Manuscript prepared for Nonlin. Processes Geophys.
with version 5.0 of the LaTeX class copernicus.cls.
Date: 26 September 2014

Multiple scale error growth in a convection-resolving model

F. Uboldi[1] and A. Trevisan[2]

[1]Independent researcher, Milano, Italy
[2]CNR-ISAC, Bologna, Italy

Correspondence to: F. Uboldi (uboldi@magritte.it)

Abstract. The properties of the multiple scale instabilities present in a non-hydrostatic forecast model are investigated. The model simulates intense convection episodes occurring in Northern Italy. A breeding technique is used to construct ensembles of perturbations of the model trajectories aimed to represent the instabilities that are responsible for error growth at various time and space scales. It is found that for initial errors of the order of present-day analysis error, a non-negligible fraction of the forecast error can be explained by a bred vector ensemble of reasonable size representing growth of errors at intermediate scales. In contrast, when the initial error is much smaller, the spectrum of bred vectors representing the fast convective-scale instabilities becomes flat and the number of ensemble members needed to explain even a small fraction of the forecast error becomes extremely large. The conclusion is that as the analysis error is decreased, it becomes more and more computationally demanding to construct an ensemble that can describe the high-dimensional subspace of convective instabilities and that can thus be potentially useful for controlling the error growth.

1 Introduction

In the last 10–15 years there has been operational interest for non-hydrostatic, convection-resolving models (e. g. Dixon et al., 2009; Seity et al., 2011; Baldauf et al., 2011) and for the possibility of making use of data assimilation (e. g. Zhang et al., 2003; Kain et al., 2010; Schenkman et al., 2011; Claussnitzer et al., 2011) to improve short range forecasting and nowcasting of weather fields, or the accuracy of a past event trajectory reconstruction. Some attention has been devoted to evaluate whether a higher resolution and a more realistic description of convective events are in fact associated to a quantitative improvement in forecasting their location, timing and intensity (e. g. Weisman et al., 2008; Schwartz et al., 2009).

Prediction and data assimilation systems for the ocean and the atmosphere are largely funded upon the theory, developed over the last decades, of predictability and state estimation in chaotic dynamical systems. Atmospheric NWP (Numerical Weather Prediction) models that predict the synoptic scale evolution of midlatitude weather systems are based on the hydrostatic assumption and include parameterisation of convective processes. Ensemble forecasting and data assimilation for such models have reached a mature stage, whereas non-hydrostatic models face the hard task of dealing with the fast instabilities typical of convection. Predictability in the presence of multiple scales has been investigated in several studies starting with the pioneering work of Lorenz (1969). Errors in the small-scale weather features such as thunderstorms grow very rapidly and reach saturation within a timescale comparable to their brief lifetime, while errors in larger scales may still be in a linear stage of growth. A comparison of the time scales involved in prediction of hydrostatic versus non-hydrostatic models has been performed by Hohenegger and Schär (2007a), shedding some doubts on the feasibility of borrowing algorithms developed in the context of the former for the application to the latter models.

In the present study we will try to look more deeply into this question, using as a "laboratory" a realistic model (MOLOCH, Sect. 3) and applying to a real case study the classical breeding method. The relation between bred vectors and Lyapunov vectors, developed in a rigorous theoretical framework, states that as the rescaling amplitude goes to zero Lyapunov vectors are recovered from bred vectors (BVs hereafter). This argument can be exploited when there fundamentally is one typical instability scale, as in synoptic scale models. In a convection-resolving model we expect to find a very large number of positive, competing Lyapunov exponents with very large values, making it unfeasible to compute all or even a reasonable number of their corresponding modes.

The flow-dependent character of forecast uncertainty and its estimation has become an important issue in NWP in the last decades starting from the early works of Kalnay and Dalcher (1987) and Lorenz (1996), among others. In the forecasting practice, before Lorenz clarified the dynamical mechanism of selection of instabilities in his 1996 model (Lorenz, 1996), researchers at the National Meteorological Centre (NMC, now NCEP) were well aware of this problem and the breeding method (Toth and Kalnay, 1993, 1997) was devised for its solution. When instabilities in a wide range of scales is present, breeding provides a useful tool for selecting the instability scale one wants to investigate by tuning the breeding parameters, i. e. the rescaling amplitude and time interval. For instance, if the rescaling amplitude is larger than the typical saturation amplitude of a given scale, unstable structures characteristic of that scale will not appear in the bred vectors.

While the importance for prediction of a rigorous description of flow instabilities in terms of the Lyapunov vectors with positive exponents that define the unstable subspace has long been recognised, the introduction of these concepts and the use of Lyapunov vectors in the state estimation problem and data assimilation algorithms is more recent. Starting with the paper by Trevisan and Uboldi (2004), a number of studies (Uboldi and Trevisan, 2006; Carrassi et al., 2007, 2008a,b;

Palatella et al., 2013) have shown that, by confining the Assimilation of observations in the Unstable Subspace (AUS) of the system, one can construct data assimilation algorithms that are computationally more efficient (EKF-AUS, Trevisan and Palatella, 2011b) and in some cases more accurate (4DVar-AUS, Trevisan et al., 2010) than their classical counterparts: the Extended Kalman Filter (EKF) and 4-Dimensional Variational assimilation (4DVar). A review on the role of instabilities in weather forecasting and data assimilation is found in Trevisan and Palatella (2011a).

For geophysical models, the huge number of degrees of freedom constitutes the main difficulty whenever a description of the error probability distribution is attempted, in particular in the application of data-assimilation algorithms. During the forecast stage, errors grow in the unstable subspace, which can be of much lower dimension than the full phase-space dimension of the model: the algorithms that perform the assimilation in the unstable subspace take full advantage of the reduced dimension in which they operate. In high resolution non-hydrostatic models, the presence of a large number of fast instabilities and of spatial scales as small as convective cells makes the problem of high dimensionality particularly critical.

One of the goals of the present paper is to investigate, in a realistic high-resolution, convection-resolving model, what is the structure of the forecast error in relation to the unstable modes at various scales that can be estimated by breeding. One question we try to address is the following: in line of principle, can an assimilation algorithm keep under control a reasonable fraction of the errors that grow on such a wide spectrum of space and time scales?

2 Preliminary remarks

Errors intrinsic to the model dynamics are present in operational systems and constitute an important limitation for all improvements that can be obtained by assimilating observations. Moreover, limited-area models are forced by lateral boundary conditions, that also represent an error source. The goal of this work is, however, to characterise the error growth at various dynamical scales. This is obtained by studying the evolution and growth of errors that are only introduced at initial time: both model and boundary conditions are then "perfect", i. e. they are the same in the "true" trajectory and in all, control and perturbed, trajectories. The reason of this choice is that the introduction of model and/or boundary error would obscure the interpretation of results: it is important not to attribute to newly introduced errors any behaviour which is only and specifically due to error dynamical evolution. At the same time we feel the need to caution the reader against too optimistic an interpretation of any results obtained in the absence of model and boundary errors.

A reader not particularly familiar with the perturbation breeding technique, besides the original papers (Toth and Kalnay, 1993, 1997) may refer to the book by Kalnay (2003), where the relation between bred and Lyapunov vectors is discussed. See also Trevisan and Pancotti (1998); Legras and Vautard (1996); Wolfe and Samelson (2007). A classical method to compute Lyapunov exponents

and vectors is to make use of frequent Gram-Schmidt re-orthogonalisation of small perturbations (Benettin et al., 1980). The review by Trevisan and Palatella (2011a) discusses the role of unstable Lyapunov vectors in atmospheric predictability and data assimilation. Finally, the perturbative equations relevant for breeding in both an unforced dynamical system and a in system forced by the assimilation of observations are presented in Section 2.1 of the paper by Uboldi and Trevisan (2006).

Although throughout this work we broadly associate fast growth to small scales and slow growth to large scales, this should not be intended as a very rigid spatial characterisation of error and perturbation fields, especially in the nonlinear regime of growth.

