A route to systematic error in forecasts of Rossby waves

Article

Accepted Version

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

A route to systematic error in forecasts of Rossby waves

O. Martínez-Alvarado[1,2], E. Madonna[3,4], S. L. Gray[2] and H. Joos[3]

July 22, 2015

1. National Centre for Atmospheric Science–Atmospheric Physics, United Kingdom

2. Department of Meteorology, University of Reading, United Kingdom

3. Institute for Atmospheric and Climate Science, ETH Zurich, Switzerland

4. Oeschger Centre for Climate Change Research and Institute of Geography, University of Bern, Switzerland

Correspondence to:
Oscar Martínez-Alvarado
Department of Meteorology, University of Reading
Earley Gate, Reading, RG6 6BB, United Kingdom
E-mail: o.martinezalvarado@reading.ac.uk
Tel: +44 (0) 118 378 8951
Fax: +44 (0) 118 378 8905

Abstract

Recent work has shown that both the amplitude of upper-level Rossby waves and the tropopause sharpness decrease with forecast lead time for several days in some operational weather forecast systems. In this contribution, the evolution of error growth in a case study of this forecast error type is diagnosed through analysis of operational forecasts and hindcast simulations. Potential vorticity (PV) on the 320-K isentropic surface is used to diagnose Rossby waves. The Rossby-wave forecast error in the operational ECMWF high-resolution forecast is shown to be associated with errors in the forecast of a warm conveyor belt (WCB) through trajectory analysis and an error metric for WCB outflows. The WCB forecast error is characterised by an overestimation of WCB amplitude, a location of the WCB outflow regions that is too far to the southeast, and a resulting underestimation of the magnitude of the negative PV anomaly in the outflow. Essentially the same forecast error development also occurred in all members of the ECMWF Ensemble Prediction System and the Met Office MOGREPS-15 suggesting that in this case model error made an important contribution to the development of forecast error in addition to initial condition error. Exploiting this forecast error robustness, a comparison was performed between the realised flow evolution, proxied by a sequence of short-range simulations, and a contemporaneous forecast. Both the proxy to the realised flow and the contemporaneous forecast a were produced with the Met Office Unified Model enhanced with tracers of diabatic processes modifying potential temperature and PV. Clear differences were found in the way potential temperature and PV are modified in the WCB between proxy and forecast. These results demonstrate that differences in potential temperature and PV modification in the WCB can be responsible for forecast errors in Rossby waves.

Keywords

1 Introduction

Despite remarkable advances in resolution and formulation in recent years, facilitated by enhanced computer power, global numerical weather prediction and climate models still rely on parameterisations of diabatic processes and will continue to do so for the foreseeable future. These parameterisations introduce inaccuracies into models and contribute to the occurrence of systematic forecast error. The reduction of systematic errors has multiple potential benefits. These include (i) extended accurate lead times for numerical weather prediction forecasts, (ii) improved lateral boundary conditions for high-resolution domains nested within global models, and (iii) improved statistical properties of climate integrations. In this paper we demonstrate one route to systematic forecast error. We associate, in a case study, forecast error in the tropopause-level Rossby-wave development to the forecast error in a warm conveyor belt (WCB). The error in the WCB is shown to arise from errors in the modification of potential temperature (θ) and potential vorticity (PV) as air ascends along it.

Many previously published studies describe research aimed at understanding and reducing systematic error. Jung (2005) documents the systematic errors in European Centre for Medium-Range Weather Forecasts (ECMWF) model simulations of the winter atmospheric circulation, describing how they evolve out to a month time range and how they have changed with model developments over the previous two decades. In the extratropics one focus of attention has been on the representation of the stratosphere (see review article by Tripathi et al., 2015). In the tropics attention has focused on aspects including the representation of the Madden-Julian Oscillation (e.g, Holloway et al., 2013; Kim et al., 2009; Jung et al., 2005) and the sources of climatic sea surface temperature biases (Vannière et al., 2014) and their impact on El Nino Southern Oscillation predictability and variation (Manganello and Huang, 2009). Each of these efforts quantifies model error development at short timescales and determines how this error then limits prediction on monthly to seasonal timescales. These errors ultimately result in biases in climate models. Phillips et al. (2004) describe a parameterisation testbed initiative designed to support the running of climate models in numerical weather prediction configuration to diagnose deficiencies in parameterisations. More recently, Ma et al. (2014) demonstrate that most systematic errors in precipitation, clouds, and radiation processes in long-term climate runs from the fifth phase of the Coupled Model Intercomparison Project (CMIP5) are present by day five in ensemble average hindcasts in all models, and that these errors typically saturate after few days of hindcasts with amplitudes comparable to the climate errors. Rodwell and Palmer (2007) propose using this approach to assess the likelihood of climate predictions derived from perturbed model experiments as a computationally cheaper alternative to assessing the skill of each experiment in reproducing the present-day climate. The use of numerical weather prediction to diagnose sources of errors in climate models is enabled by the capability to run climate models in numerical weather prediction configuration. This capability is implicit in models that use a seamless approach to modelling the weather and climate timescales, and consequently has been used to analyse systematic forecast errors in the Met Office Unified Model (Martin et al., 2010).

In general, forecast error increases with forecast lead time although forecasts can degrade substantially at critical points in their evolutions (Langland et al., 2002). There are occasions when the forecast is particularly poor, known as forecast busts. In many cases, forecast busts are not simply 'bad luck', but associated with particular flow regimes that are both highly sensitive to small perturbations and associated with conditions that are poorly represented by the forecast model. As an example, a common type of European forecast bust has been linked to a particular large-scale flow regime occurring six days prior to the forecast bust episode by Rodwell et al. (2013). This precursor large-scale flow regime is characterised by a trough over the Rocky Mountains together with poor representation of strong organised thunderstorm activity over the US Great Plains. Although a clear precursor for European forecast busts is described by Rodwell et al. (2013), a complete account of the dynamical link between the precursor and forecast bust is not given. By mechanistically linking such forecast busts to the precursor flow regime, it should be possible to identify specific aspects of the model where improvement would yield maximum benefit.

Sources of systematic forecast error include inconsistencies in the dynamical core, insufficient resolution and errors due to the parameterisations of diabatic processes. As an example of dynamical core error, Whitehead et al. (2014) demonstrate inconsistency between model dynamical cores and tracer advection schemes in idealised breaking baroclinic wave simulations. Tracer advection schemes are required to passively advect the many tracers required in models, particularly chemical tracers in climate models, and hence such inconsistencies can lead to systematic errors. As an example of resolution errors, Holloway et al. (2013) considered the representation of the Madden-Julian oscillation in simulations with grid spacings of 40, 12 and 4 km. The best representation of the Madden-Julian oscillation was found with the highest resolution simulation, but this was found to depend critically on the representation of the convection and the boundary layer in the model. In contrast, considering coarser resolutions (of about 210 km and 40 km), Berner et al. (2012) found increased horizontal resolution had little impact in the tropics whereas improving the deterministic parameterisations and introducing stochastic parameterisations reduced several systematic errors; all three model refinements led to a decrease in the systematic bias of the Northern Hemispheric circulation.

The focus of this paper is error arising from the misrepresentation of diabatic processes due to the necessity to parameterise them. In a study of the Superstorm '93 (12–14 March 1993) using the NCEP operational Medium Range Forecast Model Dickinson et al. (1997) were able to show that the failure to reproduce the initial stages of the storm could be linked to the model's inability to produce sufficient deep convection due to deficiencies in the model's convection parameterisation scheme. Using examples from the CMIP5 integrations, Stevens and Bony (2013) argue the case for focusing efforts to improve climate models on improving the numerical representations of how clouds, moist convection and heating couple to the general circulation. In the extratropics cloudy air in ascending WCBs in extratropical cyclones experiences strong diabatic heating. Diabatic cooling also occurs in the associated evaporating downdraughts and due to long-wave cooling at cloud top or at strong humidity gradients. Madonna et al.

(2015) introduced a three-component feature-based verification measure for WCBs and applied it to operational medium-range forecasts from the ECMWF. While recent model versions are shown not to systematically over- or under-represent WCB amplitude (as measured by the amount of ascending air) or tropopause-level negative PV anomalies, individual forecasts are found to be associated with large errors particularly in the amplitude of WCBs. Notice that this lack of systematic error does not imply that diabatic processes are perfectly represented in current versions of the ECMWF model. WCBs are the result of processes over periods of at least two days. During these periods, complex interactions between diabatic processes and dynamics take place. These complex interactions are not yet fully understood (e.g. the relationship between cyclone strength and WCB strength). Under these circumstances, the misrepresentation of either component (diabatic processes or dynamics) in a model could in one case lead to an overestimation and in another case to an underestimation of the WCB amplitude. Chagnon et al. (2013) and Chagnon and Gray (2015) partition the modification of PV in WCBs into different parameterised model processes and demonstrate how these processes can enhance the PV gradient across the tropopause. This enhanced PV gradient modifies the phase speed of planetary Rossby waves propagating on the tropopause; the westwards propagation of the Rossby waves counter to the eastwards jet is enhanced, but the local jet speed is also enhanced such that the overall effect on the phase speed cannot easily be inferred. Errors in Rossby-wave propagation and structure propagate downstream leading to errors in the genesis and development of cyclones and anticyclones (Davis, 1992). Gray et al. (2014) diagnosed systematic errors in the Rossby waves in operational medium-range winter-season Northern-hemisphere forecasts. Tropopause sharpness adjacent to ridges and ridge amplitude systematically decreased out to about five-day lead time. A strong effect of model resolution was diagnosed but the authors concluded that more work is necessary to isolate the contribution of diabatic processes. Davies and Didone (2013) used a PV perspective to identify five mechanisms that can cause forecast errors by generating or enhancing Rossby waves. Two of these mechanisms, lower-stratospheric PV anomalies and anomalously-low values of tropospheric PV beneath the tropopause due to deep convection, directly relate to diabatic processes.

The aim of this study is to produce a detailed anatomy of one type of forecast error, that associated with WCBs. This is achieved through analysis of a specific case study that demonstrated large forecast error in the development of a tropopause-level ridge at lead times exceeding three days. The consistency between the error characteristics in this case and those found to occur systematically in medium-range operational forecasts by Gray et al. (2014) suggest that this route to forecast error could also be systematic.

The structure of the remainder of this paper is as follows. The models and diagnostics used are described in Section 2. The diagnostics used in this paper are the feature-based verification measure for WCBs introduced by Madonna et al. (2015) (which includes trajectory analysis) to characterise the forecast error in the WCB, and on-line θ and PV tracers to assess the diabatic modification of these variables by the WCB. The case study is introduced in Section 3. The link between tropopause-level ridge error and error in the WCB that outflows into the ridge is made

using output from operational forecasts and the error is shown to be reproducible in different forecast systems. The role of diabatic processes in the error development is investigated in Section 4. A summary and conclusions are presented in Section 5.

2 Data, tools and methodology

In this Section, we describe the models and methods used in this work. The output from the operational ECMWF high-resolution model (analyses and forecast) is used to compute WCB trajectories. The feature-based method of Madonna et al. (2015) is applied to this dataset, as described in Section 2.1. The operational ensemble forecast data from the ECMWF and the Met Office are also examined to investigate the representation of the case study in these ensemble prediction systems, as described in Section 2.2. Finally the Met Office Unified Model (MetUM) is used to produce hindcasts of the case study, including diabatic tracers, as described in Section 2.3.

2.1 Trajectory analysis and PAL

WCB trajectories are calculated from a 6-hourly output from the ECMWF operational analyses and the ECMWF high-resolution forecast initialised at 1200 UTC 19 January 2011. First the data is interpolated in the horizontal from the original grid (T1279) to a regular grid with 1° grid spacing, where the vertical levels (L91) remain unchanged. Using the Lagrangian Analysis Tool LAGRANTO (Wernli and Davies, 1997) trajectories that experience a rapid ascent (more than 600 hPa in 48 hours) are selected. To be identified as WCBs, the trajectories have to ascend in the vicinity of an extratropical cyclone. Moreover, a filter which removes double counts of the same air mass is applied. More details on the calculation of WCBs can be found in Madonna et al. (2015).

