Manuscript prepared for Nonlin. Processes Geophys.
with version 5.0 of the LaTeX class copernicus.cls.
Date: 28 May 2014

Implications of model error for numerical climate prediction

O. Martínez-Alvarado

Department of Meteorology, University of Reading, Earley Gate, P.O. Box 243, Reading RG6 6BB, UK

Correspondence to: O. Martínez-Alvarado (o.martinezalvarado@reading.ac.uk)

Abstract. Numerical climate models constitute the best available tools to tackle the problem of climate prediction. However, their value is hampered by the presence of biases with respect to present-day climate, which renders the interpretation of results ambiguous. Two assumptions lie at the heart of a climate model suitability: (1) a climate attractor exists, and (2) the numerical climate model's attractor lies on the actual climate attractor or, if the model's dimension is lower than that of the actual climate, on the projection of the climate attractor on the model's phase space. In this contribution, the Lorenz '63 system is used both as a prototype system and as an imperfect model to investigate the implications of the second assumption. The imperfect model is initialised with perfect initial conditions with respect to the prototype system's attractor while the perfect model is initialised with imperfect initial conditions. Thus, the orbits generated by both perfect and imperfect models depart from prototype system's orbits starting from the same initial states. However, it is shown that there are fundamental differences in the behaviour at short lead times of perfect and imperfect model orbits. These differences arise purely from model error and not from sensitivity to initial conditions. Identifying these differences would be useful to evaluate climate models but this is only possible if the perfect model is known, which is obviously an unrealistic requirement. It is also shown that if a model is a perfect model then the attractor, constructed by sampling a sufficiently large collection of initialised model orbits (forecast orbits), will be invariant to forecast lead time. This conclusion is strictly valid only for the case of constant parameters. Nevertheless, in the absence of perfect models, it has potential as a tool not only for understanding model error sources but also for the assessment of models and for understanding model biases. The implications of using imperfect models for the prediction of climate and the potential for these applications are discussed by comparing the results drawn from the Lorenz '63 system and from operational models.

1 Introduction

One of the principal aims of numerical climate models is to provide a reliable tool for the prediction of climate change following pre-defined future forcing scenarios. The suitability of numerical climate models to study the climate is based on two assumptions. The first assumption is the existence of a climate attractor. There is no rigorous proof that this attractor exists (e.g. Lorenz, 1991). However, the observations available on long-term records provide some certainty about the validity of this assumption (e.g. Essex et al., 1987). The second assumption is that the solutions provided by a numerical climate model lie on the actual climate attractor, or at least on the projection of the infinite-dimensional climate attractor on the model's phase space. Only under this assumption, numerical climate model solutions can be used to study the properties of the climate under present-day conditions. In order to study climate change, an additional assumption must be made: it must be assumed that the model climate attractor responds in the same way as the actual climate attractor does under changing forcing conditions. The last assumption has been subject to extensive investigation through studies that have shown that, rather than producing new patterns of variability, the effects of anthropogenic forcing project onto already existing patterns of natural variability (e.g. Palmer, 1993b; Corti et al., 1999).

The presence of biases in climate models with respect to observations and reanalysis datasets under present-day conditions (Randall et al., 2007) indicates that the actual climate attractor (as inferred from observations and reanalysis datasets) and the attractors of available climate models are different even if just slightly, i.e. the second assumption is not fully satisfied. While climate models will never be perfect, we can make use of the limit posed by the second assumption to devise useful measures for the assessment of climate models. The relation between errors in the parameterisation of fast physics processes and

errors in long-term simulations has been investigated through techniques such as "initial tendencies" (see Klocke and Rodwell, 2013, and references therein). However, there is no mathematically rigorous theory to explain the relationship between phenomena developing in short timescales and observed long term trends. Such a theory would also help to relate errors in weather prediction and biases in climate projections. For example, it would help to explain why climate models exhibit biases in the tilt of cyclone tracks (Zappa et al., 2013) even though they are capable to simulate realistic extratropical cyclones (Catto et al., 2010).

The objective of this contribution is to show some implications of the second assumption for long-term integrations of a "simple" dynamical system in a three-dimensional phase space: the Lorenz '63 system (Lorenz, 1963). The Lorenz '63 system has been used as an archetype system in several previous studies of weather and climate (e.g. Palmer, 1993a,b; Mu et al., 2002). Palmer (1993a) used the Lorenz '63 system (including several modified versions) to investigate extended-range predictability of nonlinear systems. Palmer (1993a) also introduces the concept of state-dependent predictability by showing that the predictability of the Lorenz '63 system strongly depends on the initial position on the Lorenz attractor. Palmer (1993b) showed that the effects of climate change are expected to modify already existing patterns of atmospheric variability. Mu et al. (2002) investigates three problems on predictability related to maximum prediction time, maximum prediction error and maximum admissible errors in initial values and parameter values. Unlike those studies, in which the Lorenz '63 system was used to infer properties of the climate or the properties of climate models separately, in this contribution the Lorenz '63 system is used to investigate relationships between a system and an imperfect model (e.g. a model with a similar structure to that of the system but with different parameters). This is similar to the approach taken by Orrell et al. (2001), who used system/model combinations to investigate shadowing of target orbits in low- and high-dimensional systems. Instead of trajectory shadowing, the focus in this study are the cumulative effects of model error for long-term integrations, so that no orbit could be expected to shadow any target trajectory.

To avoid confusion, in this article the term "prototype system" refers to a system as part of a system/model combination. Thus, a prototype system and its model are both dynamical systems and in this work they will be instances of the Lorenz '63 system, differing only on the values of their parameters. The methodology is fully described in Sect. 2 while the results for the Lorenz '63 system are discussed in Sect. 3.

Clearly, there are several important differences between the prototype system/imperfect model combination using the Lorenz '63 system and the combination formed by the climate system and numerical climate models. For example, the Lorenz '63 system is perfectly known whereas our knowledge of the climate system relies on observations which are of limited temporal extent and subject to observational error. Another important difference is that the imperfect model for the Lorenz '63 system (as constructed here) share its dimensionality with the prototype system, whereas numerical weather and climate models have necessarily a lower dimensionality than the climate system. Despite these and other differences, there are several implications that can be transferred between both systems. These implications are discussed in Sect. 4. Finally, a summary and concluding remarks are given in Sect. 5.

2 Methodology

The Lorenz '63 system is defined by the equations (Lorenz, 1963)

$$\dot{x} = \sigma(y - x), \tag{1}$$

$$\dot{y} = rx - y - xz, \tag{2}$$

$$\dot{z} = xy - bz. \tag{3}$$

The variables x, y and z define the phase space of the system while σ, b and r are constant parameters. For a range of these parameters the trajectories of the system tend asymptotically towards the well-known two-winged Lorenz attractor. The shape of the attractor depends on the values given to the parameters σ, b, and r. Thus, two fixed points (defined as the points in phase space for which $\dot{x} = 0$, $\dot{y} = 0$, $\dot{z} = 0$) for $r > 1$ are located at $(x_0, y_0, z_0) = \left(\pm\sqrt{b(r-1)}, \pm\sqrt{b(r-1)}, r-1\right)$. A third fixed point is located at the origin. The three fixed points are unstable for

$$r > r_\mathrm{c} = \sigma\frac{\sigma + b + 3}{\sigma - b - 1}. \tag{4}$$

In this region of the parameter space the system has no other attractors but a strange attractor.