3 Model and case study

The model is MOLOCH, developed at CNR-ISAC, Bologna, Italy (Buzzi et al., 2004; Zampieri et al., 2005; Malguzzi et al., 2006; Davolio et al., 2006, 2009a)[1]. MOLOCH is a convection-resolving, non-hydrostatic atmospheric model, running at about 2.2 km resolution with 50 hybrid vertical levels. State variables include temperature, pressure, three velocity components, specific humidity, concentrations of cloud water and ice and of three precipitation phases: rain, snow, and hail (plus temperature and water content at 5 ground levels). The domain includes the Alps, Northern Italy and Corsica, with portions of Ligurian, Tyrrhenian and Adriatic Seas. The reference trajectory is a simulation of the real case of 26 September 2006 (Davolio et al., 2009b), with initial and boundary conditions from GFS (Global Forecasting System)[2] and BOLAM (Buzzi et al., 1998, 2004; Zampieri et al., 2005; Malguzzi et al., 2006). In this case-study, the circulation at 500 hPa is characterised by a deep trough West of the Alps. Near the surface, a pressure low is located in the Genoa Gulf, and strong winds from South–East over the Adriatic Sea. The Venice area is invested by intense convective precipitation, mostly during the morning. Four main episodes can be identified in the model trajectories used in this work: (1) 00:00–05:00 UTC: scattered convection over the Po Plain and the Alps; (2) 05:00–09:00 UTC: no convection in the Plain, orographic convection only; (3) 09:00–12:00 UTC: organised convection associated to a mesoscale front with intense precipitation, mainly concentrated on the Alps, Eastern Po Plain (Venice area) and the Northern Adriatic Sea; (4) 12:00–18:00 UTC: various convergence areas with intense winds and precipitation progressively extending to the whole domain.

A BOLAM trajectory provides boundary conditions every hour from 12:00 UTC of 25 September 2006 to 18:00 UTC on 26 September 2006. Initial conditions for MOLOCH can be chosen among these BOLAM fields.

[1] http://www.isac.cnr.it/dinamica/projects/forecasts/index.html (web, last access: 24 September 2014b)

[2] http://www.emc.ncep.noaa.gov/index.php?branch=GFS (web, last access: 24 September 2014)

4 Characterising the fastest instabilities: an extension of previous results

The results reported in this Section regard the fastest instabilities, associated to the smallest dynamical scales resolved by the model. Results presented in this Section were obtained by simulations performed in a smaller domain, not including Corsica and the Western Alps (Uboldi, 2010; Uboldi et al., 2010), and the integrations only covered the time interval between 00:00 and 12:00 UTC on 26 September. The time scales that characterise the growth rate and saturation of instabilities associated to convection in the 26 September case are investigated in the line of Hohenegger and Schär (2007a).

First, a perturbation breeding technique has been used: two initial random, independent perturbations, with very small initial amplitude, are added to the nonlinear model reference trajectory at 00:00 UTC. The RMS (Root-Mean-Square) value is $0.05\,\mathrm{m\,s^{-1}}$ for horizontal velocity at model level 5 (about 925 hPa over sea), with each variable scaled by its estimated variability (estimated values, $5\,^{\circ}\mathrm{C}$ for temperature and $5\,\mathrm{m\,s^{-1}}$ for horizontal velocity, are obtained by performing an average over the control trajectory). Perturbed states evolve following the model nonlinear dynamics and perturbations are rescaled to the initial amplitude every $5\,\mathrm{min}$. Bred vectors quickly get organised in spatially coherent structures and, after about $90\,\mathrm{min}$, the two bred vectors (not shown here) show very similar spatial structures, localised in dynamically active areas, where intense winds and convective precipitation are located, differences only appearing in small scale structures. The emergence of a larger scale unstable pattern shows that the linear regime of growth is not disrupted by the strongly non-linear processes present in moist convection.

In order to estimate the linearity time, two perturbed states are obtained by adding, to the state of the reference trajectory, two perturbations with the same structure (an organised bred vector), but with opposite signs. These perturbed states are non-linearly evolved, and the spatial correlation of the perturbations (i. e. the cosine of the angle between the perturbation vectors) is used as a linearity indicator (Hohenegger and Schär, 2007a). The two perturbations, that initially (01:30 UTC) have the same direction (the cosine is -1), progressively loose their alignment as non-linearity becomes important. Conventionally, linearity is considered lost when the spatial correlation exceeds the threshold -0.25 (for random states it would be $+0.5$, Hohenegger and Schär, 2007b). The scalar product used to compute spatial correlation is extended to the whole domain, and to variables T, U and V after scaling each of them with its own variability (estimated as described above). A minor dependence on variable has been found when the spatial correlation is computed for each variable separately: global linearity appears to be lost when it is lost in the horizontal velocity field. Estimates obtained at different times along the same control trajectory, i. e. starting the perturbed trajectories, in two different experiments, at 01:30 and at 06:00 UTC (26 September 2006), indicate that the time required to loose linearity is shorter during the night ($T_{\mathrm{lin}} \simeq 2.5\,\mathrm{h}$), when the main activity is scattered convection, and longer ($T_{\mathrm{lin}} \simeq 4.0\,\mathrm{h}$) when organised convection, forced by a mesoscale front, takes place during the morning.

Recalling that the perturbed states evolve according to the full non-linear dynamics, the growth rate has been estimated by averaging during the linear growth regime, when the logarithm of the amplification factor is approximately a linear function of time (and the growth rate is approximately constant). The estimates of the doubling time are $T_{\mathrm{D}} \simeq 2.5\,\mathrm{h}$, approximately the same value as T_{lin}, during the night, and $T_{\mathrm{D}} \simeq 2.0\,\mathrm{h}$, smaller than T_{lin}, during the morning.

Comparing T_{D} with T_{lin} is important because when T_{lin} is shorter than T_{D}, errors may reach non-linear saturation so quickly that one can hardly see a linear growth phase. In principle, a reduction of the initial error cannot guarantee a reduction of the forecast error after a time when non-linear error saturation has been reached (Hohenegger and Schär, 2007a).

By comparing T_{D} with T_{lin} in the present case, the morning episode of organised convection, though more intense, appears more stable and predictable than the scattered convection episode occurring during the night.

The values obtained here (for this system and case study) for the ratio $T_{\mathrm{lin}}/T_{\mathrm{D}}$ appear a bit more promising than those obtained by Hohenegger and Schär (2007a), at least in the perspective of very short-range forecasting or nowcasting.

5 Experiments

In Sect. 4 we described perturbation growth and saturation properties with regard to the fastest-growth instabilities associated to the smallest dynamical scales. This was obtained by breeding perturbations, with a very high rescaling frequency and a very small rescaling amplitude. Such amplitude is much smaller, in fact, than the order of magnitude presently achievable in the operational practice for analysis error.

With a rescaling amplitude of the same order of magnitude as the analysis error, and an appropriately lower rescaling frequency, the bred vectors do not describe the fast instabilities, which then have sufficient time to saturate between successive rescaling times.

As an example of a saturated small-scale error, the position of a particular, localised and isolated convective feature can be off by some distance in a "true" and in a "control" trajectory. As long as this offset is shorter than the spatial width of the disturbance itself, the error may grow in a linear fashion. When the offset is larger than the disturbance, though, the error is determined by the presence, in the two state fields, of two separate, non-overlapping signals of approximately the same size. As they drift away from each other, the RMS state difference (any variable) does not change steadily, so that this error component does not effectively grow anymore. Saturated small-scale instabilities are not directly responsible for further error growth, but are expected to affect, by nonlinear interaction, larger scales that may still be in a regime of linear growth. These large scales are mainly responsible for error growth at an error level comparable to the analysis error.

Under these premises, we perform an experiment where the error, i. e. the state difference between

a trajectory representing the truth and a control run representing the forecast, has an initial amplitude (norm) compatible with the present-day analysis error. We then examine the ability of bred vectors to capture the forecast error. In order to do so and in view of the different scales that may be involved in the growth of forecast errors, a range of values for rescaling amplitude and frequency is taken into consideration. In a second parallel experiment, an hypothetical analysis error an order of magnitude smaller is considered.