To verify WCB outflows in forecasts against those in analyses, the feature-based method PAL (PV-anomaly–Amplitude–Location) (Madonna et al., 2015) is used. In this method, three different aspects of the forecast of WCBs are considered, as described in detail in Madonna et al. (2015). The first component, P, measures the error of the PV anomaly in the WCB outflow region. PV is strongly modified within WCBs due to diabatic processes. However, air masses in the WCB start and end their ascent to the upper troposphere with low PV values ($\sim$ 0.5 PVU) (Methven, 2015). As a result, WCBs reach the upper troposphere with low PV values, generating a negative PV anomaly when compared to the climatological PV values. The negative PV anomaly can be important for the subsequent upper-level flow evolution. The P component is not normalised and can assume positive and negative values. A negative P value means that the PV anomaly is too weak in the forecast. The second component, A, measures the amplitude error of the WCB, since the amount of rising air is crucial for the modification of the upper-level flow. The A term is defined as the difference in the number of WCB trajectories normalized by the average number of WCB trajectories in the analysis and forecast. The A component can assume values between -2 and $+2$ with positive values

indicating an overestimation of the number of WCB trajectories in the forecast. Finally, the L component is a measure of the horizontal location error of the WCB outflow. This component consists of the error of the distance between the centre of mass of all WCBs and the error of the averaged distance between each WCB and the overall centre of mass (for more details see Wernli et al., 2008; Madonna et al., 2015). It can only assume positive values between 0 and 2. A value close to 2 indicates that the WCB outflows are located at the opposite edges of the chosen domain (here the North Atlantic 15°N–90°N, 100°W–10°E).

2.2 TIGGE

THORPEX (The Observing System Research and Predictability Experiment) Interactive Grand Global Ensemble (TIGGE) (Park et al., 2008) is an archive of ensemble forecast data from ten operational ensemble prediction systems (EPSs), starting from October 2006. In this work, we examine two of these EPSs, namely the 15-day forecasts from the Met Office Global and Regional Ensemble Prediction System (MOGREPS-15) (Bowler et al., 2008, 2009; Tennant et al., 2011) and the ECMWF EPS (Molteni et al., 1996; Buizza et al., 1999). Relevant features of these two EPSs are summarised in Table 1. These aspects will be discussed in Section 3.3, once the context of the study has been further developed.

2.3 MetUM

The Met Office Unified Model (MetUM) is an operational finite-difference numerical weather prediction model that solves the non-hydrostatic deep-atmosphere dynamical equations with a semi-implicit, semi-Lagrangian integration scheme (Davies et al., 2005). It uses Arakawa C staggering in the horizontal (Arakawa and Lamb, 1977) and is terrain-following with a hybrid-height Charney–Phillips (Charney and Phillips, 1953) vertical coordinate. Its package of physical process parameterisations includes longwave and shortwave radiation (Edwards and Slingo, 1996), boundary layer mixing (Lock et al., 2000), cloud microphysics and large-scale precipitation (Wilson and Ballard, 1999), and convection (Gregory and Rowntree, 1990).

In this study, the MetUM version 7.3 has been used in its global configuration to produce hindcasts of the case study. The model has been complemented by the inclusion of θ and PV tracers, as detailed in Section 2.3.1. The model has been initialised using its native Met Office operational global analyses to minimize the effects of 'initial shocks' (Klocke and Rodwell, 2013). The initialisation dates are detailed in Section 2.3.2. The simulations were performed with the MetUM version 7.3 at a horizontal resolution of N512 ($\sim$ 25 km grid spacing) and 70 vertical levels with the top of the model around 80 km. The model was run with its standard physical parameterisation package and climatological SST.

2.3.1 Diabatic tracers

Diabatic tracers are tracers of changes in θ and PV due to diabatic processes. Potential temperature tracers enable identification of the processes that bring air parcels to their current

isentropic level through cross-isentropic motion. PV tracers enable identification of modifications to the circulation and stability and the diabatic processes responsible for such modifications. The θ tracers used have been described in Martínez-Alvarado and Plant (2014) and Martínez-Alvarado et al. (2014), while the PV tracers have been described in Stoelinga (1996), Gray (2006), Chagnon and Gray (2009) and Chagnon et al. (2013).

The idea behind both kinds of tracers consists of the separation of the variable of interest φ (representing either θ or PV) into the sum of a materially conserved component, φ_0, a diabatically-generated component, φ_d, and a residual r_φ, i.e.

$$\varphi(\boldsymbol{x}, t) = \varphi_0(\boldsymbol{x}, t) + \varphi_\mathrm{d}(\boldsymbol{x}, t) + r_\varphi(\boldsymbol{x}, t), \tag{1}$$

where $\boldsymbol{x}$ represents spatial location and t is time. Each part is governed by the following equations

$$\frac{D\varphi_0}{Dt} = 0 \tag{2}$$

$$\frac{D\varphi_\mathrm{d}}{Dt} = S_\varphi, \tag{3}$$

where S_φ represents a source due to diabatic processes. In other words, φ_0 is unaffected by diabatic processes, and therefore conserved following an air parcel, while φ_d is affected by diabatic processes. The sum of (2) and (3) yields the full evolution equation for φ. At initialisation time, t_0, $\varphi_0(\boldsymbol{x}, t_0) = \varphi(\boldsymbol{x}, t_0)$ and $\varphi_\mathrm{d}(\boldsymbol{x}, t_0) = 0$.

Equations (2) and (3) are solved using the same numerical methods implemented in the MetUM to solve the evolution equations of the model's prognostic variables. However, there are details within such numerical methods which lie far beyond the scope of this article, such as the treatment of vertical interpolation in the advection of potential temperature, that are only strictly valid for the prognostic variables themselves. Therefore, there is an unavoidable mismatch between the computation of the prognostic variables and the computation of diabatic tracers. This mismatch is specially important in the computation of advection. Thus, the residual term in (1) arises from the mismatch between the advection of φ and that of each individual tracer (see e.g. Whitehead et al., 2014). The residual grows over time in both cases, but at a slower rate for the θ-tracers than for PV-tracers. However, restricting the length of the simulations to a maximum of two days also restricts the residual growth to a maximum of ± 2 K for θ and ± 0.5 PVU for PV within the region of interest. Notice that, apart from the just described issues arising from their numerical implementation, the diabatic tracers are fully consistent with the parent model (i.e. the MetUM) in terms of the grid-scale effects of each parameterisation scheme (including convection) and in terms of the advection of these effects due to the grid-scale wind components. A similar discussion can be found in Martínez-Alvarado et al. (2014) regarding the related method of trajectory analysis.

2.3.2 Sequence of hindcast simulations

Figure 1 is a graphical representation of the method used to generate a forecast proxy to the evolution of the realised flow constrained by analyses. The analysis of forecast error requires knowledge of the realised flow evolution and the availability of a contemporaneous forecast (e.g. Davies and Didone, 2013). Here, the contemporaneous forecast simply consists of a five-day hindcast starting from the analysis on 1200 UTC 19 January 2011 (segment A-F' in Fig 1). However, our knowledge of the realised flow evolution is restricted to the analyses at given times (points A,B,...,F in Fig. 1). Thus, we have an approximation to the actual state of the atmosphere at given times, but not the flow evolution joining atmospheric states from one time to the next. To circumvent this problem, the realised flow evolution is proxied by a sequence of short-range (up to two-day) hindcasts. A necessary assumption underlying the use of a short-range hindcasts as a proxy to the realised flow is that such hindcasts can be trusted as a reliable representation of the physical processes governing the atmospheric evolution. No model is perfect and no initial conditions are free of error. As a result, forecast error starts growing as soon as a simulation starts. For example, Zhang et al. (2003) (see also Zhang et al., 2007) have demonstrated how forecast errors in idealised simulations of moist baroclinic waves grow from small convective scales up to the large-scale flow in timescales of the order of 12 hours. However, the assumption of model reliability as a representation of the atmospheric evolution is justified by noticing that forecast error is relatively small in two-day forecasts. For example, Frame et al. (2011) showed that the forecast skill of transitions of the North Atlantic eddy-driven jet of both the ECMWF EPS and MOGREPS-15 was very good for the first three forecast days, but rapidly deteriorated beyond this time.

The duration of the short-range forecasts is restricted to a maximum of 48 hours so that they are sufficiently short to remain close to the phase space orbit defined by the sequence of analyses, but sufficiently long so that the diabatic tracers can accumulate a useful amount of information about changes due to diabatic processes. As will be explained in Section 4, the error development will be subdivided in three stages, each one lasting 24 hours. The first stage, corresponding to segment C–D in the realised flow (refer to Fig. 1 for all segment definitions), will be studied using segment C–d as proxy to the realised flow; the second stage, corresponding to segment D–E, will be studied using segment D–e, and the third stage, corresponding to segment E–F, will be studied using segment D–f. A common initialisation time for the study of the second and third stages enables the tracking of the same air masses labelled by the tracer θ_0 throughout these two stages. The tracer $\theta_0 = \theta_0(\mathbf{x}, t)$ represents the potential temperature that air at position $\mathbf{x}$ and time t had at the time of reinitialisation. Because the proxy to the realised flow is a sequence of short-range hindcasts, the tracers in both the proxy to the realised flow and the contemporaneous forecast have been re-initialised every day at 1200 UTC from 19 January to 22 January, inclusive.

All the hindcast simulations constituting the proxy to the realised flow and the contemporaneous forecast were performed using the MetUM with the configuration described in Section 2.3. This method enables process evaluation between states (up to the model's temporal and spatial

resolution) using θ and PV tracers for both the proxy to the realised flow and the contemporaneous forecast. Therefore, this method enables a comparison of the modification of θ and PV between these two hindcast sequences. However, the method does not allow us to definitely associate forecast error growth with individual parameterised processes.

3 Case study

3.1 Overview of the operational forecast error

The forecast error development case presented here consists of the formation of a pronounced ridge over the central North Atlantic which was noticeably underestimated at a lead time beyond three days in several operational forecast systems, namely the ECMWF high resolution forecast, the ECMWF EPS and MOGREPS-15. The case is characterised by the absence of large errors in the upper-level flow in the North Atlantic prior to three days lead time and by the absence of error propagation from upstream into the North Atlantic region.

In this section, we describe the development of forecast error in PV on the 320-K isentropic level as a comparison between the ECMWF operational analysis and the ECMWF high-resolution forecast. In subsequent sections, we shall describe the relationship of this error to other important atmospheric features and fields: section 3.2 shows the relationship between the forecast error on the 320-K PV field and forecast errors in the structures of collocated WCBs; Section 3.4 shows the relationship between the errors in the 320-K PV and WCBs and errors in mean sea level pressure.

Figure 2 illustrates the 320-K PV forecast error evolution between the forecast starting at 1200 UTC 19 January 2011 and the analysis for the four-day period ending on 1200 UTC 24 January 2011. The intersection of the tropopause and the isentropic surface in forecast and analysis is also indicated in Fig. 2 by the corresponding 2-PVU contours (black for the analysis, green for the forecast). Figure 2a shows PV error one day after the initialisation of the forecast (D+1) at 1200 UTC 20 January 2011. The tropopause structure is similar between analysis and forecast over the whole Northern Hemisphere and the PV errors are small. Two days later (D+3, Fig. 2b) the location of the tropopause in analysis and forecast are still similar and the PV errors are small (also valid for D+4, not shown). At 1200 UTC 24 January 2011 (D+5, Fig. 2c), we observe two main regions with strong positive PV errors (> 5 PVU), one over Alaska and one in the North Atlantic between 70°W and 20°W. Moreover, strong negative PV errors are also found in the North Atlantic. In this region (indicated by the grey box) the tropopause is located further south and west in the forecast than in the analysis, meaning that the ridge building is underestimated in the forecast. The error in the PV signal in the North Atlantic is thus related to the wrong representation of the ridge structure in the forecast. Over the Pacific, the tropopause structures are much more similar in the forecast and analysis and the PV errors are much weaker than over the North Atlantic.

The underestimation of the ridge area in the forecast after five days agrees well with the results presented in Gray et al. (2014). They analysed seven winters and showed that the am-

plitude of Rossby waves is reduced with increasing lead times up to five days. After that, a saturation point is reached. Gray et al. (2014) notice that the reduced amplitude is consistent with the under-representation of diabatic processes and transport of tropospheric air with reduced PV into the upper troposphere. As shown by Joos and Wernli (2012) and Chagnon et al. (2013) one possible mechanism for the modification of upper-level ridge structure is the strongly diabatic cross-isentropic flow of an ascending WCB. On the other hand, all the numerical models considered here are highly nonlinear systems. Therefore, in addition to the potential misrepresentation of diabatic processes and their interactions with the model's dynamical core, the models exhibit high sensitivity to initial conditions which would be present in the development of forecast error even if the forecasts were produced by a perfect model.