The standard values of these parameters (as used by Lorenz, 1963) are $\sigma = 10$, $b = 8/3$, and $r = 28$. In this work, the Lorenz '63 system characterized by these parameter values will be regarded as the prototype system. The error-doubling times for this system have been determined to be between $0.15\,\mathrm{t.u.}$ and $8\,\mathrm{t.u}$ (Smith et al., 1999). To construct an imperfect model of the prototype system the values of $\sigma = 10$ and $b = 8/3$ will be kept but the value $r = 25$ will be used instead. By doing so, $r_\mathrm{c} = 24.74$ is valid for both the prototype system and the imperfect model. However, the position of the fixed points differs between the prototype system and the imperfect model. For the prototype system $C^\mathrm{S}_{1,2} = (x_{1,2}, y_{1,2}, z_{1,2})^\mathrm{S} = (\pm 8.49, \pm 8.49, 27)$ while for the imperfect model $C^\mathrm{M}_{1,2} = (x_{1,2}, y_{1,2}, z_{1,2})^\mathrm{M} = (\pm 8, \pm 8, 24)$.

The prototype system and the imperfect model have been initialised with the same random initial conditions drawn from a uniform distribution between 0 and 1 for the three phase-space variables. Then, the system and the imperfect

model have been integrated for $10025\,\mathrm{t.u.}$ ($1\,\mathrm{t.u.}$ = 1 time unit) with a sampling rate of $\Delta t = 0.01\,\mathrm{t.u.}$ The first $25\,\mathrm{t.u.}$ have been discarded from both orbits to eliminate initial transients. The remaining $10^4\,\mathrm{t.u.}$ (i.e. 10^6 points) in each integration are considered here as the attractors of the system and the imperfect model. Time zero is then defined as the initial time of the 10^4-t.u. integrations.

The separation between any two points in phase space is measured throughout this work using the Euclidean distance (which is referred to simply as distance). The distance between the given point and the attractor is defined here as the minimum distance between a given point in phase space and the points in an attractor.

3 Model error in the prototype system/imperfect model combination

The attractors of the prototype system and the imperfect model are shown in Fig. 1. Figure 1a shows the attractors as orbits in phase space. For the sake of clarity only the first $500\,\mathrm{t.u.}$ have been plotted. Even at this relatively short time the characteristic features of both attractor have emerged, allowing comparison. The structures of both attractors appear similar, but the size of the imperfect model attractor appears smaller (Fig. 1a). Figure 1a shows the joint probability density functions (PDFs) for (x, z) characterising both attractors. These PDFs have been constructed using the full length of the long integrations. The comparison shows again the similarity in the structure of the attractors and the difference in location and size between them. Thus, the second assumption is not satisfied in this case, i.e. the attractor of the model does not lie on the attractor of the system.

Let us assume that we observe the prototype system at regular intervals, e.g. every $5\,\mathrm{t.u.}$. This observation rate produces 2000 observations from the 10^4-t.u. integrations. Figure 2a shows these observations during the interval $79 < t < 96\,(\mathrm{t.u})$ on the x subspace as black points on top of the prototype system's orbit (black line). Let us attempt to forecast the state of the prototype system using the imperfect model and those observations as initial conditions. Given that the prototype system is perfectly known, the observations are perfectly accurate apart from round-off error. Under these conditions, the forecast will tend to move away from the prototype system attractor towards the imperfect model's attractor due to two separate albeit related effects. First, the accurate initialisation of the imperfect model with respect to the prototype system's orbit moves the forecast trajectory away from the imperfect model's attractor. Figure 2b shows that the distance between observations and the imperfect model's attractor during the interval shown is small but not negligible. A more comprehensive view of this aspect of the imperfect model initialisation is given by the PDF of the distance between observations and the imperfect model's attractor (Fig. 3). The initialisation of the imperfect model from a point on the prototype system's attractor induces a transient period during which the orbit tends towards the imperfect model's attractor. Second, sensitivity to initial conditions will pull the imperfect model orbit away from the prototype system orbit and into a different segment on the imperfect model's attractor. This effect is particularly evident when the initial conditions are close to the imperfect model attractor, as occurs at $t = 85\,\mathrm{t.u.}$ in Fig. 2, for example.

Figure 2a also shows the orbits of the imperfect model for the forecast cycles during the interval $79 < t < 96\,[\mathrm{t.u}]$ projected onto the x subspace (red lines). These orbits appear to closely follow the orbit of the system very well for about $1\,\mathrm{t.u.}$ immediately after each initialisation. However, computing the distance between the prototype system's orbit and the imperfect model forecast orbit, $d = d(t_L)$, reveals that the forecast error is actually much larger than the apparent distance in x (red line, Fig. 2c). For comparison, a perfect model, given by a Lorenz '63 system with the same parameter values as the prototype system, was also used to forecast the state of the prototype system. The perfect model was initialised with imperfect initial conditions given by the observations randomly perturbed assuming that the observational error in each variable x, y, and z is independent and normally distributed with standard deviation $\sigma_\mathrm{O} = 0.1\,\mathrm{l.u.}$ ($1\,\mathrm{l.u.}$ = 1 length unit in phase space). The perfect model forecast orbits projected onto the x subspace are also shown in Fig. 2a (grey lines). Even though in this case we should expect the orbits to diverge from the prototype system's orbit due to sensitivity to initial conditions, the divergence appears to be much slower than in the imperfect model case. This effect becomes clearer when looking at the distance between the prototype system's orbit and the perfect model forecast orbits (black lines in Fig. 2c): the perfect model forecast orbits remain close to the prototype system's orbit for about $2\,\mathrm{t.u.}$ immediately after each initialisation. After this initial interval, sensitivity to initial conditions takes over and the perfect model forecast orbits move away from the prototype system's orbit.

