

# *Systematic model forecast error in Rossby wave structure*

Article

Published Version

Creative Commons: Attribution 3.0 (CC-BY)

Open Access

Gray, S. L. ORCID: <https://orcid.org/0000-0001-8658-362X>,  
Dunning, C. M. ORCID: <https://orcid.org/0000-0002-7311-7846>, Methven, J. ORCID: <https://orcid.org/0000-0002-7636-6872>, Masato, G. and Chagnon, J. M. (2014) Systematic model forecast error in Rossby wave structure. *Geophysical Research Letters*, 41 (8). pp. 2979-2987. ISSN 0094-8276 doi: 10.1002/2014GL059282 Available at <https://centaur.reading.ac.uk/36404/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

Published version at: <http://dx.doi.org/10.1002/2014GL059282>

To link to this article DOI: <http://dx.doi.org/10.1002/2014GL059282>

Publisher: American Geophysical Union

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

[www.reading.ac.uk/centaur](http://www.reading.ac.uk/centaur)

**CentAUR**

Central Archive at the University of Reading

Reading's research outputs online

## RESEARCH LETTER

10.1002/2014GL059282

## Key Points:

- Tropopause sharpness adjacent to ridges decreases with forecast lead time
- Rossby wave amplitude also decreases with forecast lead time
- Systematic errors are consistent with misrepresentation of diabatic processes

## Supporting Information:

- ridge\_error\_supplementary.pdf
- Figure S1
- Figure S2
- Figure S3

## Correspondence to:

S. L. Gray,  
s.l.gray@rdg.ac.uk

## Citation:

Gray, S. L., C. M. Dunning, J. Methven, G. Masato, and J. M. Chagnon (2014), Systematic model forecast error in Rossby wave structure, *Geophys. Res. Lett.*, 41, 2979–2987, doi:10.1002/2014GL059282.

Received 13 JAN 2014

Accepted 24 MAR 2014

Accepted article online 27 MAR 2014

Published online 16 APR 2014

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

## Systematic model forecast error in Rossby wave structure

S. L. Gray<sup>1</sup>, C. M. Dunning<sup>1</sup>, J. Methven<sup>1</sup>, G. Masato<sup>1</sup>, and J. M. Chagnon<sup>1</sup>
<sup>1</sup>Department of Meteorology, University of Reading, Reading, UK

**Abstract** Diabatic processes can alter Rossby wave structure; consequently, errors arising from model processes propagate downstream. However, the chaotic spread of forecasts from initial condition uncertainty renders it difficult to trace back from root-mean-square forecast errors to model errors. Here diagnostics unaffected by phase errors are used, enabling investigation of systematic errors in Rossby waves in winter season forecasts from three operational centers. Tropopause sharpness adjacent to ridges decreases with forecast lead time. It depends strongly on model resolution, even though models are examined on a common grid. Rossby wave amplitude reduces with lead time up to about 5 days, consistent with underrepresentation of diabatic modification and transport of air from the lower troposphere into upper tropospheric ridges, and with too weak humidity gradients across the tropopause. However, amplitude also decreases when resolution is decreased. Further work is necessary to isolate the contribution from errors in the representation of diabatic processes.

## 1. Introduction

The Observing System Research and Predictability Experiment (THORPEX) of the World Meteorological Organization grew out of a recognition that, despite continued improvements in numerical weather prediction forecasts (THORPEX specifically considered 1–14 day lead times), further improvements needed to be made. However, the strong influence of initial condition uncertainty on ensemble forecast spread and error diagnostics renders it difficult to isolate systematic contributions from model error and the physical processes responsible. Identification of such systematic errors is the first step to determining model improvements to reduce them.

The approach taken here is to exploit the Lagrangian and global conservation properties of potential vorticity (PV) to isolate systematic forecast errors associated with model formulation (as opposed to initial condition uncertainty). PV is materially conserved in adiabatic, frictionless flow, and the tropopause is identified with a strong PV gradient on isentropic surfaces. Taken together, these facts imply that an air mass with anomalous PV, relative to its surroundings, arises primarily through adiabatic advection of the tropopause PV gradient. Patterns of disturbances of PV contours away from zonal symmetry are described as “large-amplitude Rossby waves,” even though they can be far from sinusoidal in structure.

Diabatic processes in extratropical cyclones can significantly modify tropopause-level ridges and troughs. Chagnon *et al.* [2013] used a PV-tracer technique to show that diabatic processes create positive “diabatic PV anomalies” above and negative diabatic PV below the tropopause. Although not directly modifying the tropopause elevation in their case study, the resulting enhancement of the PV gradient across the tropopause (tropopause sharpening) would contribute to greater Rossby wave amplitude at the tropopause and a stronger tropopause-level jet. Both effects would modify the downstream Rossby wave development via advection and propagation. Other work [e.g., Plant *et al.*, 2003; Stoelinga, 1996] has proposed that the diabatically reduced upper tropospheric PV directly modifies the tropopause structure by erosion of the tropopause trough and associated enhancement of the upper level ridge or retardation of the trough due to advection of the PV gradient by the divergent flow.