3.2 Relationship between the 320-K PV operational forecast error and WCB

The formation of the upper-level PV anomaly described in Section 3.1 is associated with WCBs, whose outflows reach into the upper-level ridges, both in the analysis and in forecasts. Even though PV is strongly modified during the ascent due to diabatic processes, air masses in the WCB start and end their ascent to the upper troposphere with low PV values (~ 0.5 PVU) (e.g. Wernli and Davies, 1997; Grams et al., 2011; Methven, 2015). Therefore, the WCBs contribute to the building of the large upper-level ridges with low PV values which start to form around 1200 UTC 22 January 2011.

The right panels of Fig. 2 show an enlargement of the domain between 90°W–0°W and 30°N–80°N (grey boxes on left panels). In addition to the 320-K 2-PVU contour and the PV forecast errors, WCB intersection points from the trajectories satisfying the ascent criteria are displayed. The intersection points include all WCBs in the analysis and forecast that are located at 320 K at the given time, independent of their starting time.

At 1200 UTC 20 January (D+1) no WCB trajectories are located at 320 K (Fig 2a), the tropopause structure is similar in the forecast and analysis and the PV errors are very small. Two days later (D+3, Fig. 2b, right panel), the first WCB trajectories cross the 320-K isentropic surface and at similar locations in the analysis (black crosses) and forecast (green circles) and the tropopause structure is still quite similar. At D+5 (Fig. 2c, right panel) more WCB intersection points are found. They are located in the pronounced ridge, which extends from 40°N-70°N and 20°W-70°W, thus covering a large part of the central North Atlantic (Fig. 2c, right panel). The WCB trajectories in the analysis are located further north and are co-located with the region of the large positive PV error, whereas the WCBs in the forecast are located further south and east. Due to their importance for the formation or amplification of upper-level ridges (Wernli and Davies, 1997; Grams et al., 2011; Methven, 2015), an error in the WCB outflows can result in an error in the tropopause structure and the related PV field, as shown in Fig. 2c (right panel).

For each analysis time , the PV anomaly, amplitude and location errors of the WCB outflows in the forecast are determined for the domain over the North Atlantic (15°N–90°N, 100°W–

10°E). The time series of the errors are shown in Fig. 3. To calculate the components P, A and L we always use the WCB outflows, i.e. we use the position and PV values 48 hours after the start of the ascent. Since the forecast is initialised at 1200 UTC 19 January 2011, the first PAL values are computed 48 hours later, at 1200 UTC 21 January 2011. At this time, the P value is positive, meaning that the PV anomaly in the forecast is too strong, but the amplitude is well forecast (A close to zero), as is the location. However, at this stage, only a few WCB trajectories (less than 100) are detected and the effect of the WCB on the upper-level PV pattern is small.

The P, A and L values stay between -0.5 and 0.5 until 0000 UTC 24 January 2011, meaning that the WCB outflows are quite well represented in the forecast. This is consistent with the small PV error shown in Fig. 2b. However, starting from 0600 UTC 24 January 2011 a drastic drop in the P term to negative values of about -1.5 PVU is observed, in conjunction with an increase of the A value, whereas the L term remains small. This situation persists for several analysis times until 0000 UTC 25 January 2011, with the maximal negative P value observed at 1200 UTC 24 January 2011. The negative P value indicates that the PV anomaly in the forecast is too weak. This is related to the meridional and vertical location of the WCB outflow. In the analysis, the WCB outflow is further north and at similar location reaches a higher isentropic level compared to the forecast (not shown). Further to the north as well as higher up the climatological PV values are higher and therefore the resulting PV anomaly is stronger in the analysis than in the forecast. At the same time, a positive A term implies an overestimation of the number of WCB trajectories in the forecast. The A term is normalised and the value of 0.712, found at 1200 UTC 24 January 2011, indicates that the forecast has more than twice as many trajectories as the analysis. The L term is very small because it is scaled with the domain size and so the small value implies that we are comparing the same WCB systems. However, the total horizontal displacement between the centre of mass of the WCB in the analysis and that in the forecast is still about 1200 km. A similar evolution of the PAL values is also observed for forecast initialisation times of 1200 UTC 20 January and 21 January 2011 (not shown), and the development of the wrong ridge structure is consistent over the different lead times (see also Fig. 4a).

Summarizing, starting at 0600 UTC 24 January 2011, there is an error in the strength of the PV anomaly produced by the WCB outflows. The WCB outflow in the forecast is too far south and east and its amplitude is overestimated. The poorly represented WCB outflow, which persists over several consecutive analysis times, results in an erroneous ridge structure over the North Atlantic and large PV errors, which can propagate and have a pronounced downstream impact.

3.3   Reproducibility of operational forecast error development

Figure 4a shows 320-K 2-PVU contours in the analysis (black line) and in the ECMWF EPS control members at different lead times (coloured lines), all valid at 1200 UTC 24 January 2011. It shows general agreement between the different 2-PVU contours except in two areas: North

America (Alaska and the north of Canada) and the North Atlantic. The horizontal extent of the large ridge over the North Atlantic tends to decrease as forecast lead time increases. This tendency leads to a configuration of the five-day forecast, that is very similar to that in the high-resolution forecast for the same lead time (cf. Fig. 2c). Thus, this figure illustrates that the development of the 320-K PV forecast error in the operational ECMWF high-resolution forecast described in Section 3 occurs in a remarkably similar manner in other operational forecast systems. In addition to the ECMWF high-resolution forecast and the control member of the ECMWF EPS, the 320-K PV forecast error also develops in a similar manner in the control member of MOGREPS-15. The configuration characteristics of these forecast systems are shown in Table 1. Each forecast system is characterised by a given discretization method (spectral in the ECMWF models versus finite-difference in the MetUM), horizontal grid spacing (ranging from ~ 20 km in the ECMWF high resolution forecast to ~ 60 km in MOGREPS-15), number of vertical levels, and location of upper-most level. There are also differences in the physical parameterisation packages used by each model and the data used to provide surface boundary conditions, such as sea surface temperature (SST). The reproducibility in the development of the 320-K PV forecast error across different forecast systems indicates common systematic errors across these systems.

The reproducibility of the forecast error development is not restricted to the control members of the ECMWF-EPS and MOGREPS-15. Figure 4b shows 320-K 2-PVU contours in the analysis (black line), the five-day forecast control member (red line) and the fifty ECMWF-EPS members (grey lines) for the same forecast, all valid at 1200 UTC 24 January 2011. Even though there is a range in Rossby-wave amplitudes, the ensemble members cluster around the control member, with no member resembling the structure of the Rossby-wave in the analysis. An equivalent figure constructed using the corresponding MOGREPS-15 output (not shown) exhibits the same features just described with the ensemble clustering around the control member and no member depicting the Rossby-wave structure in the analysis.

The ECMWF-EPS and MOGREPS-15 have been developed independently. As a result, there are important differences between the formulations of these two ensemble prediction systems. The ECMWF-EPS uses singular vectors for the generation of initial condition perturbations (Molteni et al., 1996) whereas in MOGREPS-15 these perturbations are generated using an ensemble transform Kalman filter (Bishop et al., 2001; Bowler et al., 2008). Moreover, both systems implement methods to represent variability due to model uncertainty, but these methods are again different: while the ECMWF-EPS uses stochastic perturbation of physics tendencies (Buizza et al., 1999), MOGREPS-15 uses random parameters and stochastic kinetic energy backscatter to perturb the model physics (Bowler et al., 2008, 2009; Tennant et al., 2011). Taking these differences into consideration, and assuming the absence of large errors in the initial (beyond the reach of the initial condition perturbations) and boundary conditions, the results just presented lead to two conclusions. First, in this case the forecast error develops similar characteristics independently of the initial conditions or the strategy taken to generate the initial condition perturbations. Therefore, even though the case exhibits sensitivity to ini-

tial conditions leading to the spread around the control member in the five-day forecast, this sensitivity does not explain the forecasts error. Thus, we can conclude that, in addition to the inherent nonlinearity of the system, the development of forecast error in this case is potentially heavily influenced by model error. Second, in spite of the use of stochastic perturbations, the models show a strong tendency towards their own preferred state. This preferred state appears to be similar for the ECMWF's and Met Office's models despite substantial differences in model formulation and implementation, as described above.

As a final remark to this section, we emphasize that the forecast error development described here appears to be a recurrent feature of the models, with other cases exhibiting the same qualitative characteristics. For example, Martínez-Alvarado (2014) showed a very similar error development in the forecast of a large-amplitude Rossby-wave ridge on 25 November 2009. Martínez-Alvarado et al. (2014) showed that, like in this case, the Rossby-wave ridge was also associated with a strong WCB.

3.4 Forecast error in MetUM hindcasts

We have shown that the development of the 320-K PV forecast error is highly reproducible. We take advantage of this quality to carry out a hindcast simulation in which a realisation of this same forecast error occurs. As described in Section 2.3, the simulations were performed with the MetUM version 7.3 at a horizontal resolution of N512 ($\sim$ 25 km grid spacing) and 70 vertical levels with the top of the model around 80 km. The model was run with its standard physical parameterisation package and climatological SST (cf. Table 1). In support of the error reproducibility argument, despite the differences in the set up of this model compared to those already considered, the 320-K PV forecast developed an error with similar features to those described in Sections 3.1 and 3.3. In this section we give an account of forecast error development in terms of mean sea level pressure and relate this to forecast error in 320-K PV (Section 3.1) and associated WCBs (Section 3.2). The following description has a very similar counterpart in the three operational forecast systems discussed in Section 3.3.

Figure 5 shows three snapshots of the evolution of mean sea level pressure and 320-K PV in the proxy to the realised flow (left column) and the contemporaneous forecast (right column). The times shown are 1200 UTC 22 January 2011 (upper row, cf. Fig. 2(a,b)), 1200 UTC 23 January 2011 (middle row) and 1200 UTC 24 January 2011 (lower row, cf. Fig. 2(e,f)). On 22 January, the error in the 320-K PV is small relative to the error on 24 January. A comparison of mean sea level pressure also shows small differences on 22 January. However, some of these differences in mean sea level pressure are very important for the subsequent development of the flow. In the proxy to the realised flow, there is a low-pressure system with a central pressure around 1004 hPa, located off the east coast of Florida, United States (Fig. 5a). In contrast, the equivalent system in the contemporaneous forecast appears as an elongated low-pressure region, with lowest pressure around 1006 hPa, extending from Yucatán, Mexico to Florida (Fig. 5b). For this low-pressure system, mean sea level pressure closed contours first appear in both simulations at 0000 UTC 22 January 2011, i.e. 12 hours before the first time shown

12

in Fig. 5. At that time, the mean sea level pressure minima were located over Florida and Yucatán in the proxy to the realised flow and the contemporaneous forecast, respectively. This location mismatch represents an error of 1100 km. In the proxy to the realised flow, the low-pressure system forms as a secondary cyclone over the trailing cold front attached to the deep low-pressure system (central pressure around 974 hPa) over Nova Scotia, Canada (Fig. 5a). This low-pressure system over Nova Scotia is also present in the contemporaneous forecast with the same intensity (Fig. 5b). However, the corresponding trailing cold front has a weaker temperature gradient in this case (not shown), especially at its southern end, where the new low-pressure system forms.

Figures 5(c,d) and 5(e,f) show the effects of the small cyclogenesis differences for the subsequent flow after 24 and 48 hours, respectively. The low-pressure system deepened and travelled to the north in both simulations. However, it reached 986 hPa on 23 January (Fig. 5c) and then 966 hPa on 24 January (Fig. 5e) in the proxy to the realised flow. In contrast, it only reached 994 hPa on 23 January (Fig. 5d) and then 978 hPa on 24 January (Fig. 5f) in the contemporaneous forecast. These differences in intensification rates are accompanied by differences in tropopause structure. A comparison of the 2-PVU PV contour on 23 January shows that even though the amplitudes of the upper-level waves were similar in both simulations, their longitudinal widths were different, ranging between 25°W and 70°W in the proxy to the realised flow and between 25°W and 75°W in the contemporaneous forecast.

In terms of location, the low-pressure centre travelled faster in the contemporaneous forecast to end up located only 500 km southeast relative to its location in the proxy to the realised flow on 24 January. The development of the upper-level ridge followed that described in Sections 3.1 and 3.3. The ridge is more strongly influenced by cyclonic circulation induced by the low-pressure system in the proxy to the realised flow than in the contemporaneous forecast, so that its western flank wraps around the parent cyclone.