The experimental framework is that of a twin experiment, were a model trajectory is considered to represent the "truth" and is compared with one or more control trajectories. Here the true trajectory is a free MOLOCH evolution initiated by an external model state (from the same BOLAM trajectory providing boundary conditions) at 21:00 UTC on 25 September 2006 and integrated until 18:00 UTC on 26 September. Two different control trajectories have been considered, both computed using the same model and the same boundary conditions as the true trajectory. The first one is obtained by simply initiating a MOLOCH trajectory from the BOLAM state at 18:00 UTC on 25 September (3 h before the initial condition of the true trajectory) and integrating for 24 h, with the same boundary conditions as the "truth". At 00:00 UTC, its (RMS) TUV error (Sect. 5.1) is about 0.3, corresponding to about 0.4 °C for temperature and 1.9 ms^{-1} for horizontal velocity at level 8 (about 1500 hPa over sea), and 0.25 °C, 1.8 ms^{-1} at level 25 (about 500 hPa). This first control trajectory, labelled 18H, is used for breeding vectors meant to estimate instabilities at a larger scale. The second control trajectory has been obtained from the first one, by rescaling by 0.1 all components of its error vector at 21:00 UTC on 25 September in order to have a smaller initial error and again using the same boundary conditions. This second control trajectory, labelled R21, is used to construct bred vectors intended to estimate small-scale, fast instabilities (still slower, though, than those studied in Sect. 4). In any case no comparison is done before 00:00 UTC on 26 September, so that all trajectories have time enough to develop the dynamical scales typical of a MOLOCH non-hydrostatic, convection resolving simulation.

5.1 Norm and scalar product definition

For both control trajectories, 18H and R21, the error vector is orthogonally projected onto orthonormalised bred vectors and the square error component, which is a function of the number of bred vectors used, is expressed as a percentage of the square norm of the whole error vector (Sects. 5.2 and 6.2). In order to orthogonalise the bred vectors and to compute the error orthogonal projection, the definition of a scalar product is needed. Our choice, appearing as the most natural, is the sum of component products extended to the whole three-dimensional domain (actually vertical levels 5–45 only) for the variables temperature and horizontal velocity, each normalised with its own variability (Sect. 4). This is called TUV scalar product hereafter. The norm of a vector is by definition the square root of the scalar product of the vector with itself; the TUV norm is then dimensionless; it can be interpreted as a fraction of the overall state variability. Since the state vector is composed by

many other variables besides T, U, and V (pressure, specific humidity, vertical velocity, concentration of condensed water phases, etc.), this is more properly a seminorm: it is considered sufficient, though, to evaluate errors and state differences, because the chosen variables have the most direct impact on the dynamics. The error size is defined as the TUV norm of the vector difference between control and true state.

5.2 Breeding and orthonormalisation procedure

A breeding scheme is characterised by a rescaling period and amplitude. The rescaling period is set to 30 min for trajectory 18H (larger initial error), and, to estimate faster instabilities, to 15 min for trajectory R21 (smaller initial error). Several different values of the rescaling amplitude have been tested for each trajectory. Boundary conditions are the same for all perturbed and unperturbed integrations.

All perturbations are randomly generated, point by point on the 3-D domain, on variables T, U, and V at 21:00 UTC on 25 September with the same amplitude used for rescaling (see below). Every three hours all bred vectors are Gram–Schmidt orthonormalised with the TUV scalar product; the error, i. e. the vector difference between control and true state, is then projected onto the orthonormal bred vectors. The ratio between the sum of square error projections and the square norm of the whole error is calculated and expressed as percentage: this is an increasing function of the number of bred vectors used (it would be 100 % for a complete orthonormal basis).

In each experiment, 12 bred vectors are computed (this was the maximum possible allowed by computer memory). A second set of 12 bred vectors was computed in some cases, to assess the dependence on increasing subspace dimension of the projected error component.

For trajectory 18H, having fixed the rescaling period to 30 min, five different rescaling amplitudes have been tested, each in a different run: 0.20, 0.30, 0.36, 0.40, and 0.50 (RMS TUV norm values); For R21, with rescaling period 15 min, six rescaling amplitudes have been tested: 0.033, 0.067, 0.100, 0.133, 0.167, and 0.200 (these values are still an order of magnitude larger than those used for the fastest instabilities of Sect. 4). In both cases, the chosen amplitude values span the range of variability of the corresponding control trajectory error, as it can be seen in Fig. 1.

6 Results

6.1 Multiple scale instabilities in the forecast error and in Bred Vectors subspaces

As stated in the introduction, one of the aims of the present work is to extract information about the instabilities at play from the inspection of bred vectors and their associated growth rates. A second related question is whether the forecast error can be described by BVs in a model such as the present one, which includes the fast small scales typical of convection. In the case of a positive answer, the algorithms developed for ensemble weather forecasting and ensemble data assimilation could

be successfully used also in the context of nonhydrostatic models. We start discussing Fig. 1, that shows the forecast error as a function of time for the two control trajectories starting from different initial errors, one, R21, an order of magnitude smaller than the other, 18H. For both trajectories the growth rate is positive in the initial phase, until about 05:00 UTC for trajectory 18H and 09:00 UTC for trajectory R21, approximately corresponding to episode (1) of Sect. 3, and in the last part of the simulation, from 11:00 UTC to the end, when the intense phenomena progressively extend from North–Eastern Italy to almost the whole domain.

We observe that, whereas the ratio of the initial errors in the two control trajectories was initially as large as 10 (in view of the hypothetical assumption that the analysis error could be reduced by an order of magnitude) after 21 h the ratio of the forecast errors has become smaller than 1.5. The error, during the first 10 h in fact grows much faster in the small initial error trajectory (R21), whereas during the second episode the error growth of the two trajectories is very similar. This is because the growth rate is dominated by (small scale) fast instabilities when the initial error is small, but, as forecast error grows in time, these components saturate and slower instabilities associated to larger dynamical scales become dominant[3]. This is the mechanism suggested in the Lorenz (1996) paper, as well as in Toth and Kalnay (1993), that we here discuss relative to a real case study. We now will see how bred vectors can be useful to understand this behaviour and to quantify the characteristics of instabilities at various scales.

Figure 2 shows the "spectrum" of BVs (3-h averaged) growth exponents for the two control trajectories of Fig. 1. Here, for trajectory 18H (large initial error, top panel), BVs are constructed with rescaling period 30 min and TUV amplitude 0.36; for trajectory R21 (small initial error, bottom panel) BVs are constructed with period 15 min and TUV amplitude 0.100. In Fig. 2, different colors refer to different forecast times.

The time variability of the exponents reflects the time variability of the forecast errors: BV growth rates are larger when the forecast error increases, during the two periods of intense convection activity. The number of BVs with positive growth rate at each time provide some indication on how complex the structure of the error is, i. e. how many directions are unstable and possibly useful to describe errors on the various scales present in the forecast error. We observe that in the set of BVs that has been constructed for the control trajectory with larger initial error (18H) and meant to describe larger scale errors present in the forecast, the growth rate generally decreases with BV index. In general, these BVs are more active (have larger growth rates) at the end of the integration period, when slower scales become dominant, and less active at the beginning. There are not many BVs with positive growth rate in this set except during the second episode of error growth.

In contrast, all the BVs of the set for the control trajectory with small initial error (R21) have large growth rates during the entire period, except for a time interval between the error growth episodes;

[3] In fact they dominate the growth of the whole vector norm, even if small-scale signals corresponding to saturated unstable structures may still be present and, locally, some of them may even still be growing.

furthermore the growth rate in this set of BVs does not decrease at all with BV index, indicating that a large number of directions are present in the small scale unstable subspace. In Sect. 6.2 it is shown that this number is much larger than 12.

Figure 3 shows the growth rate of the leading bred vector of each set together with the growth rate of the forecast error shown in Fig. 1. The close correspondence in the time evolution of the growth rate for the forecast error and the associated BV suggests that the BVs embed the same instabilities that are active in the same time interval in the forecast error. Only in the first stages of the trajectory with small initial error (R21) the fast instabilities dominate, whereas, toward the end of the forecast period, when they have saturated, only those associated to larger scales, with smaller growth rates, contribute to the error in both experiments; the values of the growth rates at this stage are all very close to each other.

In Sect. 6.2 we examine in more detail whether the forecast error of the two trajectories can be described in terms of the two sets of BVs.

6.2 Forecast error component in orthonormalised Bred Vectors subspaces

As described in Sect. 4, we performed the orthogonal projection of the forecast error on the subsets of BVs that were constructed for each of the two control trajectories. The breeding parameters were varied and we explored a wide range of values. Results are shown for several values of the amplitude and the appropriate renormalisation frequency.

Figures 4–6 show these results. In all these figures, the top panel refers to BVs constructed with rescaling period 30 min for control trajectory 18H, while the bottom panel refers to BVs constructed with rescaling period 15 min for control trajectory R21. Notice that, here, percentages of the square error are shown, rather than RMS values.