In order to show that these results are robust, similar analyses were conducted for imperfect models with $r = 26$ and $r = 27$ and for initial conditions with $\sigma_\mathrm{O} = 0.2\,\mathrm{l.u.}$, $\sigma_\mathrm{O} = 0.5\,\mathrm{l.u.}$ and $\sigma_\mathrm{O} = 1\,\mathrm{l.u.}$ Figure 4a shows the evolution of the PDFs, represented by median and interquartile range, of the distance between the prototype system's orbit and the forecasts obtained with these models. For long lead times (i.e. $t_\mathrm{L} = 5\,\mathrm{t.u.}$) the effects of a relatively large observational error (e.g. $\sigma_\mathrm{O} = 1\,\mathrm{l.u.}$) and a relatively small model error (e.g. $r = 27$) are apparently similar (see Fig. 4a). At shorter lead times, however, there are important behaviour differences between models. The two imperfect models show a short period of very fast divergence from the prototype system's orbit followed by a plateau and a second period of fast divergence. It is hypothesized that the first period of fast divergence is induced by the approach of the imperfect model

forecast orbit to the imperfect model's attractor. Following Nicolis et al. (2009), the mean of the square distance between the prototype system's orbit and the forecasts are given for the Lorenz '63 system by

$$\langle d^2(t_L)\rangle = t_L^2 \delta r \langle \bar{x}^2 \rangle \quad (5)$$

where $\delta r = r_\mathrm{s} - r_\mathrm{mod}$ is model error (given as error in model parameters), r_s is the value of the parameter r in the prototype system, r_mod is the value of the parameter r in the model, $\bar{x}$ is the initial value of x on the prototype system's attractor and $\langle\cdot\rangle$ indicate average over the ensemble of forecasts. This expression is strictly valid only in the limit of small model error and short forecast lead times (Nicolis et al., 2009). However, it still provides a good approximation for the cases under analysis here, characterised by relatively large model error. Figure 5a shows a comparison between the obtained curves following Eq. (5) and those obtained from the ensemble of forecasts. The agreement is indeed very good, but only for very short forecast lead times ($t_L < 0.1\,\mathrm{t.l.}$). Due to the nonlinearity of the underlying dynamics mean and median do not follow each other as the PDFs rapidly become highly asymmetric as forecast lead time increases (see also Fig. 8 of Nicolis et al., 2009).

In contrast to the forecasts produced by imperfect models, the forecasts produced by a perfect model with imperfect initial conditions show periods of slow divergence from the prototype system's for short lead times. In fact, Fig. 4b, which shows the rate of change of the median of the distance between the prototype system's orbit and model orbits with respect to forecast lead time, reveals that the perfect model runs undergo a short period during which the distance between model orbit and prototype system's orbit tends to decrease. Indeed, this shrinking period occurs as a consequence of the prototype system's orbit being part of the prototype system's attractor and having initial conditions with finite observational error. This is consistent with the theory of Nicolis et al. (2009) which shows that, assuming unbiased and uncorrelated initial condition errors, the square distance between the prototype system's orbit and the forecasts of a perfect model will reach a minimum at a certain time $t_L^* > 0$. Considering terms up to $O(t_L^2)$, t_L^* is given by (Nicolis et al., 2009)

$$t_L^* = -\frac{\langle \mathrm{Tr}\mathbf{J}\rangle}{\langle \mathrm{Tr}[(\mathbf{J}^T+\mathbf{J})\mathbf{J}]\rangle}, \quad (6)$$

where $\mathbf{J}$ is the Jacobian matrix of the prototype system. Equation (6) is strictly valid for short forecast lead times and small initial condition errors. However, it still provides a reasonably accurate estimate for the time of minimum $\langle d^2\rangle$ in the cases considered here: Eq. (6) predicts $t^*_\mathrm{Lt} = 0.024$; the values found for $\sigma_\mathrm{O} = \{0.1, 0.2, 0.5, 1.0\}\,(\mathrm{l.u.})$ are $t^*_\mathrm{Le} = \{3,4,4,3\}\times 10^{-2}\,(\mathrm{t.u.})$, respectively. Notice that while $\langle d^2\rangle$ attains minimum values at these times, $\mathrm{median}(d)$ shrinks slightly further to attain minimum values at $t^*_\mathrm{Lmedian} = \{7,7,7,8\}\times 10^{-2}\,(\mathrm{t.u.})$, respectively.

Figure 4b provides a summary of the difference between imperfect models with perfect initial conditions and perfect models with imperfect initial conditions. At the beginning of the forecast cycle, the imperfect models are characterised by a positive and comparatively large rate of change in the median of the distance between the orbits of the prototype system and the model with respect to forecast lead time; on the other hand, the perfect models are characterised by a negative and comparatively small rate of change of the same variable.

There are two drawbacks to this analysis that makes it difficult to apply to real situations. The first such drawback is that it can lead to erroneous conclusions about the difference between perfect and imperfect models. For example, one might argue, by pointing at the gray and red lines in Fig. 4a, that having small model error ($r = 27$) or large observational error ($\sigma_\mathrm{O} = 1\,\mathrm{l.u.}$) leads to very similar model behaviour. Following this line of thought, one could try to eliminate the initial period of fast divergence in the imperfect model by following a suitable strategy to project "unbalanced" initial conditions onto the surface on which the attractor of the imperfect model evolves (e.g, the strategy suggested by Anderson, 1995). However, focusing on the distance between orbits alone gives only a partial view of the situation: two points could be at a similar distance from a third point, and nevertheless be placed at very dissimilar locations. The second but most important drawback is that the analysis relies on a perfect knowledge of the prototype system, which is, as indicated before, an unattainable requirement.

Figure 6 highlights a different aspect of the comparison between perfect and imperfect models: the location of the attractor in phase space. This aspect is fundamental for climate prediction, in which we are not interested in predicting the state of a system at a particular time, but in the statistical properties of the system during a time interval of a given duration at a particular starting time. Figure 6 shows the evolution of the PDF of z, represented by median and interquartile range and computed using forecasts, as forecast lead time increases. For comparison, it also shows these same quantities computed using the attractors of the prototype system and the imperfect model shown in Fig. 1. At $t_\mathrm{L} = 0\,\mathrm{t.u.}$, both perfect and imperfect models produce very similar statistics to those produced by the prototype system. The small difference between statistics at $t_\mathrm{L} = 0\,\mathrm{t.u.}$ is due to the difference in sample size between the prototype system's attractor and the forecasts. As forecast lead time increases the differences between perfect and imperfect model become more apparent. The imperfect model forecast orbits tend to the imperfect model's attractor so that in less than about $0.5\,\mathrm{t.u.}$ the statistics produced by the imperfect model for $t_\mathrm{L} > 0.5\,\mathrm{t.u.}$ are closer to those produced by the imperfect model's attractor than to those produced by the prototype system's attractor. Moreover, the PDF appears to oscillate around that of the imperfect model's attractor. These features are also observed with imperfect models

with $r_{\mathrm{mod}} = 26$ and $r_{\mathrm{mod}} = 27$ (not shown). In contrast, the perfect model forecast orbits tend towards the prototype system's attractor. Therefore, the statistics produced by the perfect model forecasts remain around those produced by the prototype system throughout the whole forecasting cycle even though this model was initialised with imperfect initial conditions.