An additional difference between the proxy to the realised flow and the contemporaneous forecast lies in the development of the preceding low-pressure system located to the north of Newfoundland. Its location is very similar on 24 January in both simulations. However, this system decays more rapidly in the proxy to the realised flow than in the contemporaneous forecast: the increase in central pressure is 24 hPa in the former while it is only 8 hPa in the latter between 22 January and 24 January.

4   Forecast error and diabatic processes

In this section we investigate the diabatic processes that are associated with the forecast error development by performing a comparison of the modification of θ and PV between the proxy to the realised flow and the contemporaneous forecast. The error development is subdivided in three stages. The first stage corresponds to the early development of the low-pressure system off Florida (Fig. 5a) between 1200 UTC 21 January and 1200 UTC 22 January (segment C–D, refer to Fig. 1 for all segment definitions). The second stage corresponds to the intensification

of the low-pressure system (Fig. 5c) between 1200 UTC 22 January and 1200 UTC 23 January (segment D–E). Finally, the third stage corresponds to the development of the upper-level ridge (Fig. 5d) between 1200 UTC 22 January and 1200 UTC 24 January (segment E–F). The analysis of these three stages is performed through diagnosis of modifications to θ and PV. As explained in Section 2.3.2, the first stage is studied using the proxy to the realised flow between 1200 UTC 21 January and 1200 UTC 22 January (segment C–d). In order to track the same air masses labelled by the tracer θ_0 throughout the second and third stages, a common initialisation time is used for these two stages. Thus, the second and third stages are studied using the proxy to the realised flow initialised at 1200 UTC 22 January. The second stage is studied using segment D–e while the third stage is studied using segment D–f.

4.1   First stage: Early cyclone development

Figure 6 shows θ_0 on the 300-K isentropic surface at the end of the 24-hour period starting at 1200 UTC 21 January 2011 in the proxy to the realised flow (Fig. 6a) and the contemporaneous forecast (Fig. 6b), i.e. segments C–d and C'–D' in Fig. 1, respectively. The gray shading in Fig. 6 corresponds to θ values around the nominal value of the isentropic surface. Colours above the gray shading on the colour scale represent cross-isentropic descent while colours below the gray shading on the colour scale represent cross-isentropic ascent.

The regions of diabatic activity are similar in the two simulations. There is cross-isentropic ascent of up to 20 K close to the preceding northern low-pressure system (denoted C1 in Fig. 6) and along its trailing cold front. There is also cross-isentropic descent on both sides of the frontal ascent. The main differences between the simulations occur on the southern end of the trailing cold front, around the location of the developing cyclone off the coast of Florida (denoted C2 in Fig. 6a). In the proxy to the realised flow (Fig. 6a), the higher frontal intensity has induced stronger cross-isentropic ascent than in the contemporaneous forecast (Fig. 6b).

Figure 7a shows total modification of θ in the proxy to the realised flow on a vertical section through the low-pressure centre (segment A–B in Fig. 6), located at 77°W. There are two locations with strong heating, around 80°W and 70°W. The shallow column of heating around 80°W, extending from the surface to 700 hPa, forms the leading edge of the cross-isentropic ascent region wrapping cyclonically around the low-pressure centre (Fig. 6a). The boundary layer parameterisation is the main contributor to this column of heating (Fig. 7b) but the convection parameterisation also contributes, especially in the layer between 290 K and 300 K (Fig. 7c); the other contributors to the heating are small by comparison. These contributions are consistent with the rain rate partition between convective and large-scale rain, which shows that the convective rain was more intense than large-scale rain in the region of interest (not shown). The second location of strong heating, located around 70°W, is formed by a column of deep convection ascending over a katafront (see contours of equivalent potential temperature in Fig. 7c). This column is also accompanied by clouds and so resembles the cloud column along the WCB in extratropical cyclones (see relative humidity contours in Fig. 7a). The convection parameterisation is the main contributor to this column of heating (Fig. 7c). Figure 7a also

14

shows regions of cooling. This cooling is located at cloud top level and is due to the radiation parameterisation scheme (not shown).

Diabatic activity, defined by changes in θ, is distributed similarly in the contemporaneous forecast. However, there are important differences in its morphology and intensity between simulations. Figure 7d shows total modification of θ in the contemporaneous forecast on the same section as that in Fig. 7a. As in the proxy to the realised flow, there is a region of heating around 80°W in the contemporaneous forecast, and it is mainly caused by the boundary layer parameterisation. However, in this case, this region of heating is very much restricted to the surface, barely reaching 900 hPa. The parameterised diabatic processes responsible for the diabatic activity distribution are the same in both simulations, namely the parameterisations of boundary layer (Figs. 7(b,e)), convection (Figs. 7(c,f)) and radiation (not shown). There is also a region of strong heating around 70°W in the contemporaneous forecast. As in the proxy to the realised flow, this region is caused by convection. However, in this case, the region with convective influence is deeper and more upright than in the proxy to the realised flow (Fig. 7f), possibly as a response to differences in katafront structure (cf. equivalent potential temperature contours in Fig. 7c). The differences in the tropopause are due to error in the location of an intrusion of stratospheric air that in the contemporaneous forecast is located more to the west than in the proxy to the realised flow. A comparison between Figs. 7(a,d) also shows stronger diabatic activity in the realised flow than in the contemporaneous forecast.

The changes in θ are accompanied by changes to the circulation, which can be described in terms of PV modification. In the proxy to the realised flow, positive PV modification contributing to the initial cyclone development is found along the trailing cold front peaking around the 295-K isentropic surface in the proxy to the realised flow (not shown). In this case, the process that contributed the most to PV production in this region was the parameterised convection, through the generation of a heating maximum above the 295-K isentropic surface. In contrast, PV production in the contemporaneous forecast is much weaker and much less organised than in the proxy to the realised flow.

4.2   Second stage: Cyclone intensification

Figure 8 shows θ_0 on the 320-K isentropic surface at the end of the 24-hour period starting at 1200 UTC 22 January 2011 in the proxy to the realised flow (Fig. 8a) and the contemporaneous forecast (Fig. 8b), i.e. segments D–e and D'–E' in Fig. 1, respectively. Diabatic activity is located in similar regions in the proxy to the realised flow and the contemporaneous forecast, with isentropically ascending air filling the western flank of the corresponding ridge. However, at this isentropic level, cross-isentropic ascent is more intense in the proxy to the realised flow, with a larger proportion of air originating between 290 K and 300 K than in the contemporaneous forecast. Furthermore, the region of cross-isentropic ascent is more zonally elongated in the contemporaneous forecast than in the proxy to the realised flow.

A more comprehensive comparison between the proxy to the realised flow and the contemporaneous forecast can be achieved by statistical analysis of air throughout the troposphere

15

that has undergone cross-isentropic ascent and is located under the tropopause dome. This air mass is formally defined as the air satisfying the following two conditions:

$$\theta_0 > 280 \text{ K} \quad \text{and} \quad 10 < \Delta\theta < 50 \text{ [K]}, \tag{4}$$

where $\Delta\theta = \theta - \theta_0$. The quality of the following results does not change if the actual threshold values are changed within reason. Conditions (4) are applied to a region enclosing the ridge of interest in both the proxy to the realised flow and the contemporaneous forecast. This definition includes all the blue-shaded regions within the ridge of interest in Fig. 8.

Figure 9 shows joint probability density functions (PDFs) characterising these air masses. Figures 9(a,b) show the joint PDF at 1200 UTC 22 January, represented by $(\theta_0, \mathrm{PV}_0)$ (black contours), and at 1200 UTC 23 January, represented by (θ, PV) (colour shading), in the proxy to the realised flow (Fig. 9a) and the contemporaneous forecast (Fig. 9b). These PDFs represent the air mass evolution during the past 24 hours.

In the proxy to the realised flow, a largely monomodal PDF with a shallow peak centred around (0.25 PVU, 300 K) evolves into another largely monomodal PDF with a shallow peak centred around (0.1 PV, 315 K) (see Fig. 9a). Secondary peaks also appear below (around (0.15 PVU, 307 K)) and above (around (0.05 PVU, 327 K)) the primary peak. In the contemporaneous forecast, the initial PDF is also largely monomodal but its peak is sharper than in the proxy to the realised flow. This PDF evolves into a fuzzy distribution with no clear maxima (Fig. 9b).

These evolution differences result in clear differences in the mean rate of change of θ and PV over 24 hours. Figure 9c shows the joint PDFs for $(\Delta\mathrm{PV}, \Delta\theta)/\Delta t$, where $\Delta t = 24$ h, in the proxy to the realised flow (colour shading) and the contemporaneous forecast (black contours). In general, the mean rate of change of θ is well represented by the contemporaneous forecast. However, the peak in the PDF of the proxy to the realised flow is much narrower (along the $\Delta\theta/\Delta t$ axis) than in the contemporaneous forecast indicating that the latter tends to have greater variability in heating rates around the most frequent values [1]. Larger differences are found in the mean rate of change of PV with the whole forecast distribution displaced to the right with respect to the proxy to the realised flow. This displacement indicates a tendency in the proxy to the realised flow to destroy PV that is not matched in the contemporaneous forecast during this period.

4.3 Third stage: upper-level ridge development

Figure 10 shows θ_0 on the 320-K isentropic surface at the end of the 48-hour period starting at 1200 UTC 22 January 2011 in the proxy to the realised flow (Fig. 10a) and the contemporaneous forecast (Fig. 10b), i.e. segments D–f and D'–F' in Fig. 1, respectively. Notice that this period includes the period discussed in Section 4.2. At this time, the upper-level ridge has fully

---

[1]N.B. Here, the term 'variability' refers to the variability within the population of grid points satisfying conditions (4). It does not refer to variability among ensemble members as we are not discussing ensembles in this Section

developed in both simulations. In the proxy to the realised flow, the ridge is constituted primarily by air that has undergone rapid cross-isentropic ascent (up to 40 K (48 h)$^{-1}$) filling most of the ridge area towards its centre and eastern flank. The western flank of the tropopause ridge, on the other hand, is formed by air that has undergone weak cross-isentropic subsidence ($|\Delta\theta| < 10$ K (48 h)$^{-1}$). This latter feature is not found in the contemporaneous forecast on this isentropic level inside the tropopause ridge. Instead, the tropopause ridge is composed almost entirely by cross-isentropic ascent. The parameterised process responsible for cross-isentropic descent along the western flank of the ridge in this case is long-wave radiation. However, cross-isentropic descent alone is not sufficient to explain the expansion of the ridge area without a mechanism to reduce the high values of PV characterising the cross-isentropically descending air. In this case, the PV reduction is induced by the upper part of the heating maximum due to the cloud microphysics scheme. A detailed analysis of these processes is beyond the scope of this article, whose focus are processes within the WCB. However, they will be further analysed in a forthcoming article.

The cross-isentropic ascent during this period is analysed through the joint PDFs shown in Fig. 11. Figure 11a shows the evolving joint PDFs from 1200 UTC 22 January $(\theta_0, \mathrm{PV}_0)$ to 1200 UTC 24 January (θ, PV) in the proxy to the realised flow. Notice that even though the $(\theta_0, \mathrm{PV}_0)$ PDFs in this and the previous stage correspond to the same time, the PDFs are not expected to be the same because the selection criteria (4) are applied at the end of each study period. The $(\theta_0, \mathrm{PV}_0)$ PDF is largely monomodal and it is centred around the same values (0.25 PVU, 300 K), as in the previous stage. Furthermore, this PDF evolves into another largely monomodal PDF. However, the new PDF is centred around (0.05 PVU, 325 K), i.e. at a higher isentropic level than in the previous stage, indicating the further cross-isentropic ascent, by further 10 K, of air masses reaching 315 K during the previous stage. This subsequent cross-isentropic ascent is accompanied by further PV reduction.

Figure 11b shows the evolving joint PDFs $(\theta_0, \mathrm{PV}_0) \rightarrow (\theta, \mathrm{PV})$ in the contemporaneous forecast. The evolution depicted by these PDFs is similar to that in the proxy to the realised flow. However, there are important differences. First, the distribution centre of the $(\theta_0, \mathrm{PV}_0)$ PDF in the contemporaneous forecast is located at a higher isentropic level (around 307 K) than in the proxy to the realised flow (around 300 K). Second, the maximum of the (θ, PV) PDF in the contemporaneous forecast is located at a slightly higher isentropic level (around 327 K) than in the proxy to the realised flow (around 325 K). Comparing Fig. 11a and b, the peak of the distribution is broader than in the case of the proxy to the realised flow, spanning 4 K × 0.25 PVU in contrast to 2 K × 0.12 PVU.