The square norm of the error orthogonal projection onto the 12-bred-vector subset is shown as a function of simulation time in Fig. 4, where different colors correspond to different rescaling amplitudes, and as a function of rescaling amplitude in Fig. 5, where different symbols indicate different simulation times (RMS rescaling amplitude values on the x axis may be compared with the range of values spanned by the error evolution for the two trajectories in Fig. 1).

For the large initial error trajectory 18H, as it is evident in the top panel of Fig. 5, the largest error component at each time (i. e. for each set of symbols) is almost always obtained for the "optimal" rescaling amplitude 0.36. This value corresponds to the average of TUV error values of trajectory 18H at orthonormalisation times.

On the contrary, as it can be seen in the bottom panels of both Figs. 4 and 5, for the small initial error trajectory R21 the maximum error component is obtained as time progresses for an increasing rescaling amplitude. As a consequence, no "optimal value" is found in this case. The maximum error component on bred vectors is found for smaller rescaling amplitudes at the beginning of the simulation, and for larger rescaling amplitudes at later times. In other words, bred vectors corre-

sponding to larger and slower scales progressively become dominant for increasing time. This result is consistent both with the results of Sect. 5.1 and with the theory: small scale unstable structures grow fast in the initial phases and, when they begin to saturate, they trigger growth at larger and slower scales, which become more important later on.

An important result is that when the analysis error is smaller, as in the initial part (03:00 UTC) of R21, because of the many fast instabilities present in the first stages of error growth, only an insignificant fraction of the forecast error is captured by the set of 12 bred modes. This is evident in the bottom panel of Fig. 5.

12 additional bred vectors have been computed for two selected cases[4]: for 18H control trajectory with rescaling amplitudes 0.36 and period 30 min; for R21 control trajectory with 0.100 amplitude and 15 min period. A total of 24 orthonormal bred vectors is then available in these two cases, and the square norm of the error orthogonal projection onto each set is shown as a function of the subspace dimension in Fig. 6. The top panel shows that for trajectory 18H (large initial error), though very different error percentages are reached at different simulation times, increasing the number of bred vectors above 12 does not increase the projected error component in an important way. Assuming that slower growth at very large scales (that could hypothetically be found by further increasing the subspace dimension) is suppressed by boundary forcing, this means that most of the error components are not, in fact, growing when the initial error is so large. Non-growing error components can be composed of localised unstable structures which have already reached growth saturation, and, also, of wider signals, originally present in the large initial error, that are still present in areas that are dynamically rather inactive (the 18H initial state has been in fact constructed from a forecast state of the hydrostatic model BOLAM, Sect. 5). Such larger non-growing signals may appear, in the error field, as locally organised structures, which persist or drift in the domain, accounting for an important portion of the error.

Finally, in the bottom panel of Fig. 6, corresponding to trajectory R21 (small initial error), the error components appear to continue increasing, more or less steadily, with increasing subspace dimension. This means that the number of small-scale fast instabilities is large (larger than 24, at least).

7 Discussion and further developments

In a preliminary study, it was possible to evaluate growth and saturation properties of the fastest instabilities associated to the smallest (convective) dynamical scales (Sect. 4). Their doubling time, characterising linear error growth, and their tangent-linear time, characterising the duration of growth in the linear regime, resulted of comparable magnitude, leaving, on the one hand, some theoretical pos-

[4]Since 12 BV was the maximum allowed by computer memory in one run, it has been necessary to save on disk all states (i. e. all T, U, V components on the whole 3-dimensional grid) of the previous 12-vector set at all orthogonalisation times and read them from disk during the new run to peform the Gram-Schmidt orthogonalisation on the 24 vectors.

sibilities to make use of some characterisation of the linear regime to control error growth. On the other hand, both these characteristic times resulted quite short (about 2 h) with respect to lead-time lengths of interest for weather forecasting, even if, in principle, still of some interest for nowcasting (6–12 h).

A possibility that would remain open in that context would be the implementation of a sequence of dynamically consistent analysis steps, with appropriate frequency, say every hour, with the perspective of attempting control of the assimilation system trajectory.

The forecast error, built in a twin-experiment framework, has been compared with sets of orthogonalised bred vectors, and its orthogonal projection on bred vectors subspaces has been computed (Sect. 6.2).

When the initial error is comparable with the present-day analysis error, the largest forecast error component on a subspace of 12 bred vectors has been found to have a square norm of about 30 % of the whole square error norm. This appears as a remarkable result, considering the huge linear dimension of the state vector, and the limited number of bred vectors used. Another remarkable point is, though, that when the number of orthogonal bred vectors is increased, the error component on the BVs subspace does not grow much further. Slower mesoscale instabilities, not mainly convective, which remain longer in the linear regime of growth, appear then to be controllable using a small number of independent directions. The "unexplained" 70 % of (square) forecast error appears to be distributed on saturated small-scale unstable structures and on non-growing larger-scale signals present in the error field, directly inherited from the initial state: a field of the larger-scale hydrostatic model providing boundary conditions. These last error components could be reduced or eliminated by assimilating observations, but in fact they are re-introduced each time the model is restarted from an external model field.

Starting with a smaller initial error (using a higher rescaling frequency), the "explained" square percentage of forecast error grows very slowly with the number of bred vectors. Fastest growing, small scale instabilities (purely convective), which are also known to saturate quickly, are important when the error is initially small and appear to be active in a very large number, impossible to control within a reasonably small dimension subspace.

7.1 Error evolution after subtraction of the forecast error projection on the bred vectors subspace

In a successive experiment a new state at 03:00 UTC is obtained by taking the vector difference between the 18H control state and its error projection on the 12-dimensional BVs subspace. In this way, the square error of the new state is 30 % smaller than before.

Figure 7 shows the maps of the 18H control state error (top panel) at 03:00 UTC, of its orthogonal projection on the 12-BV subspace (middle panel), and of their difference, the residual error of the new state (bottom panel). Comparing top and bottom panels it is evident that many structures

disappear. The field in the middle panel shows which error components are detected by the 12 BV together. Undetected error components include many very localised signals, but also a few larger structures (for example one located in the Northern Apennines).

The (RMS) error evolution of the new model trajectory restarted from this new state is shown in Fig. 8 together with, for comparison, the error of the 18H control forecast of Fig. 1. The error appears to grow faster than before: the new trajectory error reaches in about three hours the level of the unperturbed control trajectory, even slightly exceeding it before settling on a similar behaviour.

After restart, the RMS error goes in $3\,\mathrm{h}$ from about 0.3 to about 0.4: this corresponds to a doubling time of about $7\,\mathrm{h}$, much longer than that estimated in Sect. 4 for the fastest-growing instabilities, but similar to that of the initial part of control trajectory R21 started from a small initial error (Fig. 1). Fast growth is reactivated after the error reduction obtained by subtracting the error projection on the BVs subspace.

We think that we can summarise all this by drawing the following picture.

Perturbations that grow fast in a non-hydrostatic, convection-resolving system, are associated to small dynamical scales, typical of convection. These fast-growing instabilities may be present in a large number and quickly reach linear growth saturation, and they do this at an error level smaller than the present-day analysis error. An important portion of larger scale instabilities are effectively controlled by boundary forcing. Between these "extremes", there exist some linear growth, so that part of the error can be estimated by breeding (30 % of its square norm in this case) and, in principle, could be eliminated by assimilating observations. However, after subtracting the BV-estimated component, the growth of fast instabilities is reactivated and rapidly brings the error to its previous level in the subsequent free evolution. This means that it would be necessary to assimilate observations very frequently, with an interval of the order of one hour in the forecast-analysis cycle process, in order at least to control the error growth of slower instabilities. Moreover, setting up such a forecast-analysis cycle process might in principle make unnecessary the periodical restart from an external model field (as customary in the operational practice), thus avoiding the periodical re-introduction of errors due to the different nature of the external model solution.