It has been pointed out earlier that looking into a single variable in a multidimensional system can be misleading. This is unavoidable for many systems due to their large dimensionality (i.e. the Earth's climate system). This is not the case here, where we are dealing with only three phase space variables. To get a more comprehensive overview of the situation we use the Hellinger distance, which can be interpreted as a measure of the similarity between two PDFs and whose square is defined as (Arnold et al., 2013)

$$H^2(f,g) = \frac{1}{2}\int_V \left(\sqrt{f(\boldsymbol{X})} - \sqrt{g(\boldsymbol{X})}\right)^2 dX_1 \ldots dX_n, \quad (7)$$

where f and g are two PDFs defined over the n-dimensional random variable space $\boldsymbol{X} = (X_1, \ldots, X_n)^T$. Figure 7 shows the Hellinger distance between the prototype system PDF, computed from the 10^4-t.u. integration, and the PDFs obtained from the perfect model forecasts at different forecast lead times. Each curve in the figure corresponds to a different level of observational error. Every curve shows fluctuations. However, these fluctuation are around a similar value ($H \simeq 0.11$). The approach to the prototype system's attractor, corresponding to the shrinking periods discussed with Fig. 4, is evident in the case with larger observational error ($\sigma_{\mathrm{O}} = 1.0$ l.u.).

The situation is different in the case of imperfect models. Figure 8 shows the Hellinger distance and the median of the distance between the prototype system's orbit and model orbits for the imperfect model with $r_{\mathrm{mod}} = 25$. In this case the Hellinger distance go through a period of fast growth until a maximum is reached (Fig. 8a). The time of the maximum approximately coincides with the time of the minimum in the mean of z in Fig. 6. Furthermore, it coincides with the time when the plateau in the median of the distance between the prototype system's orbit and model orbits begins (Fig. 8b). This supports the hypothesis put forward earlier that the first period of fast divergence between prototype system and model orbits is induced by the approach of imperfect model forecast orbits to the imperfect model's attractor. A series of alternating minima and maxima follows, but it is more difficult to establish a link between these and the behaviour of the distance between the prototype system's orbit and model orbits.

The imperfect model with $r_{\mathrm{mod}} = 27$ exhibits similar qualitative behaviour to that found in the imperfect model with $r_{\mathrm{mod}} = 25$ (Fig. 9). Notice that in both cases the Hellinger distance does not start at zero but at $H \simeq 0.105$, i.e. at a similar level to that around which the Hellinger distance oscillates in the perfect model case.

These results show a fundamental property of a perfect model: if a model is a perfect model, then the attractor, reconstructed by sampling a collection of initialised model orbits (forecast orbits), will be invariant to forecast lead time, provided two conditions: (1) that the model is initialised with good estimates of the system's true state based on observations and (2) that the collection of forecast orbits is a representative sample of the region in phase space accessible to the system. Condition (1) is required to ensure that the attractor described by the collection of initial conditions is an accurate representation of the prototype system's attractor. Otherwise, an initial period of adjustment should be expected. Klocke and Rodwell (2013) refer to this period as 'initialisation shock' in their discussion on the initialisation of climate models in hindcast mode from non-native analyses. Condition (2) is required to ensure that forecasts provide a full description of the prototype system's attractor in phase space.

The perfect model's attractor invariance with forecast lead time marks a clear difference between perfect and imperfect models and provides an alternative means not only for understanding model error sources, but also for climate model evaluation. Potentially, it also provides a method to understand and interpret climate model biases. It has the advantage over the distance between the orbits of the prototype system and the models that no prior knowledge of the prototype system's orbit is required (apart from initial states at suitable times). Moreover, it avoids the false impression that a perfect model and an imperfect model exhibit similar behaviour.

4 Implications for climate prediction

4.1 Attractor reconstruction

Reconstructing even part of the attractor of a system is equivalent to knowing at least part of its climate. It would be desirable to reconstruct the full climate attractor in order to completely know the climate. However, this task is impossible given the very large dimensionality of the climate system. In principle, it would be enough to collect a sufficiently large number of observations to be able to represent the system's attractor in phase space and infer its properties. However, if the only source of data available was the imperfect model, then the most we could achieve would be to represent the imperfect model's attractor in model phase space. This is related to the existence of biases in climate models when evaluated against observations and reanalysis datasets (e.g. Kim et al., 2009; Matsueda et al., 2009; Zappa et al., 2013). As discussed in Sect. 1, these biases are an expression of the mismatch between the climate attractor and the attractors

of climate models. Evidence of the existence of biases can be found even using short-term forecasts by comparing two different times in a forecast cycle, analysis time (i.e. T + 0 d) and T + 15 d, in an analogous way to that used to study the prototype system/imperfect model combination based on the Lorenz '63 system in Sect. 3. Figure 10 shows interquartile ranges of daily zonally-averaged 320-K potential vorticity (PV) for the period between December 2009 and February 2010 for these two lead times and for three forecast datasets produced with three different models: (1) the Met Office Global and Regional Ensemble Prediction System (MOGREPS, Bowler et al., 2008), (2) the European Centre for Medium-Range Weather Forecasts (ECMWF) Ensemble Prediction System (EPS, Molteni et al., 1996) and (3) the National Centers for Environmental Prediction (NCEP) Global Ensemble Forecast System (GEFS, Toth and Kalnay, 1997). These datasets have been archived by the THORPEX Interactive Grand Global Ensemble (TIGGE, Park et al., 2008). Even though the source of these datasets are operational ensemble prediction systems. only control members are considered here, i.e. the analysis is effectively made on a set of deterministic forecasts.

As shown in Sect. 3, if the models were perfect, the statistics between the forecasts at T + 0 d and T + 15 d would be similar or, in the limit of infinitely large samples, the same. However, the three datasets reveal clear statistical differences between analyses and T + 15 d forecasts. It must be noted that even though the three ensemble prediction systems produce different statistics at analysis time and at T + 15 d, the deviation shown by the ECMWF EPS (Fig. 10b) seems systematically smaller than that produced by MOGREPS (Fig. 10a) or NCEP GEFS (Fig. 10c). This effect might occur as a result of the optimisation of the ECMWF model for the specific purpose of medium-range weather prediction. However, this is only one metric and more research would be needed to give a complete comparison between these and other TIGGE models.

There are two potential caveats in these results. The first potential caveat is that the results are shown only for the season December–February (DJF) 2009–2010 in the Northern Hemisphere, which was characterized by exceptional conditions in terms of atmospheric circulation in the North-Atlantic European sector (e.g. Santos et al., 2013). However, five other DJF periods have been analysed (from 2006–2007 to 2011–2012) on both hemispheres and all of them show the same qualitative results. Moreover, Gray et al. (2014) have shown, analysing the same dataset, that the amplitude of Rossby waves and the sharpness of the PV gradient across the tropopause tend to decrease with increasing forecast lead time (see their Fig. 5). These finding confirm the existence of systematic model error in the upper-level Rossby wave structure in these models and its growth with forecast lead time.