These evolution differences result in the joint PDFs for $(\Delta\mathrm{PV}, \Delta\theta)/\Delta t$, where $\Delta t = 48$ h, shown in Fig. 11c. In terms of changes in θ, the broader tails, which during the first 24 hours were restricted to the most frequent values, are now a general feature so that the whole distribution is more elongated in the $\Delta\theta/\Delta t$ direction. In terms of PV modification, the correspondence between PDFs is better for the whole 48-hour period than for the initial 24-hour period (cf. Fig. 9c), during which PV modification was clearly overestimated. This response indicates a

compensation taking place during the second half of the 48-hour period, which is achieved by an overestimation of heating during the last 24 hours.

5 Summary and conclusions

We have performed a detailed investigation of a case study of forecast error development in 320-K PV. The development of forecast error in this case is consistent with the features described in the statistical analysis of forecast error by Gray et al. (2014). Specifically, it is characterised by a reduction in Rossby-wave ridge area with increasing forecast lead time (Figs. 2 and 4a).

It has been shown that the 320-K PV forecast error is associated with errors in WCB forecasts. Trajectories were calculated to locate air parcels in the WCB outflow inside the corresponding Rossby-wave ridge in analysis and forecasts (Fig. 2, right panels). The PAL diagnostic method was then employed to provide a quantitative assessment of the errors in the forecasts of WCBs. The PAL method showed the sudden development of forecast error in WCB during 24 January 2011, characterised by an overestimation of WCB amplitude, a location of the WCB outflow regions too far to the south and east and a resulting underestimation of the PV anomaly at the outflow (Fig. 3). These forecast error characteristics persisted for at least 24 hours, corresponding with the time span of the error in ridge structure over the North Atlantic.

The first part of the analysis was performed on data from the ECMWF high-resolution forecast. However, a key feature of the forecast error development case is its reproducibility across several operational forecast systems, which, after assuming the absence of extremely large errors in the initial and boundary conditions, can be regarded as an indication of common systematic errors across these systems, in addition to the inherent high sensitivity to initial conditions. Indeed it was found that essentially the same forecast error developed in the ECMWF EPS and the Met Office MOGREPS-15, despite differences in forecast system formulations (Table 1). Moreover, the MetUM simulations performed as part of this work also exhibited the same essential forecast error characteristics. These findings are consistent with Gray et al. (2014), who showed that the error development in the control members of ECMWF EPS and MOGREPS-15 exhibit similar behaviour, while the NCEP model (not analysed here) showed different characteristics. It has also been shown here that the reproducibility of error development is not restricted to the control members of the two considered EPSs (Fig. 4b): The forecast error development showed the same qualitative features independently of initial conditions, which suggests that, in addition to inherent chaos, model error is an important contributor to the development of forecast error in this case.

The connection of WCBs with their parent cyclones immediately indicates a link between cyclone development and 320-K PV forecast error development. In this case the large 320-K PV error on 24 January 2011 is related to a cyclone that formed two days before. It has been shown that this cyclone is present in both analysis and forecast (Fig. 5). However, the time of cyclogenesis is slightly delayed in the forecast and the subsequent cyclone intensity is underestimated. Despite this underestimation throughout the cyclone's life cycle, the associated

WCB's intensity is overestimated. This seemingly paradoxical result could be explained by differences in the moisture reservoir the cyclones in the analysis and forecast have access to. Indeed the surface cold front of the forecast cyclone reaches more southern latitudes, where moisture availability is larger, than the analysis cyclone's front. Other aspects that could have had an effect on the intensities of the cyclone and the resulting WCBs include the representation of cold-frontal precipitation bands (Lackmann, 2002) and the representation of the air-sea interactions between the cold front and the Gulf Stream oceanic front (Cione et al., 1993; Li et al., 2002). Even though further analysis would be required to confirm these hypotheses, this case illustrates the highly complex relationship between cyclone intensity and WCB intensity.

The development of forecast error has been divided into three stages: (i) early cyclone development, (ii) cyclone intensification and (iii) upper-level ridge development. Diabatic processes were studied during these three stages to determine key differences between two simulations: a proxy to the realised flow, consisting of three short-term simulations reinitialised every day, and a contemporaneous forecast, consisting of a single free-running simulation initialised at a single time (see Fig. 1).

The analysis of the first stage was carried out for the 24-hour period between 1200 UTC 21 January and 1200 UTC 22 January. During the first stage, the main differences between proxy to the realised flow and contemporaneous forecast were found near the surface. The cyclone that gave rise to the upper-level ridge developed along the southern end of a parent cyclone's trailing cold front (Fig. 6). The main differences appear precisely at this end of the front, which is more active in the proxy to the realised flow, inducing stronger cross-isentropic ascent than in the contemporaneous forecast. The strong cross-isentropic ascent in the proxy to the realised flow is achieved through contributions from the boundary layer scheme, which in the MetUM includes calls to the cloud microphysics parameterisation scheme, to balance the moisture and temperature fields, and the convection parameterisation (Fig. 7). Even though the sources of error are related to processes taking place at or near the surface, a definite attribution is not possible. Errors in surface processes can be due either to incorrect parameterisation of surface fluxes or incorrect specification of boundary conditions (e.g. sea-surface temperatures) or both. The initial forecast error in the cyclone might be related to the interaction of the cold front with the Gulf Stream oceanic front and associated with air-sea exchanges, which are generally difficult to represent in weather prediction models (Moore and Renfrew, 2002) and are important for the cyclone intensification (Cione et al., 1993; Li et al., 2002).

The analysis of the second stage was carried out for the subsequent 24-hour period. During the second stage, differences in air mass redistribution at upper levels (e.g. 320-K isentropic surface) become more evident (Fig. 8). Cross-isentropic ascent is less intense and the incipient ridge already has a more zonal structure in the contemporaneous forecast than in the proxy to the realised flow. The evolution properties of cross-isentropically ascending air within the ridge, immediately below the tropopause dome, were analysed statistically (Fig. 9). This statistical analysis revealed greater variability in heating rates around the most frequent values in the contemporaneous forecast than in the proxy to the realised flow, indicating a spuriously wide

range in the end values of θ in the contemporaneous forecast. This same analysis also revealed a tendency for the destruction of PV in the proxy to the realised flow that was not matched in the contemporaneous forecast.

The analysis of the third stage was carried out starting from the same time as for the second stage, but over a 48-hour period. During the third stage, the upper-level ridge fully develops in both simulations. Important differences appear in the origin of the air that constitutes the ridge in each simulation. The ridge is formed in both simulations mainly by cross-isentropically ascending air. However, the western flank of the ridge in the proxy to the realised flow exhibits cross-isentropically descending air; this feature is not present within the ridge in the contemporaneous forecast (Fig. 10). The cross-isentropic descent was accompanied by reduction of PV air in an interplay between the radiation parameterisation scheme inducing cooling and the cloud microphysics parameterisation scheme inducing a reduction in PV at upper levels. A statistical analysis of the evolution properties of cross-isentropically ascending air within the ridge, immediately below the tropopause dome, was also carried out for this stage (Fig. 11). This analysis also showed greater variability in heating rates in the contemporaneous forecast than in the proxy to the realised flow, and therefore the production of spuriously wide range in the end values of θ in the contemporaneous forecast. The modification of PV over the 48-hour period exhibits similar distribution between the proxy to the realised flow and the contemporaneous forecast, indicating that the underestimation of PV destruction during the first 24-hour period was corrected at the expense of an overestimation during the second 24-hour period.

Our results, linking the occurrence of error in forecasts of 320-K PV with errors in WCB forecasts and errors in the modification of θ and PV, support the conclusions of previous studies which point towards the importance of the representation of latent heat release in WCBs, driving changes in θ and determining the isentropic level of the WCB outflow (e.g. Martínez-Alvarado et al., 2014). Recent theoretical work (Methven, 2015), case studies (Wernli, 1997; Joos and Wernli, 2012) and climatological studies (Wernli and Davies, 1997; Madonna et al., 2014) show that the net modification of PV along the WCB will be small independently of the diabatic processes that the WCB air might be subject to. However, our results also show that there are clear differences in the net PV modification by the WCB between an approximation to the realised flow and a contemporaneous forecast. The joint differences in the modification of θ and PV are indicative of major differences in tropopause structure (e.g. major differences in ridge amplitude in the case presented here). Thus, we go one step further by showing that indeed differences in θ and PV modification translate into very different atmospheric evolution paths, one close to reality and one far away from it. One of the main assumptions in this work is that the model behaviour is close to the atmosphere's behaviour for short forecasting times (up to two days). Comparisons between the short-term forecasts and analyses at the end of the forecasting period (not shown) provide us with reassurance about the validity of this assumption. Understanding the development of error in one system might lead to understanding the error development in other systems. The challenge is to separate the generic findings from those particular to the forecast system under scrutiny.

There are several questions arising from this work. (1) What is the relative importance of individual parameterised/diabatic processes to the development of forecast error in this case study? (2) Here we have shown that errors in mean sea level pressure led to errors in the WCB and then to errors in the 320-K PV. Is this link systematic and is it possible to systematically link other forecast error types in routes to forecast error? (3) Different weather patterns are very likely to produce different error developments, e.g. cyclone clustering leads to unbroken wave trains that remain for several days unlike the case discussed here, in which the wave grows almost in isolation. Thus, what other routes to forecast error are there? (4) If indeed there are other routes to forecast error, what is the importance of the route to error described here relative to others in terms of frequency of occurrence and impacts? We leave these questions as motivation for future work.

Acknowledgements

References

Arakawa, A. and Lamb, V. (1977). Computational design of the basic dynamical processes of the UCLA General Circulation Model. *Methods Comput. Phys.*, 17:173–265.

Berner, J., Jung, T., and Palmer, T. N. (2012). Systematic model error: The impact of increased horizontal resolution versus improved stochastic and deterministic parameterizations. *J. Climate*, 25:4946–4962.

Bishop, C. H., Etherton, B. J., and Majumdar, S. J. (2001). Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspects. *Mon. Weather Rev.*, 129:420–436.

Bowler, N. E., Arribas, A., Beare, S. E., Mylne, K. R., and Shutts, G. J. (2009). The local ETKF and SKEB: Upgrades to the MOGREPS short-range ensemble prediction system. *Q. J. R. Meteorol. Soc.*, 135:767–776.

Bowler, N. E., Arribas, A., Mylne, K. R., Robertson, K. B., and Beare, S. E. (2008). The MOGREPS short-range ensemble prediction system. *Q. J. R. Meteorol. Soc.*, 134:703–722.

Buizza, R., Milleer, M., and Palmer, T. (1999). Stochastic representation of model uncertainties in the ECMWF ensemble prediction system. *Q. J. R. Meteorol. Soc.*, 125:2887–2908.

Chagnon, J. M. and Gray, S. L. (2009). Horizontal potential vorticity dipoles on the convective storm scale. *Q. J. R. Meteorol. Soc.*, 135:1392–1408.

Chagnon, J. M. and Gray, S. L. (2015). A diabatically-generated potential vorticity structure near the extratropical tropopause in three simulated extratropical cyclones. *Mon. Wea. Rev.*, in press.

Chagnon, J. M., Gray, S. L., and Methven, J. (2013). Diabatic processes modifying potential vorticity in a North Atlantic cyclone. *Q. J. R. Meteorol. Soc.*, 139:1270–1282.

Charney, J. G. and Phillips, N. A. (1953). Numerical integration of the quasi-geostrophic equations for barotropic and simple baroclinic flows. *J. Meteor.*, 10(2):71–99.

Cione, J. J., Raman, S., and Piatrafesa, L. J. (1993). The effect of Gulf Stream-induced baroclinicity on US East Coast winter cyclones. *Mon. Weather Rev.*, 121:421–430.

Davies, H. C. and Didone, M. (2013). Diagnosis and dynamics of forecast error growth. *Mon. Weather Rev.*, 141:2483–2501.

Davies, T., Cullen, M. J. P., Malcolm, A. J., Mawson, M. H., Staniforth, A., White, A. A., and Wood, N. (2005). A new dynamical core for the Met Office's global and regional modelling of the atmosphere. *Q. J. R. Meteorol. Soc.*, 131:1759–1782.

Davis, C. A. (1992). A potential-vorticity diagnosis of the importance of initial structure and condensational heating in observed extratropical cyclogenesis. *Mon. Weather Rev.*, 120:2409–2428.