8 Conclusions

In this paper, we investigated the properties of error evolution in the forecast of a real case study with a convection resolving model. We used the model as a laboratory where we assume that a model trajectory represents the truth and we can vary the amplitude and structure of the initial condition error. We produce forecasts that start from two hypothetical analyses, one with an error comparable to and one with an error ten times smaller than the present day analysis error. In this way it was possible to quantify the properties of instabilities with different space and time scales that are representative of the evolution of the forecast error. According to the theory, when the analysis and

the ensuing forecast are so accurate to reproduce the details of the single convective cells, the error growth rate is that typical of this phenomenon, while as the error increases and convective-scale instabilities saturate, the error growth is due to slower instabilities. This result is based on the theory of Lyapunov exponents and vectors that characterise the unstable, neutral and stable subspace of the system. In the real case study at hand, it is impossible to compute the entire set of Lyapunov vectors with positive growth rate that would give us all the information on the instabilities of the system. Therefore we used the breeding technique to quantify the instabilities that are relevant for forecast errors of a given typical amplitude. We tuned the breeding parameters and used different rescaling amplitude and frequency to adapt them to the assumed initial condition error. The results can be summarised as follows. When the error is comparable to the present day analysis error, the unstable subspace, as estimated with the breeding technique, has the following properties: the growth rate decreases with BV index and the number of BVs that are active during the convective episodes is not extremely large. By projecting the forecast error on this subspace we find that most of the forecast error projection is on the leading BVs, and it does not significantly increase when the number of BVs is raised from 12 to 24. Typical values of growth rates for the leading BVs are about 0.05–0.07 h^{-1} corresponding to doubling times of 10–14 h.

When we decrease the analysis error by an order of magnitude, we find a completely different scenario. The spectrum of BVs is flat: there is a large number of BVs with competitive growth rates, three or four times larger than those found in the previous case (large initial error), corresponding to doubling times of 2–6 h. In addition, the projection of the forecast error on the leading 24 BVs is very small and continues to increase very slowly and steadily, suggesting the need of a very large number of BVs to account even for a small fraction of the forecast error. This makes us believe that the unstable subspace at the convective scale really has a very large dimension. The conclusion is that as we attempt to significantly decrease the analysis error (e. g. by improving on the quality of observations and assimilation schemes), the need to increase the number of members in an ensemble forecast or an ensemble-based assimilation scheme may very well turn out to be so computationally demanding as to become an insurmountable obstacle.

The situation appears to be different for errors typical of present day analysis errors, when the instabilities at play are not growing so fast, because the fast growth of errors at the small convective scales is saturated. In this case we have seen that a non-negligible fraction of the forecast error is captured by the BVs.

We performed an experiment which can be interpreted as a preliminary test on the expected performance of an assimilation in the present situation. We assumed a hypothetical assimilation scheme, able to exactly subtract from the forecast error its projection on the BV subspace (once assigned the subspace where the analysis increment is confined, this would be an upper limit to scheme performance and observations quality), and looked at the forecast error of the subsequent free evolution (i. e. after this single "assimilation" step). A relevant error reduction is obtained (in this perfect-

model, perfect-boundary framework): subtracting 30 % of square norm means reducing the RMS "analysis" error to about 84 % of RMS forecast error. However, after restart, again the fast instabilities take over and the forecast error recovers its previous levels in the matter of 3 h. This means that a frequent assimilation, every hour or so, would be required to control the large-scale, slower instabilities, but rapid growth should still be expected after each analysis, due to the action of small-scale fast instabilities. Their large number makes the possibility of controlling them by data assimilation (by means of an hypothetical "multiple-scale" breeding or ensemble) particularly challenging.

Acknowledgements. We wish to thank Silvio Davolio, Piero Malguzzi, Oxana Drofa for their support in using the MOLOCH model and Andrea Buzzi for the partial financial support of this work under Project PROSA of the Italian Space Agency (ASI).

References

National Centers for Environmental Prediction, http://www.emc.ncep.noaa.gov/index.php?branch=GFS, last access: 21 September 2014.

Previsioni meteorologiche GLOBO BOLAM MOLOCH CNR-ISAC, http://www.isac.cnr.it/dinamica/projects/forecasts/index, last access: 21 September 2014.

Baldauf, M., Seifert, A., Förstner, J., Majewski, D., and Raschendorfer, M.: Operational convective-scale numerical weather prediction with the COSMO model: description and sensitivities, Mon. Weather Rev., 139, 3887–3905, 2011.

Benettin, G., Galgani, L., Giorgilli, A., and Strelcyn, J.-M.: Lyapunov characteristic exponents for smooth dynamical systems and for Hamiltonian systems; a method for computing all of them. Part 1: Theory, Meccanica, 15, 9–20, 1980.

Buzzi, A., Tartaglione, N., and Malguzzi, P.: Numerical simulations of the 1994 Piedmont Flood: role of orography and moist processes, Mon. Weather Rev., 126, 2369–2383, 1998.

Buzzi, A., Davolio, S., D'Isidoro, M., and Malguzzi, P.: The impact of resolution and of MAP reanalysis on the simulations of heavy precipitation during MAP cases, Meteorol. Z., 13, 91–97, 2004.

Carrassi, A., Trevisan, A., and Uboldi, F.: Adaptive observations and assimilation in the unstable subspace by breeding on the data-assimilation system, Tellus A, 59, 101–113, 2007.

Carrassi, A., Ghil, M., Trevisan, A., and Uboldi, F.: Data assimilation as a nonlinear dynamical systems problem: stability and convergence of the prediction-assimilation system, Chaos, 18, 023112, 2008a.

Carrassi, A., Trevisan, A., Descamps, L., Talagrand, O., and Uboldi, F.: Controlling instabilities along a 3DVar analysis cycle by assimilating in the unstable subspace: a comparison with the EnKF, Nonlin. Processes Geophys., 15, 503–521, 2008b.

Claussnitzer, A., Schartner, T., Névir, P., Stephan, K., and Cubasch, U.: The data assimilation method "Latent Heat Nudging" assessed with the Dynamic State Index, Meteorol. Z., 20, 65–172, 2011.

Davolio, S., Buzzi, A., and Malguzzi, P.: Orographic influence on deep convection: case study and sensitivity experiments, Meteorol. Z., 15, 215–223, 2006.

Davolio, S., Buzzi, A., and Malguzzi, P.: Orographic triggering of long lived convection in three dimensions, Meteorol. Atmos. Phys., 103, 35–44, 2009a.

Davolio, S., Mastrangelo, D., Miglietta, M. M., Drofa, O., Buzzi, A., and Malguzzi, P.: High resolution simulations of a flash flood near Venice, Nat. Hazards Earth Syst. Sci., 9, 1671–1678, 2009b.

Dixon, M., Li, Z., Lean, H., Roberts, N., and Ballard, S.: Impact of data assimilation on forecasting convection over the United Kingdom using a high-resolution version of the Met Office Unified Model, Mon. Weather Rev., 137, 1562–1584, 2009.

Hohenegger, C. and Schär, C.: Atmospheric predictability at synoptic versus cloud-resolving scales, B. Am. Meteorol. Soc., 88, 1783–1793, 2007a.

Hohenegger, C. and Schär, C.: Predictability and error growth dynamics in cloud-resolving models, J. Atmos. Sci., 64, 4467–4478, 2007b.

Kain, J. S., Xue, M., Coniglio, M. C., Weiss, S. J., Kong, F., Jensen, T. L., Brown, B. G., Gao, J., Brewster, K., Thomas, K. W., Wang, Y., Schwartz, C. S., and Levit, J. J.: Assessing advances in the assimilation of radar data and other mesoscale observations within a collaborative forecasting-research environment, Weather

Forecast., 25, 1510–1521, 2010.

Kalnay, E.: Atmospheric Modeling, Data Assimilation and Predictability, Cambridge University Press, 2003.

Kalnay, E. and Dalcher, A.: Forecasting forecast skill, Mon. Weather Rev., 115, 349–356, 1987.

Legras, B. and Vautard, R.: A guide to Liapunov vectors, in: Proceedings 1995 ECMWF Seminar on Predictability, vol. 1, pp. 143–156, 1996.

Lorenz, E. N.: The predictability of a flow which possesses many scales of motion, Tellus, 21, 289–307, 1969.

Lorenz, E. N.: Predictability: a problem partly solved, in: Proceedings of ECMWF Seminar on Predictability, 1–18, 4–8 September 1995, Reading, UK, 1996.

Malguzzi, P., Grossi, G., Buzzi, A., Ranzi, R., and Buizza, R.: The 1966 "century" flood in Italy: a meteorological and hydrological revisitation, J. Geophys. Res., 111, D24106, 2006.