Different processes contribute to the diabatic PV anomalies. The positive diabatic PV above the tropopause arises from longwave cooling just below the sharp humidity gradient at the tropopause; this cooling also gives rise to negative diabatic PV below the tropopause. However, Chagnon *et al.* [2013] showed that air leaving regions of latent heating associated with resolved ascent, parameterized convection, or boundary layer processes also contribute to the negative diabatic PV. Most of this ascending air is associated with the warm conveyor belt of cyclones [Browning and Roberts, 1994]: a strong flow of warm wet-bulb potential

temperature air advancing poleward ahead of the cold front. The outflow from warm conveyor belts can reach the upper troposphere across a range of potential temperature,  $\theta$ , values. For example, *Martínez-Alvarado et al.* [2014] used parcel trajectories to illustrate the splitting of a warm conveyor belt into a higher-level (higher  $\theta$ ) anticyclonically turning flow and a lower level (lower  $\theta$ ) cyclonically turning flow. The outflow level is sensitive to the representation of diabatic processes in the model with consequences for the magnitude of the PV anomaly, Rossby wave propagation, and wave-breaking behavior.

Errors in the representation of diabatic processes in extratropical cyclones, perhaps resulting from the necessity to parameterize convection and other diabatic processes in global (and most regional) weather forecast models, could thus lead to errors in forecasts. *Brennan et al.* [2008] advocate PV-based interpretation in operational forecasting to identify diabatically driven parts of the model solutions that might thus be associated with increased uncertainty. A failure to forecast Rossby wave breaking can result in so-called “forecast busts.” For example, *Rodwell et al.* [2013] hypothesize that misrepresentation of diabatic processes within mesoscale convective systems across North America leads to the most extreme forecast busts over Europe.

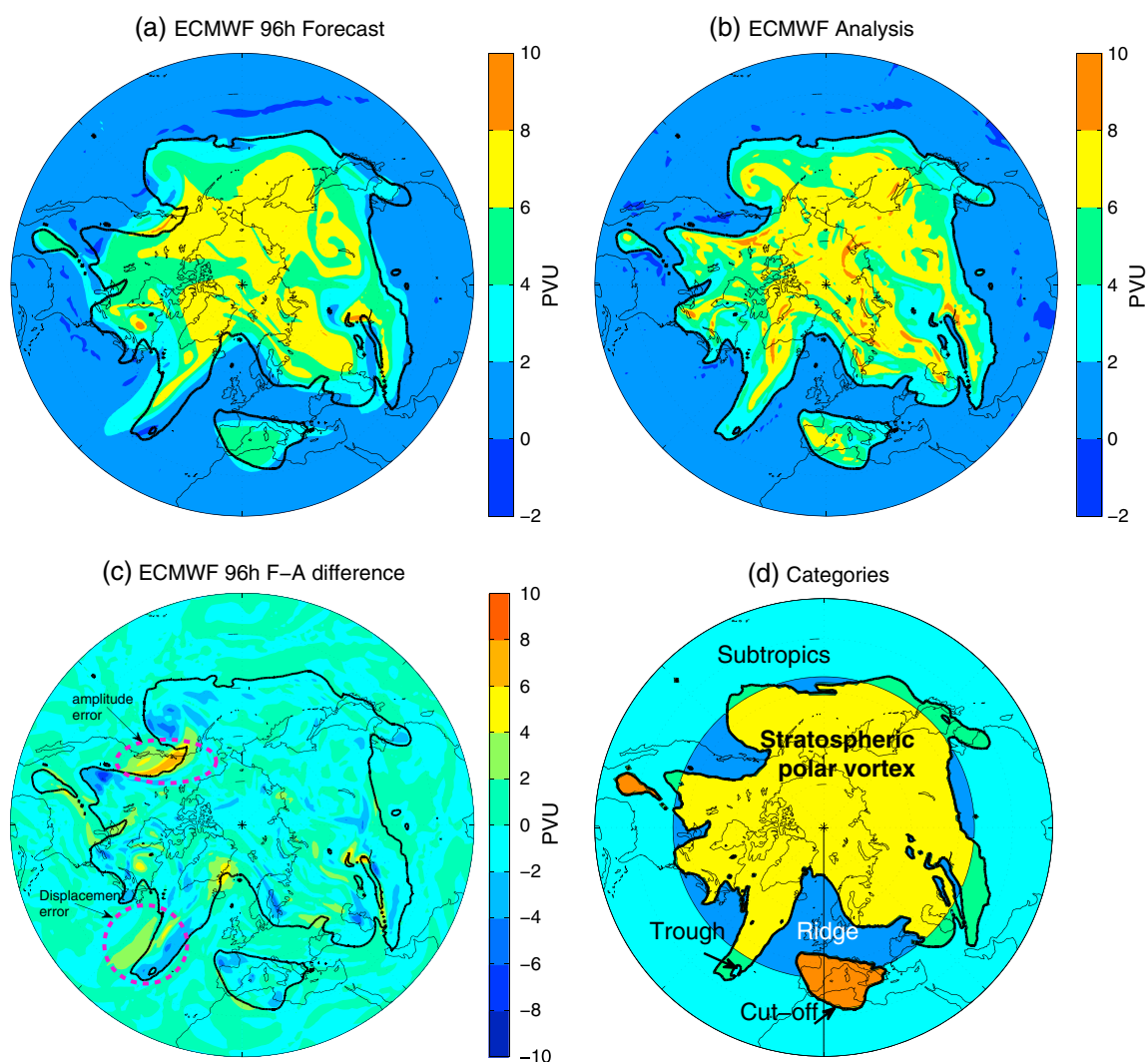
Previous research [*Dirren et al.*, 2003; *Davies and Didone*, 2013] using European Centre for Medium-Range Weather Forecasts (ECMWF) operational forecasts from the winter of 2001–2002 strikingly demonstrates the development of errors in forecast Rossby wave structure over several days lead time by comparing forecast and analysis PV fields. *Davies and Didone* [2013] consider five mechanisms for generating and/or enhancing Rossby waves; two of these mechanisms, stratospheric PV anomaly, and tropospheric PV error, relate to the diabatic generation of PV described above. However, these authors could not partition the forecast errors by process because they focused on root-mean-square (RMS) errors in PV that are dominated by phase errors.