Dickinson, M. J., Bosart, L. F., Bracken, W. E., Hakim, G. J., Schultz, D. M., Bedrick, M. A., and Tyle, K. R. (1997). The March 1993 Superstorm cyclogenesis: Incipient phase synoptic- and convective-scale flow interaction and model performance. *Mon. Weather Rev.*, 125:3041–3072.

Edwards, J. and Slingo, A. (1996). Studies with a flexible new radiation code. Part I: Choosing a configuration for a large-scale model. *Q. J. R. Meteorol. Soc.*, 122:689–719.

Frame, T., Ambaum, M., Gray, S., and Methven, J. (2011). Ensemble prediction of transitions of the north atlantic eddy-driven jet. *Q. J. R. Meteorol. Soc.*, 137:1288–1297.

Grams, C. M., Wernli, H., Böttcher, M., Campa, J., Corsmeier, U., Jones, S. C., Keller, J. H., Lenz, C.-J., and Wiegand, L. (2011). The key role of diabatic processes in modifying the upper-tropospheric wave guide: a North Atlantic case-study. *Q. J. R. Meteorol. Soc.*, 137:2174–2193.

Gray, S. L. (2006). Mechanisms of midlatitude cross-tropopause transport using a potential vorticity budget approach. *J. Geophys. Res.*, 111(D17113):14 pp.

Gray, S. L., Dunning, C., Methven, J., Masato, G., and Chagnon, J. (2014). Systematic model forecast error in Rossby wave structure. *Geophys. Res. Lett.*, 41.

Gregory, D. and Rowntree, P. R. (1990). A mass flux convection scheme with representation of cloud ensemble characteristics and stability-dependent closure. *Mon. Weather Rev.*, 118:1483–1506.

Holloway, C. E., Woolnough, S. J., and Lister, G. M. S. (2013). The effects of explicit versus parameterized convection on the mjo in a large-domain high-resolution tropical case study. Part I: Characterization of large-scale organization and propagation. *J. Atmos. Sci.*, 70:1342–1369.

Joos, H. and Wernli, H. (2012). Influence of microphysical processes on the potential vorticity development in a warm conveyor belt: a case-study with the limited-area model COSMO. *Q. J. R. Meteorol. Soc.*, 138:407–418.

Jung, T. (2005). Systematic errors of the atmospheric circulation in the ECMWF forecasting system. *Q. J. R. Meteorol. Soc.*, 131:1045–1073.

Jung, T., Tompkins, A. M., and Rodwell, M. J. (2005). Some aspects of systematic error in the ECMWF model. *Atmosph. Sci. Lett*, 6:133–139.

Kim, D., Sperber, K., Stern, W., Waliser, D., Kang, I.-S., Maloney, E., Wang, W., Weickmann, K., Benedict, J., Khairoutdinov, M., et al. (2009). Application of MJO simulation diagnostics to climate models. *J. Climate*, 22:6413–6436.

Klocke, D. and Rodwell, M. (2013). A comparison of two numerical weather prediction methods for diagnosing fast-physics errors in climate models. *Q. J. R. Meteorol. Soc.* In Press.

Lackmann, G. M. (2002). Cold-frontal potential vorticity maxima, the low-level jet, and moisture transport in extratropical cyclones. *Mon. Weather Rev.*, 130:59–74.

Langland, R. H., Shapiro, M. A., and Gelaro, R. (2002). Initial condition sensitivity and error growth in forecasts of the 25 January 2000 East Coast snowstorm. *Mon. Weather Rev.*, 130:957–974.

Li, Y., Xue, H., and Bane, J. M. (2002). Air-sea interactions during the passage of a winter storm over the Gulf Stream: A three-dimensional coupled atmosphere-ocean model study. *J. Geophys. Res.*, 107(C11):3200.

Lock, A. P., Brown, A. R., Bush, M. R., Martin, G. M., and Smith, R. N. B. (2000). A new boundary layer mixing scheme. Part I: Scheme description and single-column model tests. *Mon. Weather Rev*, 128:3187–3199.

Ma, H.-Y., Xie, S., Klein, S., Williams, K., Boyle, J., Bony, S., Douville, H., Fermepin, S., Medeiros, B., Tyteca, S., et al. (2014). On the correspondence between mean forecast errors and climate errors in cmip5 models. *Journal of Climate*, 27:1781–1798.

Madonna, E., Boettcher, M., Grams, C., Joos, H., Martius, O., and Wernli, H. (2015). Verification of North Atlantic warm conveyor belt outflows in ECMWF forecasts. *Q. J. R. Meteorol. Soc.*, 141:1333–1344.

Madonna, E., Wernli, H., Joos, H., and Martius, O. (2014). Warm conveyor belts in the ERA-Interim dataset (1979–2010). Part I: Climatology and potential vorticity evolution. *J. Clim.*, 27:3–26.

Manganello, J. V. and Huang, B. (2009). The influence of systematic errors in the Southeast Pacific on ENSO variability and prediction in a coupled GCM. *Climate Dynamics*, 32(7-8):1015–1034.

Martin, G. M., Milton, S. F., Senior, C. A., Brooks, M. E., Ineson, S., Reichler, T., and Kim, J. (2010). Analysis and reduction of systematic errors through a seamless approach to modeling weather and climate. *J. Climate*, 23:5933–5957.

Martínez-Alvarado, O. (2014). Implications of model error for numerical climate prediction. *Nonlin. Processes Geophys. Discuss.*, 1:131–153.

Martínez-Alvarado, O., Joos, H., Chagnon, J., Boettcher, M., Gray, S. L., Plant, R. S., Methven, J., and Wernli, H. (2014). The dichotomous structure of the warm conveyor belt. *Q. J. R. Meteorol. Soc.*, 140:1809–1824.

Martínez-Alvarado, O. and Plant, R. S. (2014). Parameterised diabatic processes in numerical simulations of an extratropical cyclone. *Q. J. R. Meteorol. Soc.*, 140:1742–1755.

Methven, J. (2015). Potential vorticity in warm conveyor belt outflow. *Q. J. R. Meteorol. Soc.*, 141:1065–1071.

Molteni, F., Buizza, R., Palmer, T. N., and Petroliagis, T. (1996). The ECMWF ensemble prediction system: Methodology and validation. *Q. J. R. Meteorol. Soc.*, 122:73–119.

Moore, G. W. K. and Renfrew, I. A. (2002). An assessment of the surface turbulent heat fluxes from the NCEP-NCAR Reanalysis over the Western Boundary Currents. *J. Clim.*, 15:2020–2037.

Park, Y.-Y., Buizza, R., and Leutbecher, M. (2008). TIGGE: preliminary results on comparing and combining ensembles. *Q. J. R. Meteorol. Soc.*, 134:2029–2050.

Phillips, T. J., Potter, G. L., Williamson, D. L., Boyle, R. T. C. J. S., Fiorino, M., Hnilo, J. J., Olson, J. G., Xie, S., and Yio, J. J. (2004). Evaluating parameterizations in general circulation models: Climate simulation meets weather prediction. *Bull. Amer. Meteor. Soc.*, 85:1903–1915.

Rodwell, M. J., Magnusson, L., Bauer, P., Bechtold, P., Bonavita, M., Cardinali, C., Diamantakis, M., Earnshaw, P., Garcia-Mendez, A., Isaksen, L., et al. (2013). Characteristics of occasional poor medium-range weather forecasts for europe. *B. Am. Meteorol. Soc.* doi: 10.1175/BAMS-D-12-00099.1.

Rodwell, M. J. and Palmer, T. N. (2007). Using numerical weather prediction to assess climate models. *Q.J.R. Meteorol. Soc.*, 133:129–146.

Stevens, B. and Bony, S. (2013). What are climate models missing? *Science*, 1237554:340.

Stoelinga, M. T. (1996). A potential vorticity-based study of the role of diabatic heating and friction in a numerically simulated baroclinic cyclone. *Mon. Weather Rev.*, 124:849–874.

Tennant, W. J., Shutts, G. J., Arribas, A., and Thompson, S. A. (2011). Using a stochastic kinetic energy backscatter scheme to improve MOGREPS probabilistic forecast skill. *Mon. Weather Rev.*, 139:1190–1206.

Tripathi, O. P., Baldwin, M., Charlton-Perez, A., Charron, M., Eckermann, S. D., Gerber, E., Harrison, R. G., Jackson, D. R., Kim, B.-M., Kuroda, Y., Lang, A., Mahmood, S., Mizuta, R., Roff, G., Sigmond, M., and Son, S.-W. (2015). The predictability of the extratropical stratosphere on monthly time-scales and its impact on the skill of tropospheric forecasts. *Q.J.R. Meteorol. Soc.*, 141:987–1003.

Vannière, B., Guilyardi, E., Toniazzo, T., Madec, G., and Woolnough, S. (2014). A systematic approach to identify the sources of tropical SST errors in coupled models using the adjustment of initialised experiments. *Clim. Dyn.*, pages 1–22.

Wernli, H. (1997). A Lagrangian-based analysis of extratropical cyclones. II: A detailed case-study. *Q. J. R. Meteorol. Soc.*, 123:1677–1706.

Wernli, H. and Davies, H. C. (1997). A Lagrangian-based analysis of extratropical cyclones. I: The method and some applications. *Q. J. R. Meteorol. Soc.*, 123:467–489.

Wernli, H., Paulat, M., Hagen, M., and Frei, C. (2008). SAL–A novel quality measure for the verification of quantitative precipitation forecasts. *Mon. Weather Rev.*, 136:4470–4487.

Whitehead, J. P., Jablonowski, C., Kent, J., and Rood, R. B. (2014). Potential vorticity: measuring consistency between GCM dynamical cores and tracer advection schemes. *Q. J. R. Meteorol. Soc.* In Press.

Wilson, D. R. and Ballard, S. P. (1999). A microphysically based precipitation scheme for the UK Meteorological Office Unified Model. *Q. J. R. Meteorol. Soc.*, 125:1607–1636.

Zhang, F., Bei, N., Rotunno, R., Snyder, C., and Epifanio, C. C. (2007). Mesoscale predictability of moist baroclinic waves: convection-permitting experiments and multistage error growth dynamics. *J. Atmos. Sci.*, 64:3579–3594.

Zhang, F., Snyder, C., and Rotunno, R. (2003). Effects of moist convection on mesoscale predictability. *J. Atmos. Sci.*, 60:1173–1185.

Table 1: Operational forecasting systems included in the comparison. (RTG_SST stands for Real-time, global, sea surface temperature analysis; OSTIA stands for Operational Sea Surface Temperature and Sea Ice Analysis.)