Palatella, L., Carrassi, A., and Trevisan, A.: Lyapunov vectors and assimilation in the unstable subspace: theory and applications, J. Phys. A Math. Theor., 46, 254020, 2013.

Schenkman, A. D., Xue, M., Shapiro, A., Brewster, K., and Gao, J.: The analysis and prediction of the 8–9 May 2007 Oklahoma tornadic mesoscale convective system by assimilating WSR-88D and CASA radar data using 3DVAR, Mon. Weather Rev., 139, 224–226, 2011.

Schwartz, C. S., Kain, J. S., Xue, S. J. W. M., Bright, D. R., Kong, F., Thomas, K. W., Levit, J. J., and Coniglio, M. C.: Next-day convection-allowing WRF model guidance: a second look at 2-km versus 4-km grid spacing, Mon. Weather Rev., 137, 3351–3372, 2009.

Seity, Y., Brousseau, P., Malardel, S., Hello, G., Bénard, P., Bouttier, F., Lac, C., and Masson, V.: The AROME-France convective-scale operational model, Mon. Weather Rev., 139, 976–991, 2011.

Toth, Z. and Kalnay, E.: Ensemble Forecasting at NMC: the generation of perturbations, B. Am. Meteorol. Soc., 74, 2317–2330, 1993.

Toth, Z. and Kalnay, E.: Ensemble forecasting at NCEP: the Breeding Method, Mon. Weather Rev., 125, 3297–3318, 1997.

Trevisan, A. and Palatella, L.: Chaos and weather forecasting: the role of the unstable subspace in predictability and state estimation problems, Int. J. Bifurcat. Chaos, 21, 3389–3415, 2011a.

Trevisan, A. and Palatella, L.: On the Kalman Filter error covariance collapse into the unstable subspace, Nonlin. Processes Geophys., 18, 243–250, 2011b.

Trevisan, A. and Pancotti, F.: Periodic Orbits, Lyapunov Vectors, and Singular Vectors in the Lorenz System, J. Atmos. Sci., 55, 390–398, 1998.

Trevisan, A. and Uboldi, F.: Assimilation of standard and targeted observations in the unstable subspace of the observation-analysis-forecast cycle system, J. Atmos. Sci., 61, 103–113, 2004.

Trevisan, A., D'Isidoro, M., and Talagrand, O.: Four-dimensional variational assimilation in the unstable subspace and the optimal subspace dimension, Q. J. Roy. Meteor. Soc., 136, 487–496, 2010.

Uboldi, F.: Inspecting the trajectory instabilities of a convection-resolving model, in: Poster presentation at the ECODYC workshop (Exploring COmplex DYnamics in high-dimensional Chaotic systems): from weather forecasting to oceanic flows, 25–29 January 2010, Dresden, Germany, available at: http://www.pks.mpg.de/~ecodyc10/Contributions/Uboldi.pdf, last access: 26 March 2014, 2010.

Uboldi, F. and Trevisan, A.: Detecting unstable structures and controlling error growth by assimilation of standard and adaptive observations in a primitive equation ocean model, Nonlin. Processes Geophys., 13,

67–81, 2006.

Uboldi, F., Trevisan, A., and Davolio, S.: Evolution and growth of perturbations in a convection-resolving model, in: Poster presentation at the HyMeX workshop, (HYdrological cycle in Mediterranean EXperiment), 8–10 June 2010, Bologna, Italy, available at: http://www.hymex.org/public/workshops/4/posters/poster_HPF38.pdf, last access: 26 March 2014, 2010.

Weisman, M. L., Davis, C., Wang, W., Manning, K. W., and Klemp, J. B.: Experiences with 0–36-h explicit convective forecasts with the WRF-ARW model, Weather Forecast., 23, 407–437, 2008.

Wolfe, C. L. and Samelson, R. M.: An efficient method for recovering Lyapunov vectors from singular vectors, Tellus A, 59, 355–366, 2007.

Zampieri, M., Malguzzi, P., and Buzzi, A.: Sensitivity of quantitative precipitation forecasts to boundary layer parameterization: a flash flood case study in the Western Mediterranean, Nat. Hazards Earth Syst. Sci., 5, 603–612, 2005.

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

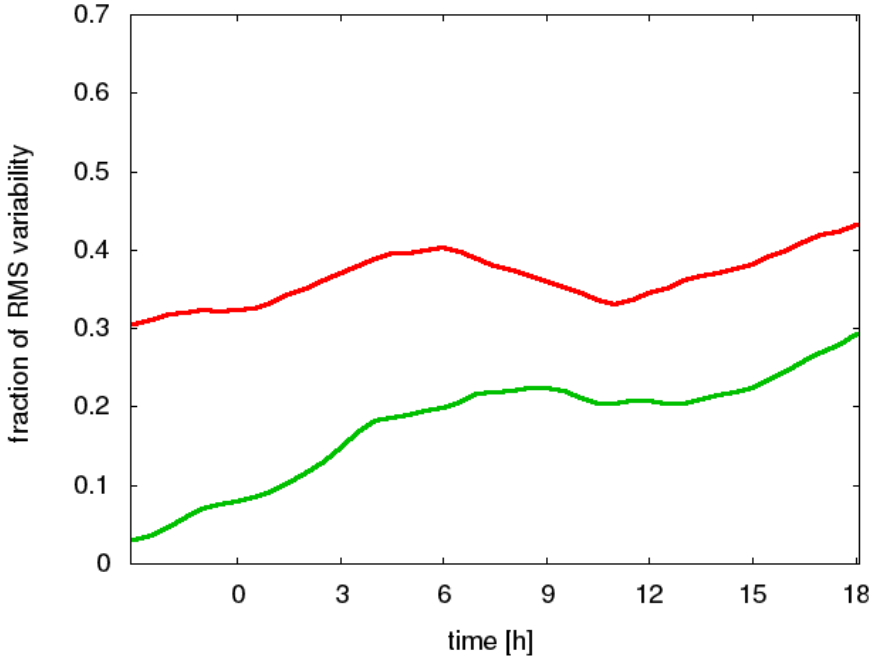

Fig. 1. RMS forecast error of control trajectores with different initial errors. Times indicated on the x axis correspond to UTC times of 26 September 2006. Values on the y axis are nondimensional (TUV norm defined in the text) and are expressed as fractions of natural variability. Red: trajectory labelled 18H (actually starting at 18:00 UTC on 25 September), which at 21:00 UTC has error 0.3, of the order of present-day analysis error. Green: trajectory labelled R21, obtained from the previous one by rescaling, at 21:00 UTC on 25 September, its error to 0.03, i.e. an order of magnitude smaller than present-day analysis error.