In this paper we identify and characterize systematic errors in Rossby wave structure by using diagnostics that are not impacted by phase error and are therefore less dominated by initial condition uncertainty. The TIGGE (THORPEX Interactive Grand Global Ensemble) archive [*Park et al.*, 2008] is utilized to examine 15 day forecasts from three operational weather forecast centers. An example of the structure of forecast errors is shown in Figure 1 for a lead time of 96 h. Figure 1c shows the difference between the PV in the forecast valid at the analysis time (Figure 1a) and the analysis (Figure 1b). Two characteristic types of error occur: (i) a phase error, e.g., the forecast trough over the North Atlantic is to the west of the analyzed position, and (ii) an amplitude error, e.g., the ridge over the west U.S. coast is much less developed in the forecast than in the analysis yielding a positive forecast minus analysis difference in PV. A third type of error can also occur that is not illustrated in Figure 1: an error in the isentropic gradient of PV across the tropopause. Here we focus on quantifying errors in total ridge amplitude and tropopause sharpness error around the poleward flank of ridges and link them with forecast model properties such as horizontal resolution and associated numerical dissipation.

The motivation for focusing on ridges rather than troughs is (i) a signal of systematic error due to diabatic processes is likely to be seen more clearly in ridges than in troughs as the outflow from warm conveyor belts intersects (and potentially modifies) the tropopause on the flanks of the ridges [*Martínez-Alvarado et al.*, 2014] and (ii) ridges tend to be much larger scale and less convoluted than troughs, and therefore, amplitude error may be larger than zonal displacement error enabling extraction of model error from the forecast data. We define ridge amplitude by ridge area, rather than the maximum poleward displacement of a PV contour, as we assert this is a more robust measure of wave activity and the potential impact of diabatic processes on Rossby waves for this study since it is not dependent on the contour shape. Also, for large amplitude disturbances a wave activity conservation law can be obtained for wave activity that is related to the mass and circulation enclosed between a disturbed PV contour and its equivalent latitude [e.g., *Methven*, 2013]. In the TIGGE data we do not have a diagnostic of density in isentropic coordinates and therefore can only use ridge area, which is equivalent to mass if density does not vary across ridges.

## 2. Methodology

Daily (12 UTC) Northern Hemisphere (20–90°N) fields of PV on the 320 K isentrope are used from the control members of medium-range ensemble forecasts of three operational centers: ECMWF, the Met Office, and National Centers for Environmental Prediction (NCEP). These were extracted from the TIGGE archive for up to seven winter seasons (December, January, and February from 2006/2007 (2008/2009 for NCEP) to 2012/2013). The resolutions of the forecast models are given in Table 1. For further details on the forecast models and data, the reader is referred to <http://tigge.ecmwf.int>.



**Figure 1.** Forecast and analysis tropopause structure in ECMWF data for 24 December 2006: (a) 96 h forecast, (b) analysis, and (c) 96 h forecast minus analysis of PV on 320 K isentropic (color shading); (d) categorization of regions defined using the equivalent latitude of the tropopause PV contour. The analyzed tropopause position is shown by the bold contour on all maps, and the equivalent latitude is shown by the circular contour in Figure 1d.

The evolution of forecast error with forecast lead time in this PV field is analyzed by partitioning it into five regions defined using the structure of the PV contour identified with the tropopause,  $PV_{trop}$ ; in the results presented here, this was taken to be 2.24 PVU (potential vorticity unit) because it is the average location of the strongest PV gradient along 320 K in the background state. The five categories are the stratospheric polar vortex, the subtropics, troughs, ridges, and cutoff lows. Each grid point is objectively assigned to one of these five categories, as illustrated in Figure 1d.

The first four categories are identified in two steps. The PV at each point is compared with  $PV_{trop}$ . The location of each point is then compared with the equivalent latitude,  $\phi_e$ , associated with  $PV_{trop}$  on the 320 K surface. The categories are stratospheric polar vortex ( $PV > PV_{trop}$ ,  $\phi > \phi_e$ ), trough ( $PV > PV_{trop}$ ,  $\phi < \phi_e$ ), subtropics ( $PV < PV_{trop}$ ,  $\phi < \phi_e$ ), and ridge ( $PV < PV_{trop}$ ,  $\phi > \phi_e$ ). Equivalent latitude was originally defined in studies of the stratosphere using the area enclosed by a PV contour [Butchart and Remsburg, 1986]. The equivalent latitude is the perimeter of a circle centered on the pole with the same area as the wavy PV contour. The method used here to obtain equivalent latitudes involves a simultaneous rearrangement of all PV contours from the meteorological analysis so that they all are zonally symmetric and enclose the same mass and circulation as in the full (wavy) state. The resulting background state defined by the equivalent latitudes of these PV contours is called the modified Lagrangian mean. Methven [2009] and Nakamura and Solomon [2011] show the slow evolution of this state obtained from ERA-Interim, which is a consequence