The second potential caveat is that only the control members (unperturbed analyses with no stochastic physics included in the forecast model) in each ensemble prediction systems have been considered in this analysis. However, the ensemble members tend to follow the behaviour of the control member. For example, Fig. 11a shows the 2-PVU contours ($1\,\mathrm{PVU} = 10^{-6}\,\mathrm{m^2\,K\,kg^{-1}s^{-1}}$) on the 320-K isentropic surface in the analysis and the control members in five forecasts produced with MOGREPS for the same validation time (00:00 UTC, 25 November 2009) but different lead times (T + 1 d to T + 5 d). Figure 11b shows the 2-PVU contours on the 320-K isentropic surface in the analysis and the ensemble members for the T + 4 d forecast for the same validation time. There are two remarks to make regarding this figure. The first remark is that the apices of the upper-level ridge (over Scandinavia in the analysis) in the control members tend towards the southeast with increasing lead time (Fig. 11a) (Sideri, 2013). The second remark is that the ensemble at the lead time shown (Fig. 11b), and in fact any other between 1 d and 5 d, clusters around the corresponding control member while failing to include the analysis (Sideri, 2013).

4.2 Short-term forecast

It has been shown that initialising the imperfect model with perfect initial conditions with respect to the system can be viewed as initialising the model with initial conditions away from its own attractor. This induces a transient period during which the model approaches its own attractor. Data assimilation blends information from the model and observations in order to provide initial conditions for the next forecast. Using data assimilation to initialise a numerical prediction model has a similar effect to initialising the imperfect model with perfect initial conditions by moving the initial model state away from the model's attractor. This induces a transient (spin-up) period until the numerical model reaches a new balance (Daley, 1991). The new balance is achieved when the model's orbit is close to the model's attractor.

The transient period and the subsequent evolution on the model attractor imply divergence between the model's orbit and the true system's orbit. This divergence is not only due to sensitivity to initial conditions. Instead, it is partly due to fundamental differences between the system and the imperfect model. The forecast of the upper-level ridge on 25 November 2009 introduced in Sect. 4.1 provides one example of this model-error related divergence (Fig. 11). As mentioned before, the apices in the forecasts tend towards the southeast as lead time increases (Fig. 11a), thus indicating that the model is diverging from the system's orbit. The fact that no member in the ensemble is close to the actual behaviour of the system (Fig. 11b) might be due to the same effect: in this particular event, an ensemble around accurate initial conditions generates an ensemble forecast with every member tending towards the model's attractor and away from the true future state of the system. This

occurs even though MOGREPS incorporates a representation of model error variability in the ensemble (Bowler et al., 2008). Furthermore, increasing the representation of model error variability to increase forecast spread could make the analysis fall into the envelope of the forecast ensemble members but it would not fix the fact that the attractors of true system and model are essentially different. There are many other examples of this type of behaviour in other models (e.g. Rodwell et al., 2013). One might argue that even though analysis and forecasts diverge from each other, they could still be part of the same attractor. However, the evidence presented (Fig. 10 and subsequent discussion) strongly suggests that indeed model and system have different attractors.

5 Summary and concluding remarks

It has been shown that, in the prototype system/imperfect model combination based on the Lorenz '63 system, imperfections in the model translated into differences in attractor structure (fixed points and apparent size) between the system and the imperfect model (Fig. 1). As a result, the second assumption for the suitability of a model (i.e. the assumption that the solutions provided by a model lie on the system's attractor, or at least on the projection of the system's attractor on the model's phase space) was not satisfied. Under these circumstances, even a perfectly accurate initialisation of the system induces a transient period during which the model orbit diverges from the system's orbit and approaches the model attractor (Fig. 2). Thus, the orbit of the model and the actual system's orbit become essentially different. This difference is purely due to model error and not to sensitivity to initial conditions. This was shown through a comparison of two imperfect models initialised with perfect initial conditions and a perfect model initialised with imperfect initial conditions subject to four levels of observational error (Fig. 4). It was shown that, even though at long lead times small model error and large observational error produced apparently similar results (Fig. 4a), there were noticeable differences at very short lead times: while imperfect model forecast orbits tend to quickly diverge from the prototype system's orbit, perfect model forecast orbits tend to undergo a short period at the beginning of the forecast cycle during which they approach the prototype system's orbit (Fig. 4b). However, these methods require the prior knowledge of the actual state of the system and its evolution, which is an unaffordable luxury for climate scientists, who are bound to deal with a system of very large dimensionality.

It has been shown that climate model biases can be interpreted as an expression of a mismatch between the climate system attractor and the numerical climate model attractor. Furthermore, it has been shown that such a mismatch can be detected even in short-term forecasts by relying on the following fundamental property of a perfect model: if a model is a perfect model, then the attractor, reconstructed by sampling a collection of initialised model orbits (forecast orbits), will be invariant to forecast lead time, provided two conditions: (1) that the model is initialised with good estimates of the system's true state based on observations and (2) that the collection of forecast orbits is a representative sample of the region in phase space accessible to the system. Deviations from this condition would constitute an alternative measure for the suitability of a climate model. This was shown for the Lorenz '63 system (Fig. 6) and for the control members of three operational ensemble prediction systems (Fig. 10). Stochastic parameterisation schemes have been shown to reduce model biases (Arnold et al., 2013). Therefore, it would be very interesting to extend the analysis to investigate the effects of the full ensembles on the stochastic model's attractor. However, this remains as future work.

These results provide the basis for a framework for the interpretation of output from numerical climate models with implications for two widely recognized needs in climate science: (1) the need for climate model improvement (e.g. Stevens and Bony, 2013) and (2) the need for new methods for the interpretation of current available climate models when contrasted against observations (e.g. Brands et al., 2012). There are projects tackling the first need although these focus primarily on errors arising in the short term in order to minimise the interaction between parameterisations (e.g. Klocke and Rodwell, 2013, and references therein). This is an important aspect for the improvement of models. However, it is equally important to understand such parameterisation interactions and therefore it would be valuable to extend that research to these situations. The approach proposed here provides a link between the fields of weather and climate prediction as it relies on the availability of forecast orbits produced by climate models.

The results presented in this contribution are consistent with the discussions by Judd and Smith (2001, 2004). They have shown that, given a set of imperfect observations in a perfect models scenario, it is possible to find a set of indistinguishable states consistent with the observations (Judd and Smith, 2001). In contrast, in an imperfect model scenario, almost certainly no trajectory of the imperfect model is consistent with any set of observations (Judd and Smith, 2004). Judd and Smith (2004) also introduce the concept of pseudo-orbits that intrinsically take into account the existence of model error. Although the discussion in Judd and Smith (2001, 2004) do require the availability of observations, the concept of pseudo-orbits might prove useful for the interpretation of climate projections; however, I can only speculate at this point, leaving this for future investigation.

Acknowledgements. The author thanks Jeffrey Chagnon, Triantafyllia Sideri, Sue Gray, John Methven, Ross Bannister and

Joaquim Pinto for insightful discussions. The author was funded by the Natural Environment Research Council (NERC) as part of the DIAMET project (NE/I005196/1) during the development of this work. The author also thanks Dr Hannah Arnold and an anonymous referee for the time invested in reviewing this article and for useful comments.

References

Anderson, J. L.: Selection of initial conditions for ensemble forecasts in a simple perfect model framework, J. Atmos. Sci., 53, 22–36, 1995.