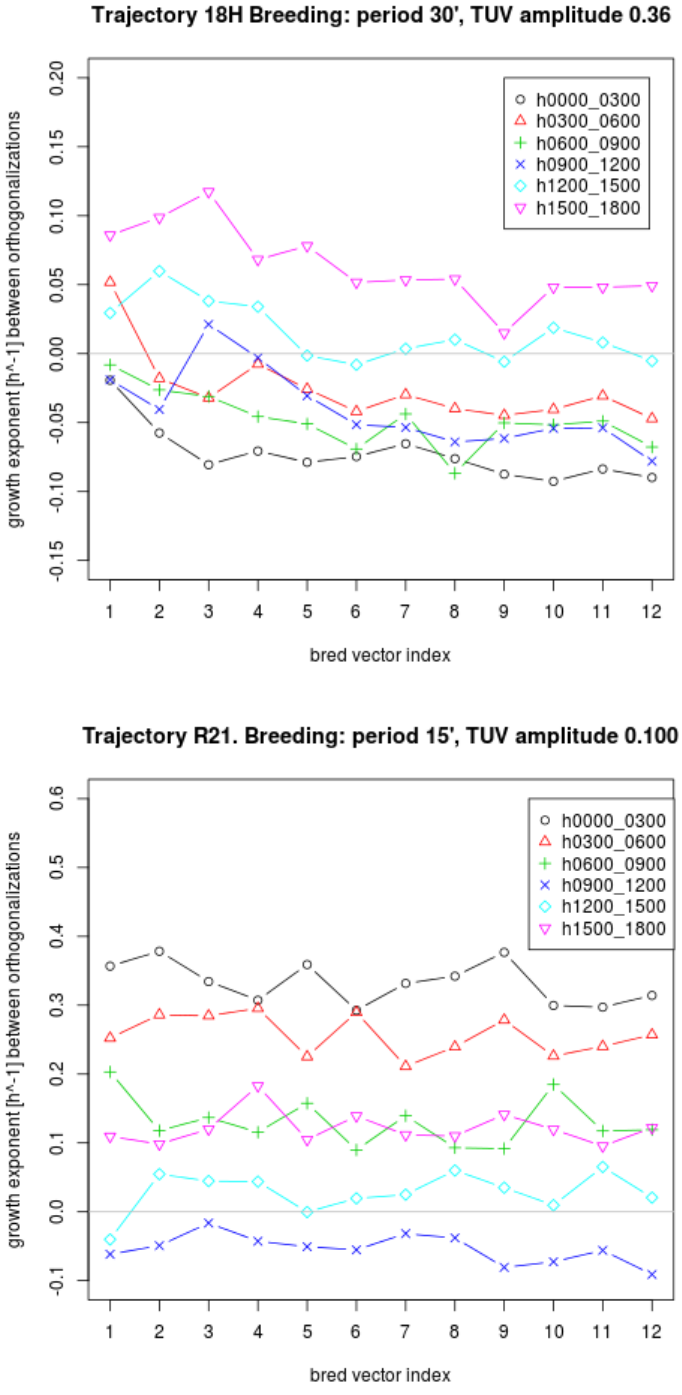

Fig. 2. Growth rates of the 12 Bred vectors at different forecast times, as a function of the bred vector index (the growth rate has been averaged over three hours intervals, as indicated in the legend box). Different colors refer to different forecast times. As indicated in panel titles, the BVs are constructed differently for trajectory 18H (large initial error, top panel) and for trajectory R21 (small initial error, bottom panel).

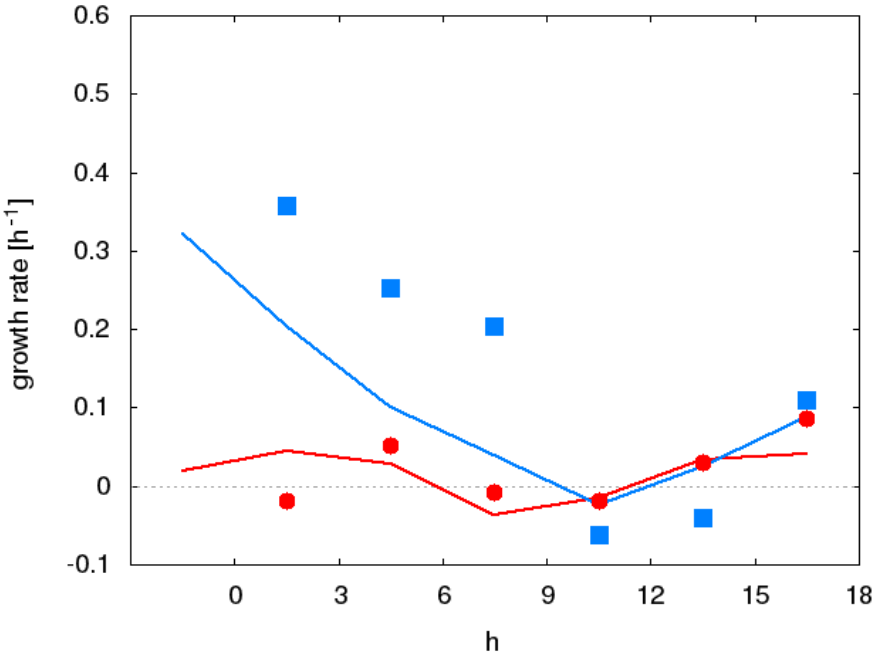

Fig. 3. Growth rates of forecast errors for the two control trajectories of Fig. 1 and of the corresponding leading Bred Vector. Times indicated on the x axis correspond to UTC times on 26 September 2006. Breeding parameters: period 30 min, amplitude 0.36 (red dots) for trajectory 18H (large initial error, red line); period 15 min, amplitude 0.100 (blue squares) for trajectory R21 (small initial error, blue line).

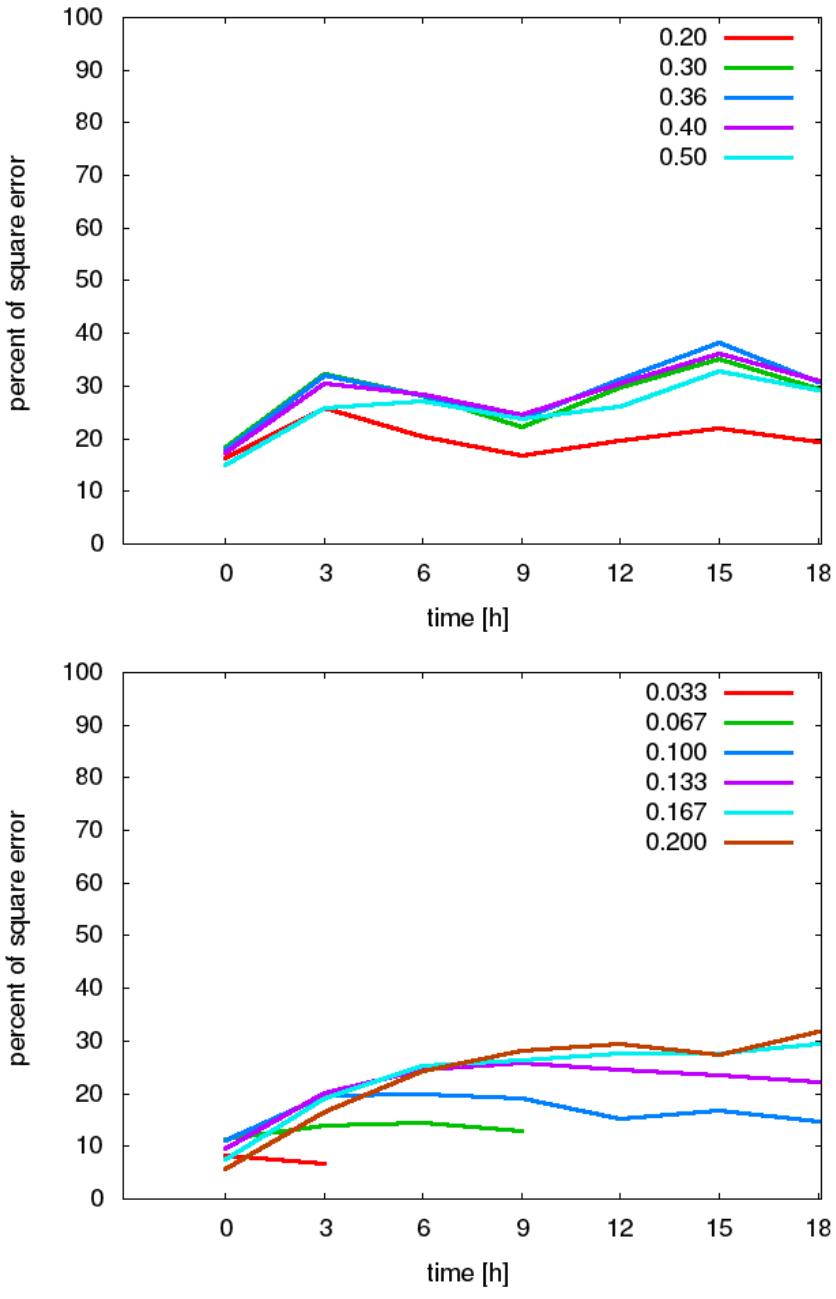

Fig. 4. Square norm of forecast error orthogonal projection onto the subspace of the 12 leading bred vectors (percentage of error square norm) as a function of time. Different curves in each panel correspond to different rescaling norms. Top panel: control trajectory 18H (large initial error), rescaling period 30 min, rescaling RMS TUV norms from 0.20 to 0.50. Bottom panel: control trajectory R21 (small initial error), rescaling period 15 min, rescaling RMS TUV norms from 0.033 to 0.200. Times indicated on the x axis correspond to UTC times on 26 September 2006. The rescaling norm range in each panel should be compared with the corresponding control trajectory error in Fig. 1.