**Table 1.** Resolutions of the ECMWF, Met Office, and NCEP Models Used for the Operational Forecast Ensembles Stored in the TIGGE Archive<sup>a</sup>

Year	ECMWF	Met Office	NCEP
2006/2007	N200L62/N128L62 <sup>b</sup> (days 1–10/10–15)	N144L38	N64L28 <sup>c</sup>
2007/2008			
2008/2009			
2009/2010	N320L62/N160L62 <sup>d</sup>	N216L70 <sup>f</sup>	N96L28 <sup>e</sup>
2010/2011			
2011/2012			
2012/2013			N128L42/N96L42 <sup>g</sup> (days 1–8/8–15)

<sup>a</sup>L = number of vertical levels. For the ECMWF and NCEP pseudospectral models, T = highest wave number retained in the spherical harmonics (see footnotes) and a linear Gaussian grid is used for transformation into grid point space. N = number of grid points from equator to pole in these models. In all models, 4N is the number of grid points around the equator. Note that all model data were output on the same  $1^\circ \times 1^\circ$  grid for diagnostics.

<sup>b</sup>T399L62/T255L62.

<sup>c</sup>T126L28.

<sup>d</sup>T639L62/T319L62, from 26 January 2010.

<sup>e</sup>T190L28, from 23 February 2010.

<sup>f</sup>from 9 March 2010.

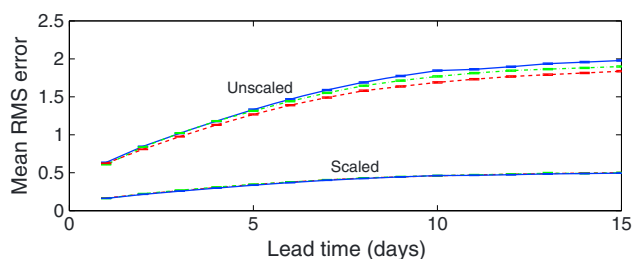
<sup>g</sup>T254L42/T190L42, from 14 February 2012.

of approximate conservation of mass and circulation. *Methven* [2013] has shown that this is a natural state with which to reference large-amplitude Rossby wave activity. The equivalent latitude for the  $PV_{\text{trop}}$  contour was calculated from ERA-Interim analyses for the fifteenth day of each month from one season and linearly interpolated to daily values yielding a slow southward progression over the winter (from  $42.90^\circ\text{N}$  to  $39.36^\circ\text{N}$ ); the same equivalent latitude values are used when comparing forecasts from all centers and for all winter seasons.

After this four-way categorization, cutoff lows are identified as regions of stratospheric PV for which the region enclosed by the delimiting PV contour does not include the pole. Some points that were initially identified as polar vortex or troughs are then recategorized as cutoff lows. Figure 1d shows these categories for the example case. Climatologically, the 320 K isentrope extends from the lower troposphere ( $\sim 700$  hPa) in the subtropics to well into the stratosphere ( $\sim 250$  hPa) near the pole in the Northern Hemisphere winter. Occasionally, this isentrope can intersect the ground, especially over the Himalayas. Where this occurs, the points have been removed from the analysis.

### 3. Errors at the Hemispheric Scale

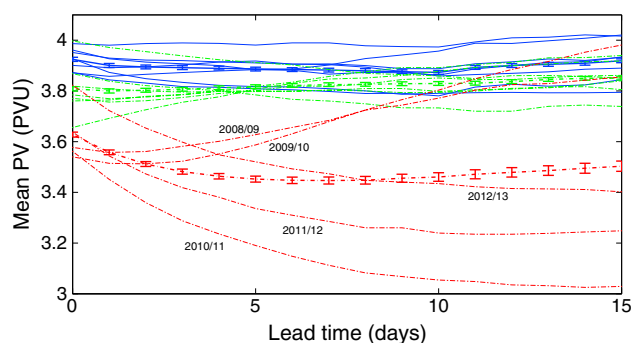
The Northern Hemispheric average of PV forecast error approaches saturation after about 9 days for all three centers (Figure 2). The shape of the error curves as a function of lead time is very similar for the three



**Figure 2.** Average root-mean-square forecast minus analysis difference in PV as a function of forecast lead time for ECMWF (solid blue line), the Met Office (dot-dashed green line), and NCEP (red dashed line): unscaled (upper lines, PVU) and scaled by the mean analysis PV for the given center (lower lines, dimensionless). Error bars are standard errors estimated from the variability.

centers with slight differences in amplitude, and the error curves collapse nearly to a single curve when scaled by mean analysis PV of the given center (this mean analysis PV is shown in Figure 3, where the analysis corresponds to lead time of 0 days). This could be taken as implying that all centers have similar model error. However, it will be shown that this is not the case and the similarity in RMS errors must primarily be associated with uncertainty in the initial conditions for all the forecasts and chaotic behavior.