Arnold, H. M., Moroz, I. M., and Palmer, T. N.: Stochastic parametrizations and model uncertainty in the Lorenz '96 system, Philos. T. R. Soc. A, 371, 2013.

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

Brands, S., Herrera, S., Fernandez, J., and Gutierrez, J.: How well do CMIP5 Earth System Models simulate present climate conditions, Clim. Dynam., 41, 803–817, 2012.

Catto, J. L., Shaffrey, L. C., and Hodges, K. L.: Can climate models capture the structure of extratropical cyclones?, J. Climate, 23, 1621–1635, 2010.

Corti, S., Molteni, F., and Palmer, T.: Signature of recent climate change in frequencies of natural atmospheric circulation regimes, Nature, 398, 799–802, 1999.

Daley, R.: Atmospheric data analysis, Cambridge University Press, UK, 1991.

Essex, C., Lookman, T., and Nerenberg, M.: The climate attractor over short timescales, Nature, 326, 64–66, 1987.

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

Judd, K. and Smith, L.: Indistinguishable states I: Perfect model scenario, Physica D, 151, 125–141, 2001.

Judd, K. and Smith, L. A.: Indistinguishable states II: The imperfect model scenario, Physica D, 196, 224–242, 2004.

Kim, D., Sperber, K., Stern, W., Waliser, D., Kang, I.-S., Maloney, E., Wang, W., Weickmann, K., Benedict, J., Khairoutdinov, M., Lee, M.-I., Neale, R., Suarez, M., Thayer-Calder, K., and Zhang, G.: Application of MJO simulation diagnostics to climate models, J. Climate, 22, 6413–6436, 2009.

Klocke, D. and Rodwell, M.: A comparison of two numerical weather prediction methods for diagnosing fast-physics errors in climate models, Q. J. Roy. Meteorl. Soc., online first, 2013.

Lorenz, E. N.: Deterministic non-periodic flow, J. Atmos. Sci., 20, 130–141, 1963.

Lorenz, E. N.: Dimension of weather and climate attractors, Nature, 353, 241–244, 1991.

Matsueda, M., Mizuta, R., and Kusunoki, S.: Future change in wintertime atmospheric blocking simulated using a 20-km-mesh atmospheric global circulation model, J. Geophys. Res., 114, D12114, 2009.

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

Mu, M., Wansuo, D., and Jiacheng, W.: The predictability problems in numerical weather and climate prediction, Adv. Atmos. Sci., 19, 191–204, 2002.

Nicolis, C., Perdigao, R. A. P., and Vannitsem, S.: Dynamics of prediction errors under the combined effect of initial condition and model errors, J. Atmos. Sci., 66, 766–778, 2009.

Orrell, D., Smith, L., Barkmeijer, J., and Palmer, T. N.: Model error in weather forecasting, Nonlin. Processes Geophys., 8, 357–371, 2001.

Palmer, T.: Extended-range atmospheric prediction and the Lorenz model, B. Am. Meteorol. Soc., 74, 49–65, 1993a.

Palmer, T.: A nonlinear dynamical perspective on climate change, Weather, 48, 314–326, 1993b.

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

Randall, D. A., Wood, R. A., Bony, S., Colman, R., Fichefet, T., Fyfe, J., Kattsov, V., Pitman, A., Shukla, J., Srinivasan, J., Stouffer, R. J., Sumi, A., Taylor, K. E., and contributing authors: Climate models and their evaluation, in: Climate change 2007: The physical science basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, edited by Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K. B. M., Tignor, M., and Miller, H. L., Cambridge University Press, Cambridge, UK, 2007.

Rodwell, M. J., Magnusson, L., Bauer, P., Bechtold, P., Bonavita, M., Cardinali, C., Diamantakis, M., Earnshaw, P., Garcia-Mendez, A., Isaksen, L., Källén, E., Klocke, D., Lopez, P., McNally, T., Persson, A., Prates, F., and Wedi, N.: Characteristics of occasional poor medium-range weather forecasts for Europe, B. Am. Meteorol. Soc., 94, 1393–1405, 2013.

Santos, J. A., Woollings, T., and Pinto, J. G.: Are the winters 2010 and 2012 archetypes exhibiting extreme opposite behavior of the North Atlantic jet stream?, Mon. Weather Rev., 141, 3626–3640, 2013.

Sideri, T.: Do clouds and precipitation in a warm conveyor belt modify the structure of a downstream ridge?, Master's thesis, University of Reading, UK, 2013.

Smith, L. A., Ziehmann, C., and Fraedrich, K.: Uncertainty dynamics and predictability in chaotic systems, Q. J. Roy. Meteor. Soc., 125, 2855–2886, 1999.

Stevens, B. and Bony, S.: What Are Climate Models Missing?, Science, 340, 1053–1053, 2013.

Toth, Z. and Kalnay, E.: Ensemble forecasting at NCEP and the breeding method, Mon. Weather Rev., 125, 3297–3319, 1997.

Zappa, G., Shaffrey, L. C., and Hodges, K. I.: The ability of CMIP5 models to simulate North Atlantic extratropical cyclones, J. Climate, 26, 5379–5396, 2013.

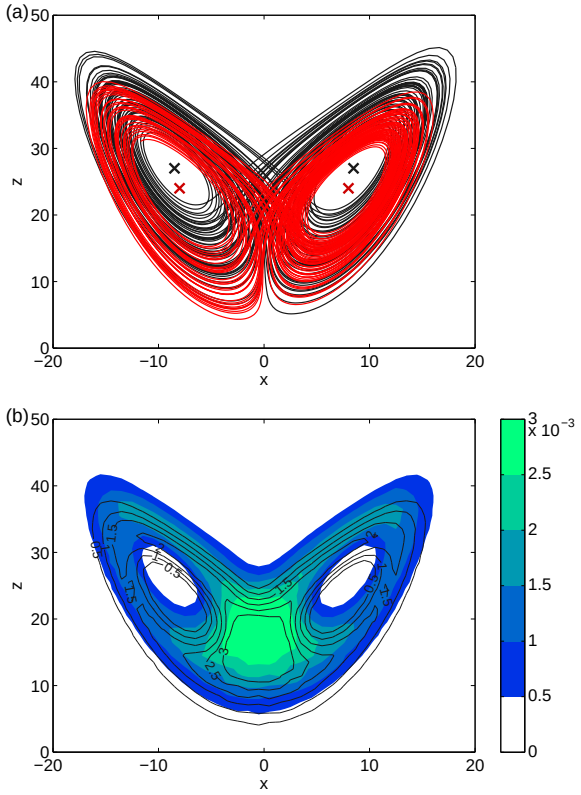

Fig. 1. (a) Prototype system attractor (black) and imperfect model attractor (red). Also shown are the fixed points for the system (black x) and for the imperfect model (red x). **(b)** Joint probability density function for (x,z) resulting from long integrations (10^4 t.u.) of the prototype system (colour shading) and the imperfect model (line contours). The contour values are the same in both cases.

