for just $24.95*—That’s nearly 50% off the list price!

Written for a general audience, Weatherwise offers a colorful and nontechnical look at recent discoveries in meteorology and climatology. Check out the latest table of contents at www.weatherwise.org.

Want your own?
Then order a personal subscription at the same great price.

Contact Member Services by e-mail at amsmem@ametsoc.org or by phone at 617-226-3998 to place all of your Weatherwise orders today!

*Cost for delivery outside of the U.S. is $40.95. Weatherwise is available to AMS Members through a cooperative agreement with Taylor & Francis Group LLC, the publishers of Weatherwise.