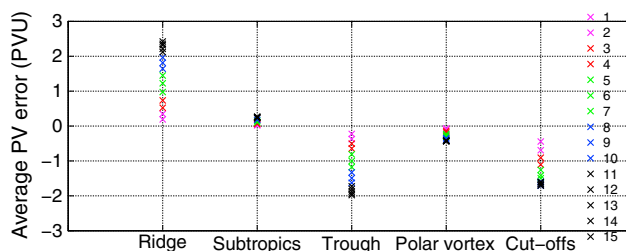
**Figure 3.** Mean PV as a function of forecast lead time for ECMWF (solid blue lines), the Met Office (dot-dashed green lines), and NCEP (red dashed lines). Lines with error bars (standard errors) are averages over all winter seasons. Lines without error bars are averages over individual seasons, labeled for NCEP model only.

PV values are nearly constant (Figure 3) which is more consistent with the slow variation of the background state and maintenance of the stratospheric PV reservoir by radiative cooling against erosion from stirring and mixing of PV to lower latitudes. Interannual variability in the mean PV is also small in these forecasts.

#### 4. Errors in Ridges and Troughs

The forecast PV has a positive bias, where the analysis shows ridges or subtropical air, and a negative bias in the locations of analyzed troughs, cutoff lows, and the stratospheric polar vortex (Figure 4); these biases increase with lead time. This could arise from phase errors alone; e.g., a displaced ridge in a forecast would lead to greater PV values in the location of the analyzed ridge. Saturation occurs when there is little skill in the forecast location of troughs and ridges. Note that PV has many fine-scale features contributing to this forecast error. There is still skill in probability forecasts of the large-scale patterns out to 15 days, as illustrated for the low-level Atlantic jet by *Frame et al.* [2011]. Despite having weaker amplitude (as indicated by the saturation error), the forecast errors in the cutoff lows are larger than those in the troughs (for a given lead time) prior to saturation. This is consistent with cutoff lows generally being smaller-scale features than their parent troughs and so being more difficult to represent accurately at analysis time. The biases are much smaller in the subtropics and polar vortex than in the ridges and troughs, consistent with the much larger areas of these regions and only a small proportion of points in these regions being close to strong isentropic PV gradients.

A reduction of Rossby wave amplitude (i.e., undulations of the tropopause) with forecast lead time can be diagnosed from the reduction in total area of the ridges in the forecast compared with that in the analysis (Figures 5a–5c). The total area of ridges is calculated for each forecast lead time averaged over all verification dates for each winter season. A reduction in ridge area increasingly occurs for lead times up to about 5 days in the multiwinter season mean in ECMWF and Met Office forecasts; this behavior also occurs in varying degrees in each of the individual winter seasons. Beyond about 5 days the evolution of the total ridge area is

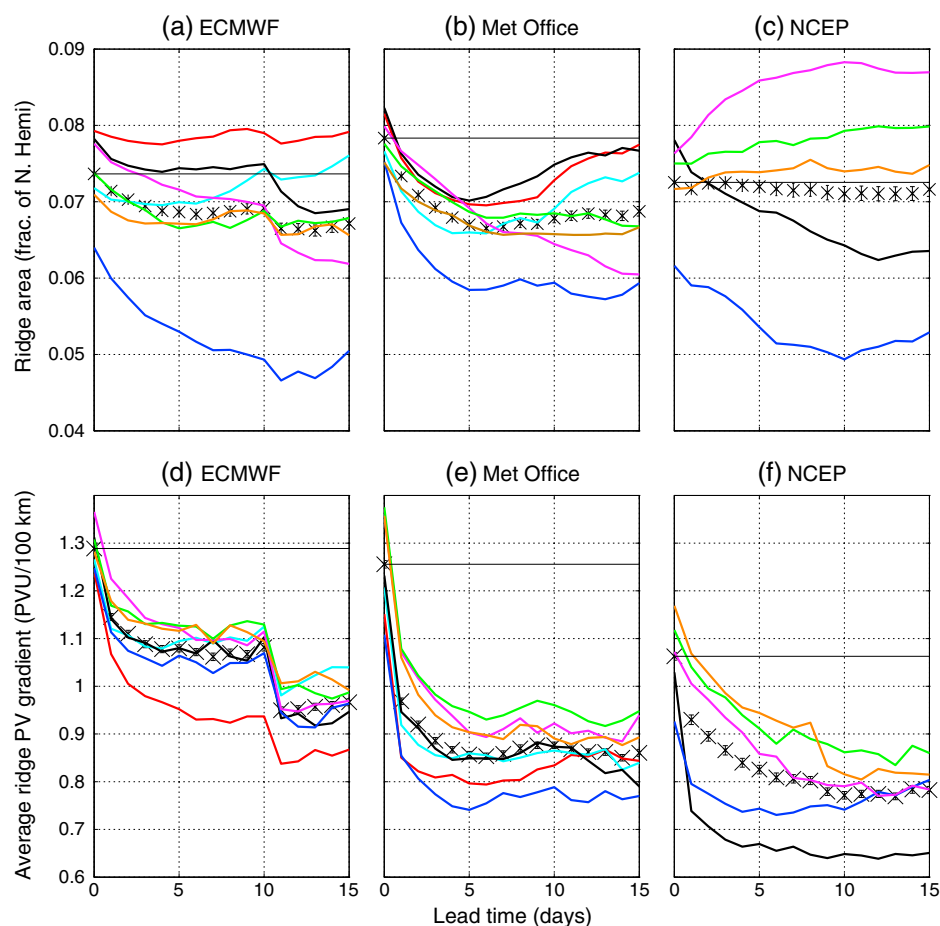


**Figure 4.** Average ECMWF forecast minus analysis difference in PV as a function of forecast lead time (days) for the five region categories illustrated in Figure 1d. The feature categories are defined from the analyses.