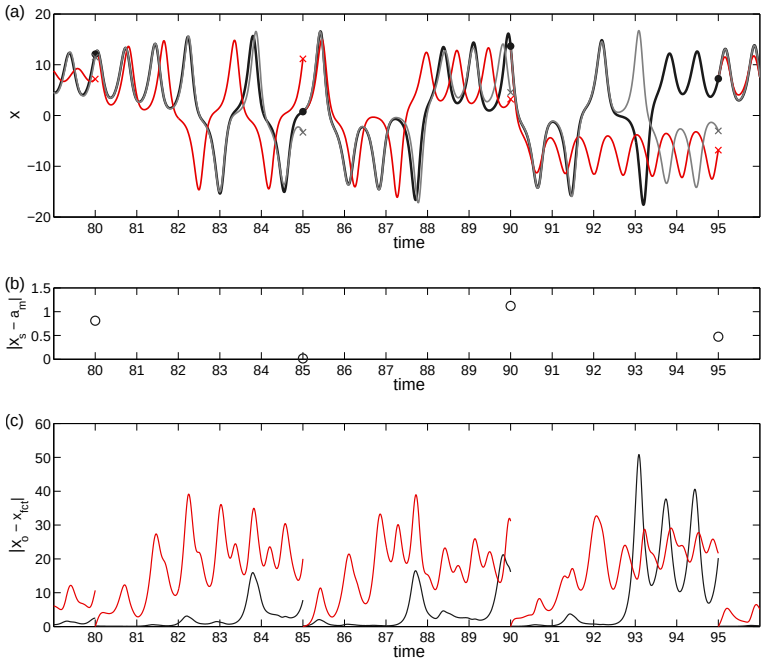

Fig. 2. (a) Prototype system orbit (black) and perfect observations on the prototype system orbit (black dots). Also shown are (1) orbits of the imperfect model with perfect initial conditions (red lines) and corresponding forecasts (red crosses), and (2) orbits of the perfect model with imperfect initial conditions (grey lines) and corresponding forecasts (black crosses). Observations and forecasts are shown at regular intervals of 5 t.u. **(b)** Distance between perfect initial conditions (black dots in (a)) and the imperfect model attractor at observation times during the interval shown. **(c)** Error between the prototype system orbit and those from the imperfect model with perfect initial conditions (red) and the perfect model with imperfect initial conditions (black).

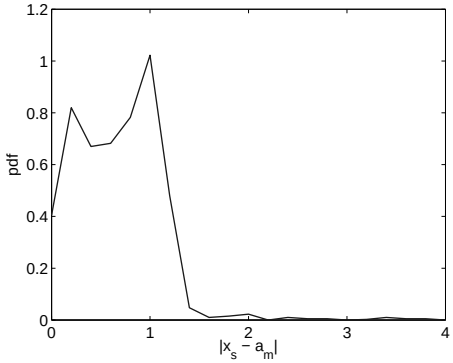

Fig. 3. Probability density function of the distance between perfect initial conditions and the imperfect model attractor.

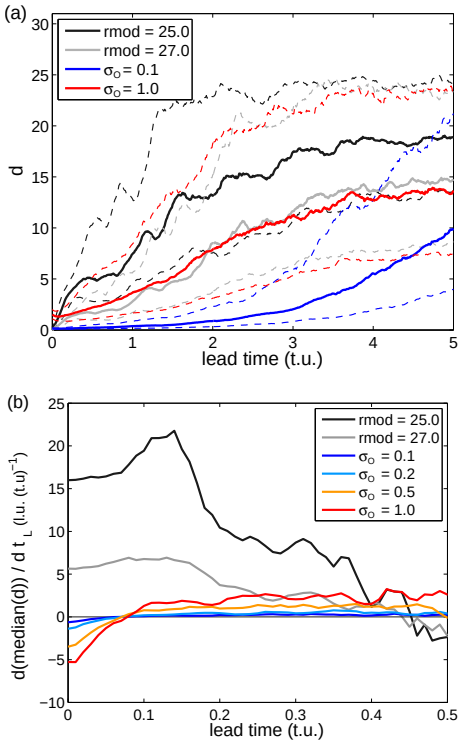

Fig. 4. (a) Median (solid lines) and interquartile range (delimited by dashed lines) of the distance between the prototype system's orbit and model orbits as functions of forecast lead time. **(b)** Rate of change of the median of the distance between the prototype system's orbit and model orbits with respect to forecast lead time.

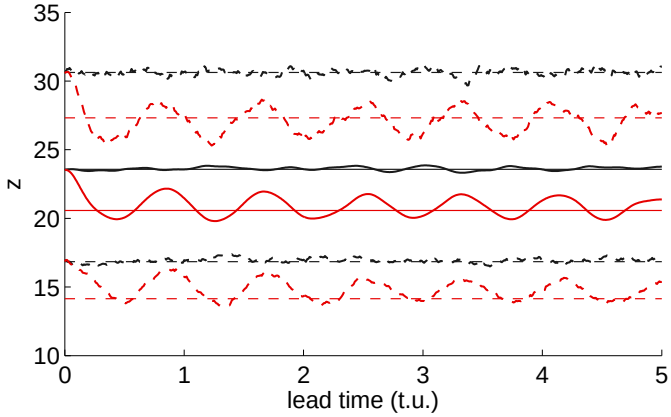

Fig. 6. Mean (solid lines) and interquartile range (delimited by dashed lines) of z as a function of forecast lead time for the imperfect model with $r = 25$ and perfect initial conditions (red lines) and for a perfect model with imperfect initial conditions ($\sigma_{\mathrm{O}} = 0.1\,\mathrm{l.u.}$, black lines). The horizontal lines represent the mean and interquartile range of z computed from attractors of the prototype system (black) and the imperfect model (red) shown in Fig. 1.

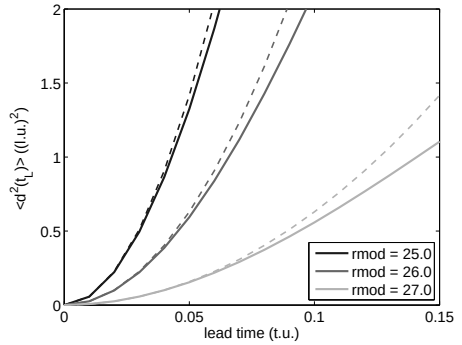

Fig. 5. Average over the ensemble of forecasts of the square distance between the prototype system's orbit and model orbits for $0 \leq t_L \leq 0.15\,(\mathrm{t.u.})$ (solid lines) for three different values of r_{mod}. Also shown are the corresponding theoretical curves according to Eq. (5) (dashed lines).

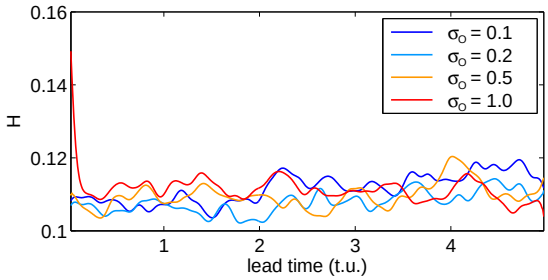

Fig. 7. Hellinger distance between the prototype system's PDF and the perfect model's PDFs with different levels of error in the initial conditions. The lines have been smoothed by applying a 5-point-average filter 14 times.