Model	Version	Discretization	Horizontal grid spacing	Vertical levels (model top)	SST
ECMWF high resolution	CY36r4	Spectral	T1279 (~ 20 km)	91 (0.01 hPa)	OSTIA / RTG_SST
ECMWF EPS	CY36r4	Spectral	T639 (~ 40 km)	62 (5 hPa)	OSTIA / RTG_SST
MOGREPS-15	10 (MetUM 7.6)	Finite differences	N216 (~ 60 km)	70 (80 km, ~ 0.005 hPa)	OSTIA
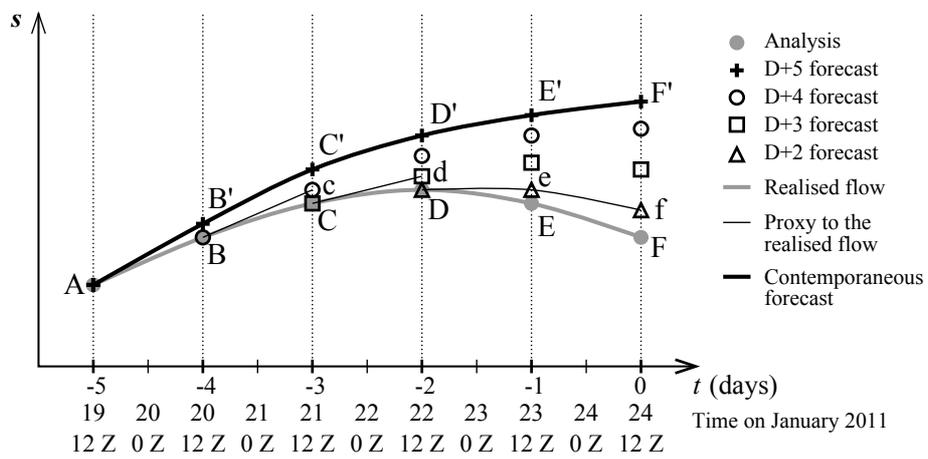
Figure 1: Schematic describing the sequence of simulations to produce the proxy to the realised flow and the contemporaneous forecast. s represents the model's phase space.
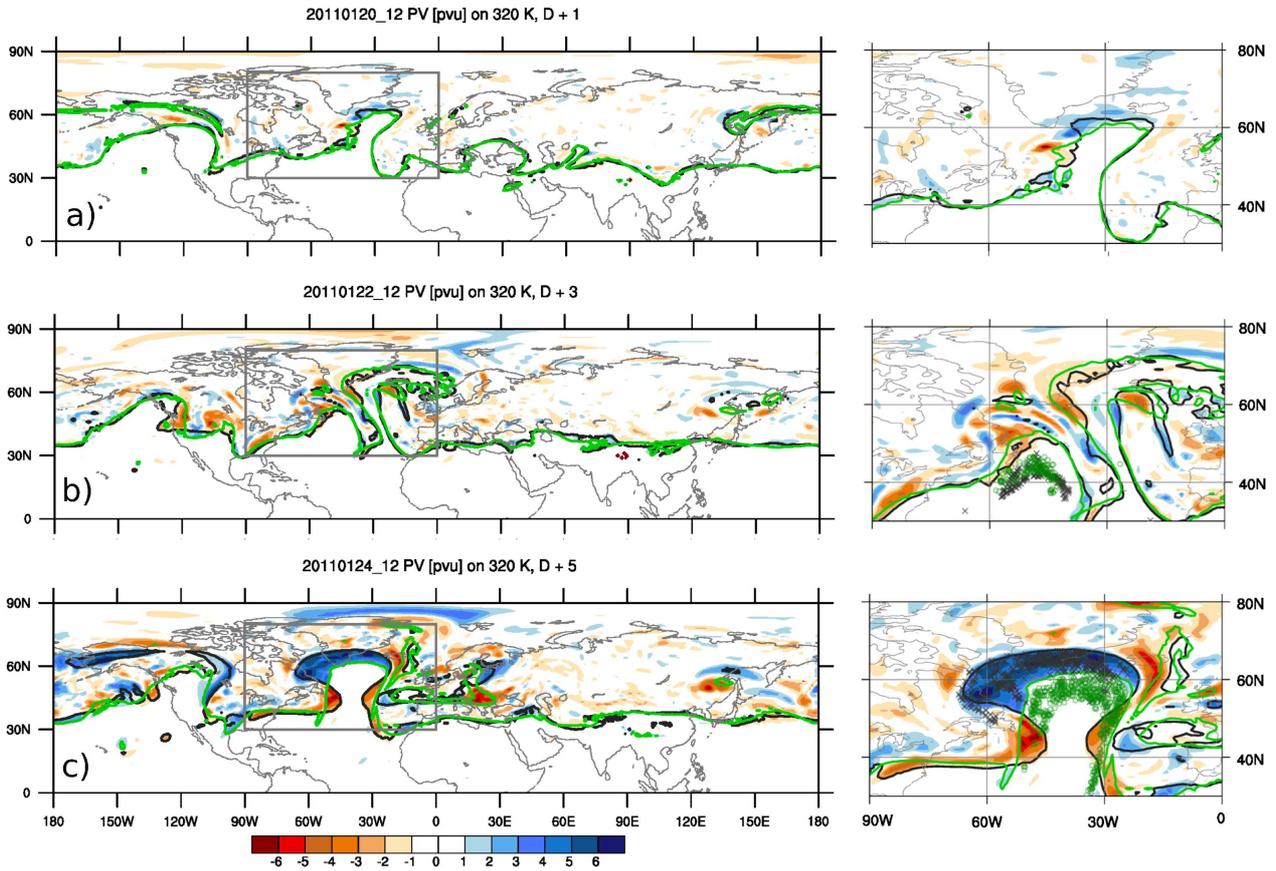
Figure 2: Forecast error (in PVU, colour shading) and 2-PVU contours in the analysis (black) and in the forecast with initialisation time 1200 UTC 19 January 2011 (green) valid at 1200 UTC on (a) 20 January 2011 (D+1), (b) 22 January 2011 (D+3) and (c) 24 January 2011 (D+5). All the fields are plotted on the 320-K isentropic surface. The right panels show zoomed-in details inside the grey boxes (30°N–80°N, 90°W–0°W) marked in the left panels. They also show WCB intersection points with the 320-K isentropic surface at the given time in the analysis (black crosses) and the forecast (green circles), independent of when they started to ascend.

PAL values as function of time for the fc initialised at 12 UTC 19 January 2011
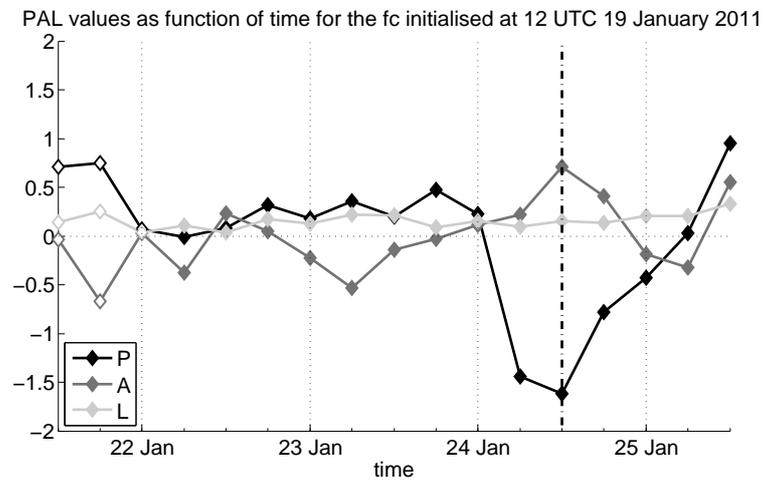
Figure 3: PAL time series for the forecast initialised at 1200 UTC 19 January 2011. The x-axis shows the validation time every 6 hours and corresponds to the WCB outflow time. The open symbols denote times when less than 100 WCB trajectories are found. The dashed line indicates the 1200 UTC 24 January 2011, when the PV error is largest (cf. Fig. 2c).
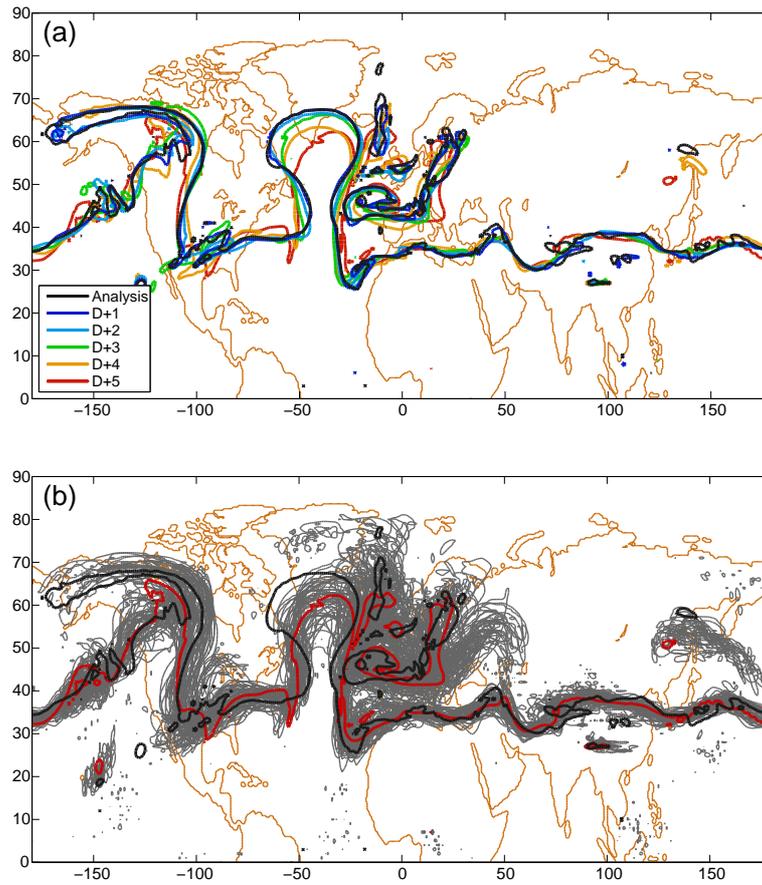
Figure 4: 2-PVU contours on the 320-K isentropic surface valid at 1200 UTC 24 January 2011 in the ECMWF EPS showing **(a)** control members for forecasts with different lead times and **(b)** analysis (black), control member (red) and ensemble members (grey) for the D+5 forecast.