For NCEP forecasts, the hemispheric mean PV (multiseason average) is lower than the other centers in the analysis and decreases with forecast lead time out to 5 days on average (Figure 3). However, there is substantial interannual variation and the PV at long lead times is substantially greater for the earlier two winter seasons than for the latter three winter seasons. This change in behavior may be a consequence of an increase in model resolution implemented at the end of the 2009/2010 winter (Table 1) and the associated changes in model configuration. By contrast, in the Met Office and ECMWF forecasts the mean

PV values are nearly constant (Figure 3) which is more consistent with the slow variation of the background state and maintenance of the stratospheric PV reservoir by radiative cooling against erosion from stirring and mixing of PV to lower latitudes. Interannual variability in the mean PV is also small in these forecasts.

different in the different winter seasons (decreasing further in some years and increasing in others). The evolution of total ridge area in the NCEP forecasts is not consistent between the different winter seasons (Figure 5c); similar to the hemispherically averaged PV (Figure 3), the earlier two winter seasons show different behavior to the latter three. A reduction of Rossby wave amplitude with forecast lead time can also be diagnosed from the general reduction



**Figure 5.** (a–c) Average ridge area and (d–f) the isentropic PV gradient flanking ridges as a function of forecast lead time for ECMWF, Met Office, and NCEP. Black markers with error bars (standard errors) are averages over all winter seasons with horizontal lines extending across all lead times from the analysis values. Colored lines are averages for the individual seasons where red is 2006/2007, cyan is 2007/2008, black is 2008/2009, blue is 2009/2010, magenta is 2010/2011, green is 2011/2012, and orange is 2012/2013. Note (as an example) that a fraction of the Northern Hemisphere of 0.05 is equivalent to an area of  $1.275 \times 10^7 \text{ km}^2$ .

in the area of cutoff lows (not shown) since cutoff lows form from wave amplification and breaking. For cutoff lows the NCEP analyses are the outlier with a notably smaller average fraction of the Northern Hemisphere comprised of cutoff lows in these analyses compared to those from ECMWF and the Met Office ( $3.5 \times 10^{-3}$  compared to  $5.6 \times 10^{-3}$  and  $5.2 \times 10^{-3}$ , respectively).

A reduction in the isentropic PV gradient across the tropopause (tropopause sharpness) around the poleward flank of ridges occurs with forecast lead time (Figures 5d–5f). Analogous to Figures 5a–5c, the isentropic PV gradient is calculated for each forecast lead time averaged over all verification dates for each winter season. The isentropic PV gradient is calculated by combining centered differences in the zonal and meridional directions (to calculate the magnitude of PV gradient) at all grid points north of the  $PV_{\text{trop}}$  equivalent latitude, and then the average tropopause gradient is calculated from the values at the grid points intersecting the tropopause at each longitude. The tropopause intersection point, for a given longitude, is taken as the first latitude point, traveling southward from the north pole, where the PV falls below  $PV_{\text{trop}}$  for three consecutive grid points (to avoid identifying small cutoff highs of low PV in the stratospheric vortex), where this point exists. This simple approach neglects the possibility of folded ridges or cutoff highs with meridional extent exceeding  $3^\circ$  and ignores ridges that do not extend poleward of the equivalent latitude by more than  $3^\circ$ . The multiwinter season average of tropopause PV gradient slackens by 15–30% over the first 5–10 days of the forecasts from the three operational centers. This characteristic is seen in each individual season (as well as in forecasts from all three centers) and is more robust than the reduction in ridge area



seen in Figures 5a–5c. In contrast to the RMS error results, this general downward trend with lead time for all models and years does indeed emphasize a systematic model error aspect.

A discontinuity occurs in the ECMWF forecast ridge area and tropopause sharpness at a forecast lead time of 10 days (clearly apparent in the multiseason average but also apparent in most of the individual seasonal averages for the ridge area and all individual seasonal averages for the tropopause sharpness). This is hypothesized to be associated with the reduction of the ECMWF forecast's horizontal resolution at 10 days forecast lead time (Table 1). A discontinuity, potentially associated with resolution reduction, is also seen for the NCEP model at 8 days lead time in the 2012/2013 winter season. All three operational centers have enhanced the resolution of their models during the period of data availability (Table 1). The last three winter seasons show the sharpest analyzed tropopauses, and the least slackening in tropopause sharpness with lead time, in forecasts from all three operational centers. However, there is no systematic effect on the total ridge area of these enhancements.

Tropopause sharpness has also been calculated across the equatorward flank of troughs for  $PV_{\text{trop}} = 2.24$  PVU. This demonstrates a similar reduction with forecast lead time to that shown in Figures 5d–5f for ridges, including the discontinuity at 10 days lead time for the ECMWF forecasts. The tropopause sharpness is systematically stronger flanking ridges than troughs. The trough area decreases with short lead times in forecasts in some winter seasons from all centers for this tropopause PV value, but the decrease is not systematic; however, increasing  $PV_{\text{trop}}$  to 3.35 PVU (which increases the areas of troughs) does yield systematic decreases with lead time in both trough area (except for one winter season) and tropopause sharpness flanking troughs for lead times of about 5 days in forecasts from all three centers. Analysis of ridge area and tropopause sharpness flanking ridges performed for this higher  $PV_{\text{trop}}$  value is similar to that shown in Figure 5 for  $PV_{\text{trop}} = 2.24$  PVU. Please see supporting information for figures.