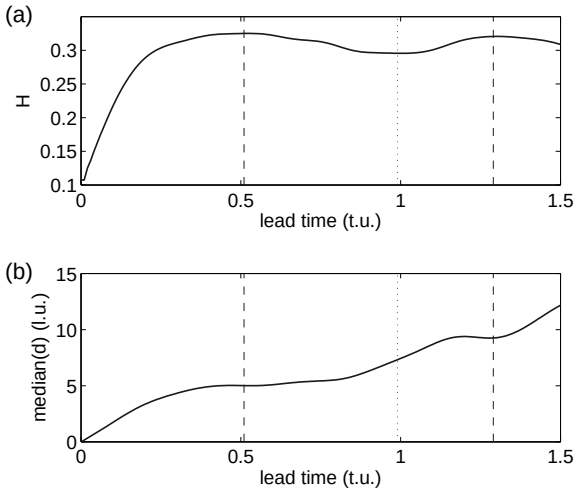

Fig. 8. (a) Hellinger distance H between the prototype system's PDF and the imperfect model's PDFs with $r_{\mathrm{mod}} = 25$. **(b)** Median of the distance between the prototype system's orbit and model orbits. In both panels, vertical dashed (dotted) lines mark the position of maxima (minima) in H. The lines have been smoothed by applying a 5-point-average filter 14 times.

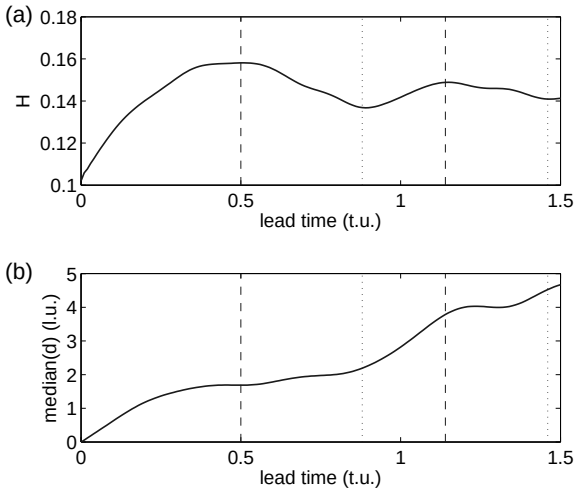

Fig. 9. As in Fig. 8, but for an imperfect model with $r_{\mathrm{mod}} = 27$. Notice that the vertical scales are different to those in Fig. 8.

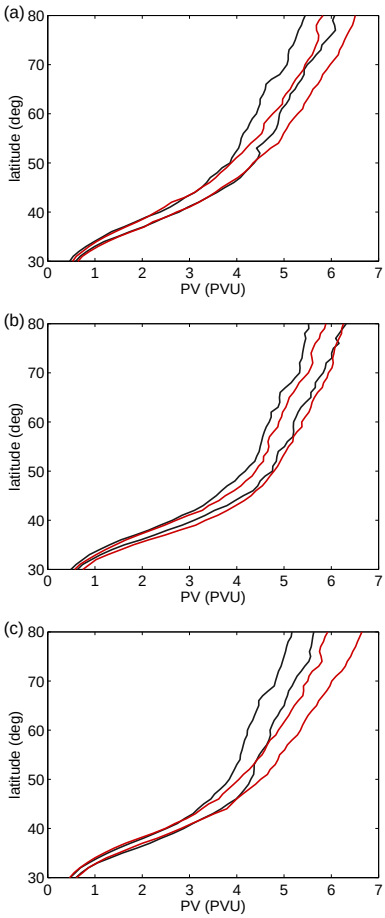

Fig. 10. First and third quartiles of daily zonally averaged 320-K PV in analyses (T + 0, black) and T + 15 d forecasts (red) for **(a)** MOGREPS-15, **(b)** ECMWF EPS, **(c)** NCEP GEFS for the season from December 2009 to February 2010.

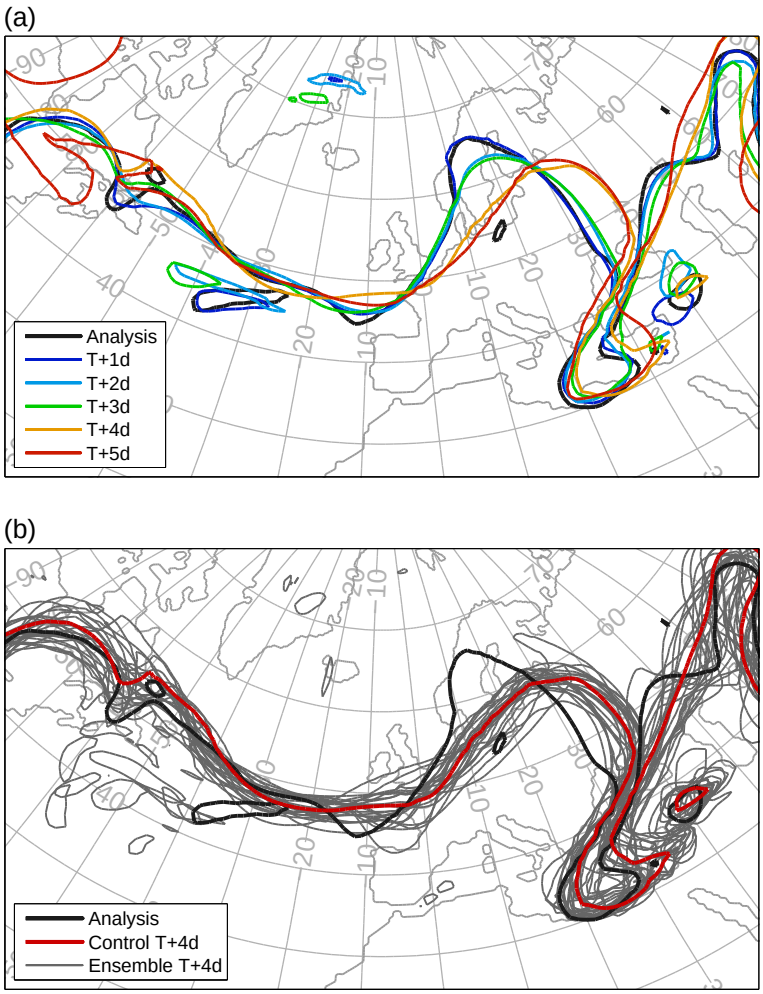

Fig. 11. 2-PVU contours on the 320-K isentropic surface valid at 00:00 UTC 25 November 2009 in MOGREPS showing **(a)** control members for forecasts with different lead times and **(b)** analysis, control member and ensemble members for the T + 4 d forecast (After Sideri, 2013).