Figure 5: 320-K 2-PVU contours (bold lines) and mean sea level pressure (thin lines, in hPa) with a contour interval of 4 hPa in **(a,c,e)** the proxy to the realised flow and **(b,d,f)** forecasts valid on **(a,b)** 1200 UTC 22 January 2011 **(c,d)** 1200 UTC 23 January 2011 and **(e,f)** 1200 UTC 24 January 2011. Forecasts initialised at 1200 UTC 19 January 2011.
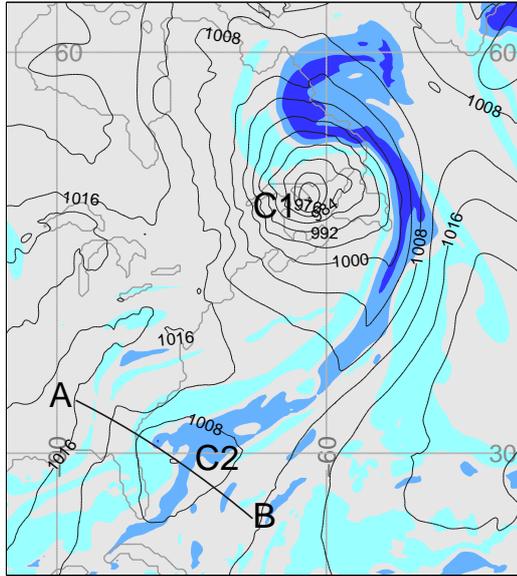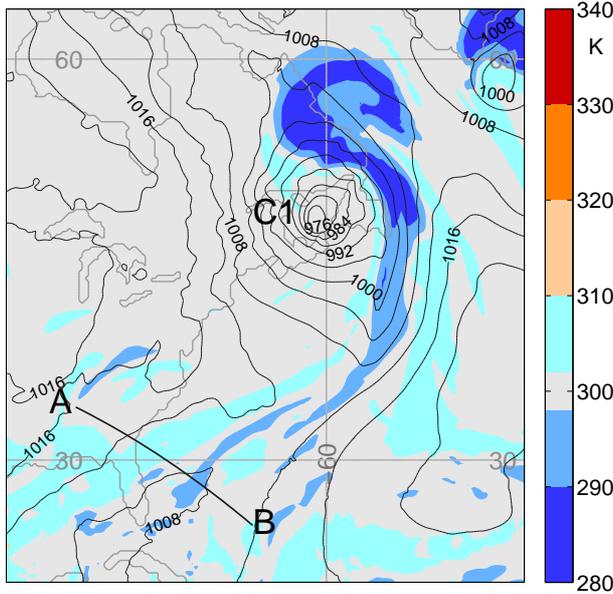
Figure 6: Air mass redistribution on the 300-K isentropic surface in **(a)** proxy to realised flow and **(b)** contemporaneous forecast for the period 1200 UTC 21 January 2011 to 1200 UTC 22 January 2011. The initialisation time for the contemporaneous forecast is 1200 UTC 19 January 2011. Thin lines represent contours of mean sea level pressure, in hPa, with a contour interval of 4 hPa. The colour scale has been chosen to make this figure comparable to Figs. 8 and 10. The gray shading indicates the nominal value of the isentropic surface. Colours above (below) the gray shading on the colour scale represent cross-isentropic descent (ascent). The segment A–B defines the position of the cross-sections in Fig. 7. C1 denotes the low-pressure system over Nova Scotia and C2 (in **(a)**) denotes the incipient cyclone.
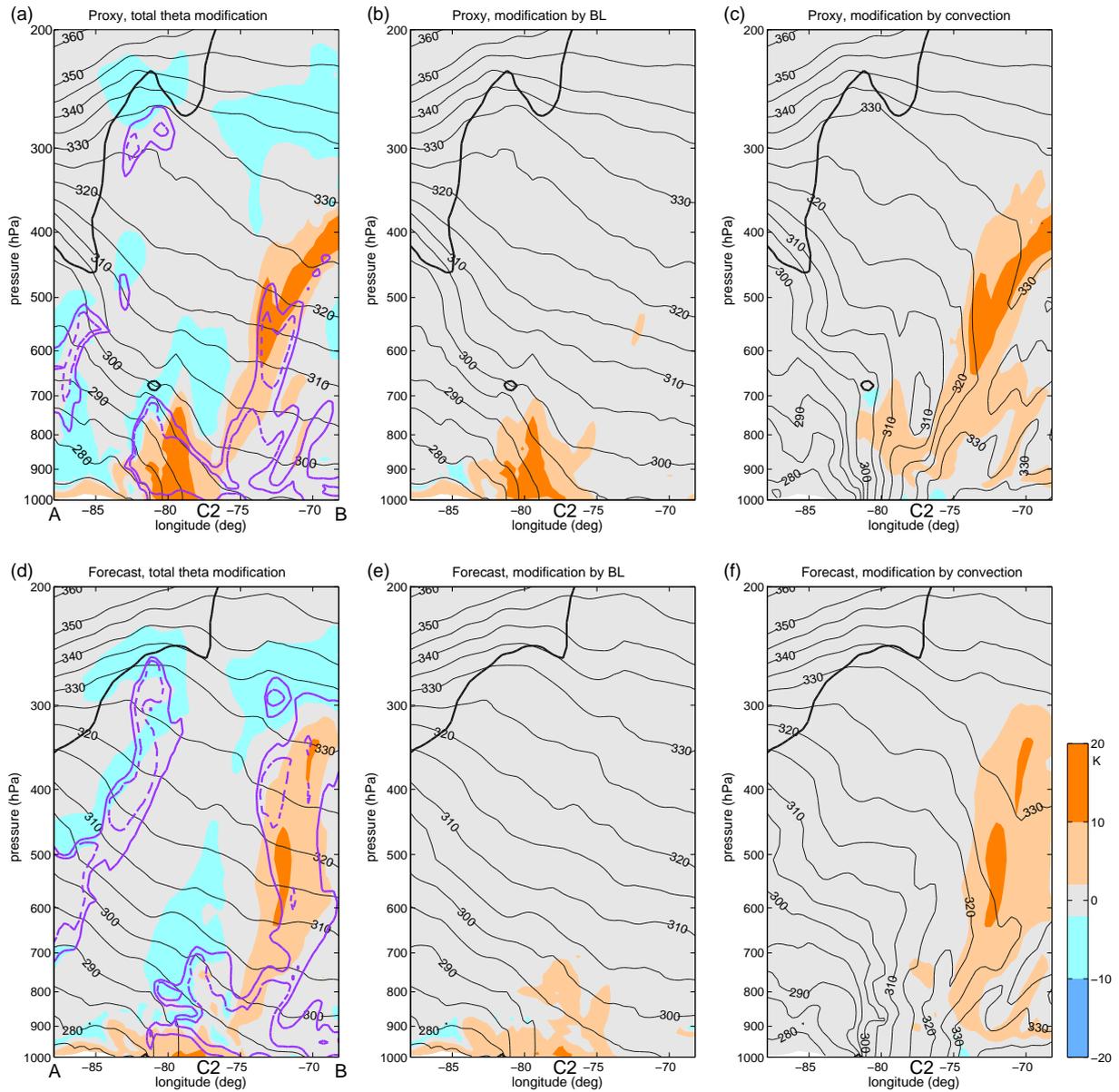
Figure 7: Cross-sections along segment A–B in Fig. 6 showing θ modification in **(a,b,c)** proxy to realised flow and **(d,e,f)** contemporaneous forecasts for the period 1200 UTC 21 January 2011 to 1200 UTC 22 January 2011: **(a,d)** total modification, **(b,e)** modification due to the boundary layer parameterisation, and **(c,f)** modification due to the convection parameterisation. Initialisation time for the forecast is 1200 UTC 19 January 2011. The cyclone's centre along the section is marked by the label C2. Thin black lines represent θ contours (in **(a,b,d,e)**) and equivalent potential temperature (in **(c,f)**), in K, with a contour interval of 5 K; bold black lines represent 2-PVU PV contours, and purple lines in **(a,d)** represent relative humidity with respect to ice (solid – 80% contour; dashed – 90% contour).

33

(a) Proxy, 23 January 2011 1200 UTC

(b) Forecast, 23 January 2011 1200 UTC
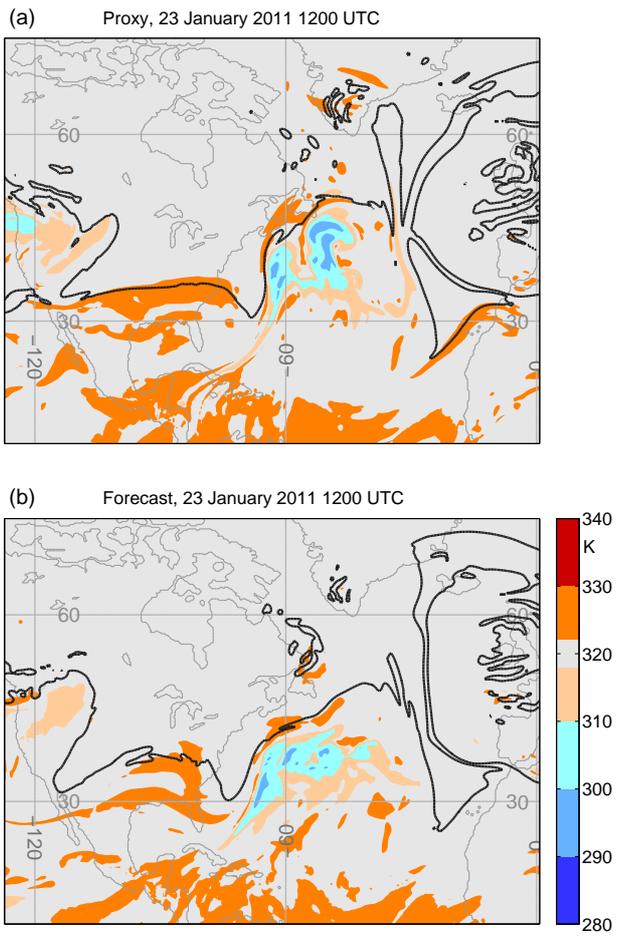
Figure 8: As in Fig. 6, but for the 320-K isentropic surface and for the period 12 UTC 22 January 2011 to 1200 UTC 23 January 2011.

34
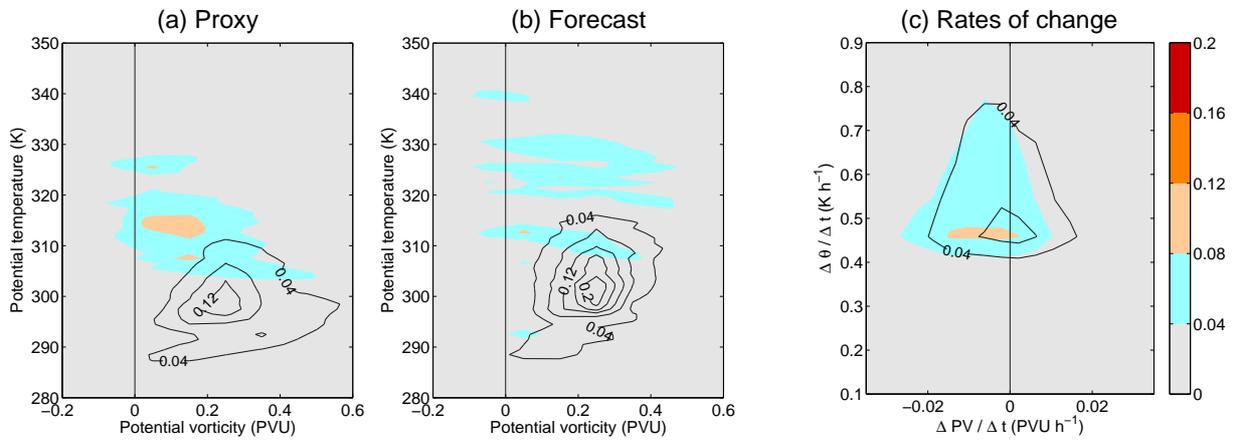
Figure 9: Joint PDFs for the period 1200 UTC 22 January 2011 to 1200 UTC 23 January 2011: **(a)** PV versus θ in the proxy to the realised flow, showing the distribution at the start (black contours) and at the end of the period (colour shading); **(b)** as in (a) but for the contemporaneous forecast; **(c)** rate of change in PV versus rate of change in θ in the proxy to the realised flow (colour shading) and in the contemporaneous forecast (black contours). The colour scale has been chosen to make this figure comparable to Fig. 11. All the PDF portions shown represent at least 85% probability.

(a) Proxy, 24 January 2011 1200 UTC

(b) Forecast, 24 January 2011 1200 UTC
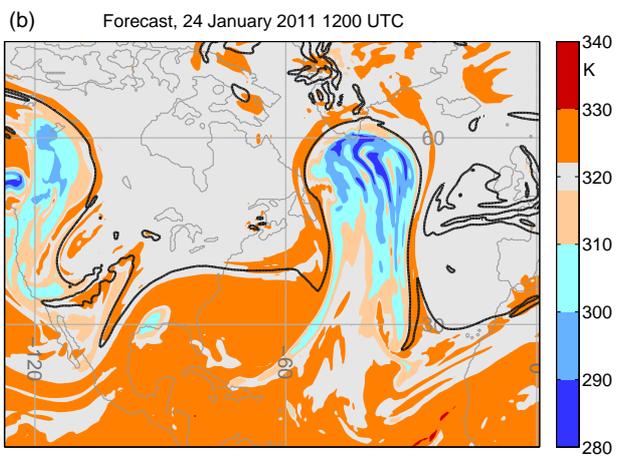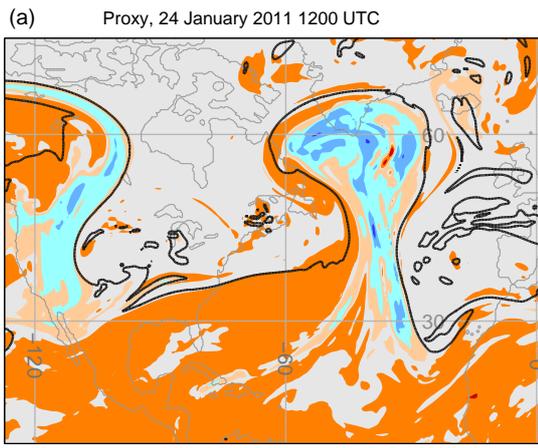
Figure 10: As in Fig. 6, but for the 320-K isentropic surface and for the period 12 UTC 22 January 2011 to 1200 UTC 24 January 2011.
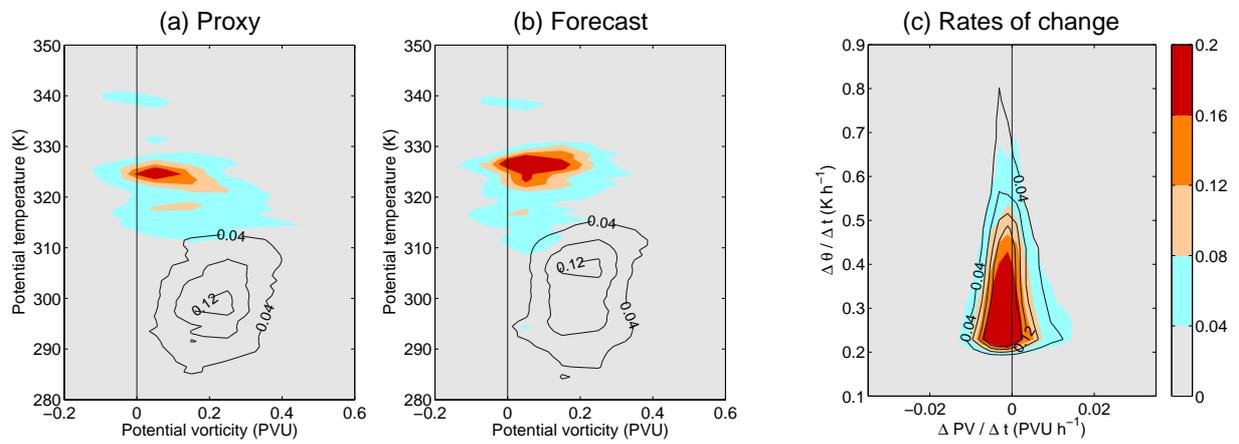
Figure 11: As in Fig. 9, but for the period 1200 UTC 22 January 2011 to 1200 UTC 24 January 2011.