Data from all centers are interpolated to the same ( $1^\circ \times 1^\circ$ ) grid for output and calculation of the isentropic gradient. Therefore, the change in the gradient reflects the decrease in resolution of the underlying dynamical simulation, not the diagnostic method. The weaker PV gradients in NCEP forecasts reflect their lower resolution (see Table 1). Similarly, the Met Office forecasts have weaker gradients than ECMWF. The ECMWF and NCEP models are pseudospectral, while the Met Office model uses a finite difference scheme and typically has stronger numerical dissipation (for the same resolution) than the other two models. All three models use semi-implicit semi-Lagrangian time-stepping schemes which introduce a degree of numerical dissipation via interpolation to departure points.

## 5. Conclusions and Broader Implications

The drive to improve medium-range weather forecasts via the World Meteorological Organization THORPEX program has led to a focus on the identification of systematic forecast errors [e.g., Jung *et al.*, 2005]. Here we have used potential vorticity to diagnose systematic errors that develop with forecast lead time in winter forecasts from ECMWF, the Met Office, and NCEP extracted from the TIGGE archive. Northern Hemisphere upper level PV forecast errors (forecast minus analysis differences) in these operational global models approach saturation after about 9 days. RMS errors are a blunt tool to investigate model error due to the dominance of phase errors from initial condition uncertainty. The Lagrangian and global conservation properties of potential vorticity (PV) are used here to draw out systematic errors related to model formulation.

The total area covered by ridges in the ECMWF and Met Office forecasts decreases consistently with forecast lead time (up to about 5 days) implying a reduction in Rossby wave amplitude. The results for NCEP vary greatly between years. The total PV on the 320 K isentrope also varies between years in the NCEP forecasts and analyses, while it is much less variable in both the ECMWF and Met Office forecasts. The decrease of average PV with lead time in NCEP forecasts in the later years indicates that the model cannot maintain the stratospheric PV reservoir through radiative cooling against erosion from stirring and mixing of PV to lower latitudes. The PV gradient across the tropopause in ridges decreases sharply with lead time in forecasts from all three centers (and in each winter season) on a similar time scale to the reduction in ridge area in the ECMWF and Met Office forecasts. A second sharp drop in PV gradient is also seen when model resolution is degraded at longer lead times. This implies that the forecast models cannot maintain the sharpness of the isentropic PV gradient at the tropopause as a result of horizontal resolution and numerical dissipation.

This has major ramifications for forecast quality. In balanced models, the quasi-geostrophic shallow water equation being the simplest example, the Rossby wave dispersion relation is different when the PV gradient is concentrated into a sharp step [Esler, 2004], as opposed to the more familiar theory for uniform planetary vorticity gradient,  $\beta$  [e.g., Vallis, 2006]. In the longwave limit,  $kL_R \ll 1$ , the dispersion relation is the same as that for uniform  $\beta$  when the magnitude of the PV step,  $\Delta q = 2\beta L_R$  (here  $L_R$  is the Rossby deformation radius and  $k$  is the zonal wave number). As  $k$  increases (wavelength decreases), the westward propagation of Rossby waves counter to the eastward flow weakens more rapidly in the  $\beta$ -plane case ( $\sim k^{-2}$ ) than on a PV step ( $\sim k^{-1}$ ), and therefore, the phase speed is greater on smooth PV gradient than a step, especially for shortwaves. Therefore, we hypothesize that the weaker gradients in the forecasts will result in overdispersion of Rossby waves. This could preclude stationary behavior and result in a greater rate of downstream development. Weaker tropopause PV gradients would also lead to reduced baroclinic growth rates (as illustrated by the Eady model [e.g., Vallis, 2006]). Dawson et al. [2012] have presented evidence that higher resolution than that used in these ensemble forecasts is required to capture the observed patterns of variability in the North Atlantic sector and their low frequency behavior. This could plausibly be related to the sharpness of PV gradients or representation of diabatic processes.

The decay of total ridge amplitude with lead time is consistent with an underestimation of the diabatic enhancement of PV anomalies. Chagnon et al. [2013] have shown how these arise through a combination of diabatic transport of lower tropospheric air through warm conveyor belts into locations near the tropopause in ridges, plus the effects of longwave cooling. Forster and Wirth [2000] have argued that the sharp drop in water vapor at the tropopause dominates the cooling signal. Since the PV gradient at the tropopause is too weak, it is likely that the humidity gradient is also too weak (since both are nearly conserved tracers) and therefore that the spike in radiative cooling just below the tropopause would also be underestimated. Leroy and Rodwell [2013] have recently shown that upper tropospheric specific humidity in analyses is uncertain by as much as 20%, associated with uncertainty in parameterized mixing in the model and the bias correction required in the assimilation of satellite observations. In their example they repeated the data assimilation cycling using a model where vertical diffusion was reduced in regions of higher static stability, again highlighting the importance of mixing at tropopause level. Therefore, several physical processes could be contributing to this forecast error, in addition to possible decay in Rossby wave amplitude associated with numerical dissipation. However, causality is not proven here. Further work needs to be done to attribute the systematic errors in Rossby wave structure shown here to model limitations in the representation of diabatic processes.

#### Acknowledgments

Gray, Chagnon, and Methven acknowledge funding from the Natural Environment Research Council (NERC) project "Diabatic influences on mesoscale structures in extratropical storms" (DIAMET) as part of the Storm Risk Mitigation program, NERC grant NE/I005196/1. Dunning thanks NERC for providing funding for a summer placement at the Department of Meteorology, University of Reading. <http://tiggie.ecmwf.int>.

The Editor thanks two anonymous reviewers for their assistance in evaluating this paper.

#### References

- Brennan, M. J., G. M. Lackmann, and K. M. Mahoney (2008), Potential vorticity (PV) thinking in operations: The utility of nonconservation, *Weather Forecast.*, 23, 168–182, doi:10.1175/2007WAF20060441.
- Browning, K. A., and N. M. Roberts (1994), Structure of a frontal cyclone, *Q. J. R. Meteorol. Soc.*, 120, 1535–1557.
- Butchart, N., and E. E. Remsberg (1986), The area of the stratospheric polar vortex as a diagnostic for tracer transport on an isentropic surface, *J. Atmos. Sci.*, 43, 1319–1339, doi:10.1175/1520-0469(1986)043<1319:TAOTSP>2.0.CO;2.
- Chagnon, J., S. L. Gray, and J. Methven (2013), Diabatic processes modifying potential vorticity in a North Atlantic cyclone, *Q. J. R. Meteorol. Soc.*, 139, 1270–1282, doi:10.1002/qj.2037.
- Davies, H., and M. Didone (2013), Diagnosis and dynamics of forecast error growth, *Mon. Weather Rev.*, 141, 2483–2501, doi:10.1175/MWR-D-12-00242.1.
- Dawson, A., T. N. Palmer, and S. Corti (2012), Simulating regime structures in weather and climate prediction models, *Geophys. Res. Lett.*, 39, L21805, doi:10.1029/2012GL053284.
- Dirren, S., M. Didone, and H. C. Davies (2003), Diagnosis of "forecast-analysis" differences of a weather prediction system, *Geophys. Res. Lett.*, 30(20), 2060, doi:10.1029/2003GL017986.
- Esler, J. G. (2004), Benjamin-Feir instability of Rossby waves on a jet, *Q. J. R. Meteorol. Soc.*, 130, 1611–1630, doi:10.1256/qj.03.74.
- Forster, C., and V. Wirth (2000), Radiative decay of idealized stratospheric filaments in the troposphere, *J. Geophys. Res.*, 105, 10,169–10,184, doi:10.1029/2000JD900052.
- Frame, T. H. A., M. H. P. Ambaum, S. L. Gray, and J. Methven (2011), Ensemble prediction of transitions of the North Atlantic eddy-driven jet, *Q. J. R. Meteorol. Soc.*, 137, 1288–1297, doi:10.1002/qj.829.
- Jung, T., A. M. Tompkins, and M. J. Rodwell (2005), Some aspects of systematic error in the ECMWF model, *Atmos. Sci. Lett.*, 6, 133–139, doi:10.1002/asl.105.
- Leroy, S. S., and M. J. Rodwell (2013), Leveraging highly accurate data in diagnosing errors in atmospheric models, *Bull. Am. Meteorol. Soc.*, doi:10.1175/BAMS-D-12-00143.1, in press.
- Martínez-Alvarado, O., H. Joos, J. Chagnon, M. Boettcher, S. L. Gray, R. S. Plant, J. Methven, and H. Wernli (2014), The dichotomous structure of the warm conveyor belt, *Q. J. R. Meteorol. Soc.*, doi:10.1002/qj.2276, in press.
- Methven, J. (2009), Diagnostics of the extratropics, *Proceedings of the ECMWF Annual Seminar 2009*, pp. 25–51, ECMWF, Reading, U. K.
- Methven, J. (2013), Wave activity for large-amplitude disturbances described by the primitive equations on the sphere, *J. Atmos. Sci.*, 70, 1616–1630, doi:10.1175/JAS-D-12-0228.1.

- Nakamura, N., and A. Solomon (2011), Finite-amplitude wave activity and mean flow adjustments in the atmospheric general circulation. Part II: Analysis in the isentropic coordinate, *J. Atmos. Sci.*, *66*, 2783–2799, doi:10.1175/2011JAS3685.1.
- Park, Y. Y., R. Buizza, and M. Leutbecher (2008), Preliminary results on comparing and combining ensembles, *Q. J. R. Meteorol. Soc.*, *134*, 2029–2050, doi:10.1002/qj.334.
- Plant, R. S., G. C. Craig, and S. L. Gray (2003), On a three-fold theory of extratropical cyclogenesis, *Q. J. R. Meteorol. Soc.*, *129*, 2989–3012.
- Rodwell, M. J., L. Magnusson, P. Bauer, P. Betchtold, M. Bonavita, C. Cardinali, and M. Diamantakis (2013), Characteristics of occasional poor medium-range weather forecasts for Europe., *Bull. Am. Meteorol. Soc.*, *94*, 1393–1405, doi:10.1175/BAMS-D-12-00099.1.
- 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.
- Vallis, G. K. (2006), *Atmospheric and Oceanic Fluid Dynamics*, 745 pp., Cambridge Univ. Press, Cambridge, U. K.