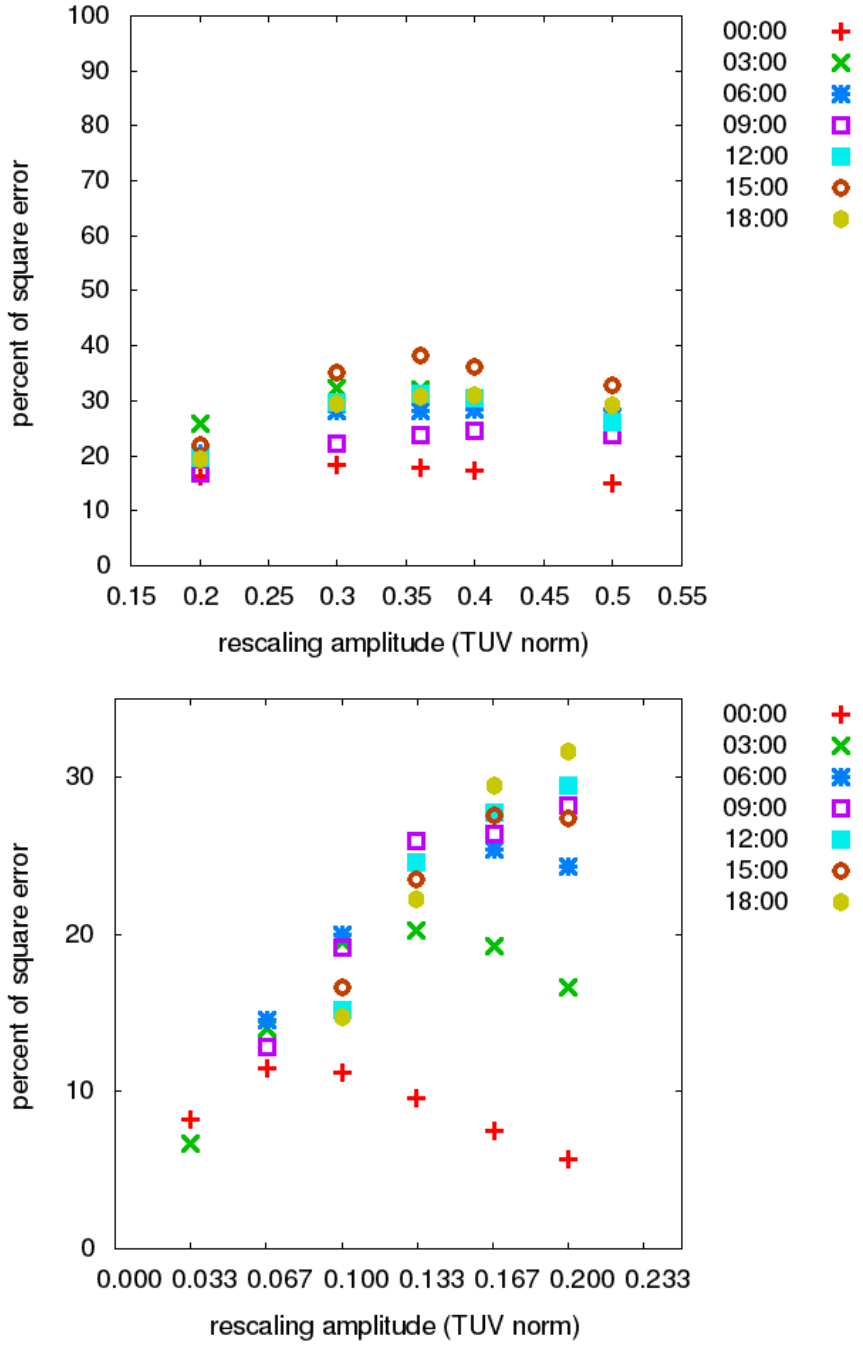

Fig. 5. Square norm of error orthogonal projection onto the subspace of the first 12 bred vectors (percentage of error square norm) as a function of rescaling RMS TUV norm. Different symbols are used in each panel to mark different times (UTC, 26 September 2006). Top panel: control trajectory 18H (large initial error), longer rescaling period 30 min. Bottom panel: control trajectory R21 (small initial error), shorter rescaling period 15 min.

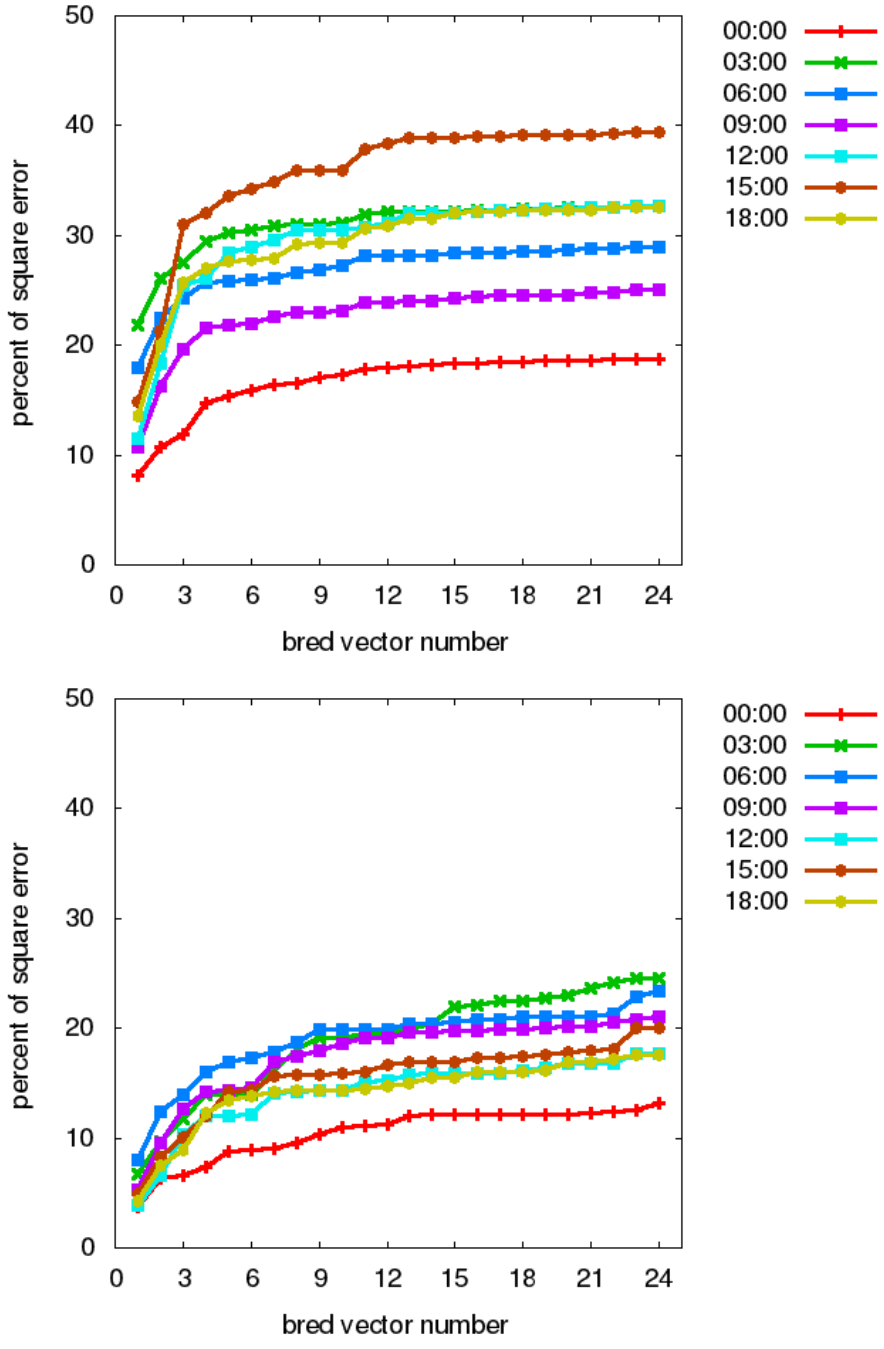

Fig. 6. Square norm of error orthogonal projection onto subsets of bred vectors (percentage of error square norm) as a function of subspace dimension. Different curves in each panel correspond to different times (UTC, 26 September 2006). Top panel: control trajectory 18H (large initial error), rescaling period 30 min, rescaling RMS TUV norm 0.36. Bottom panel: control trajectory R21 (small initial error), rescaling period 15 min, rescaling RMS TUV norm 0.10.

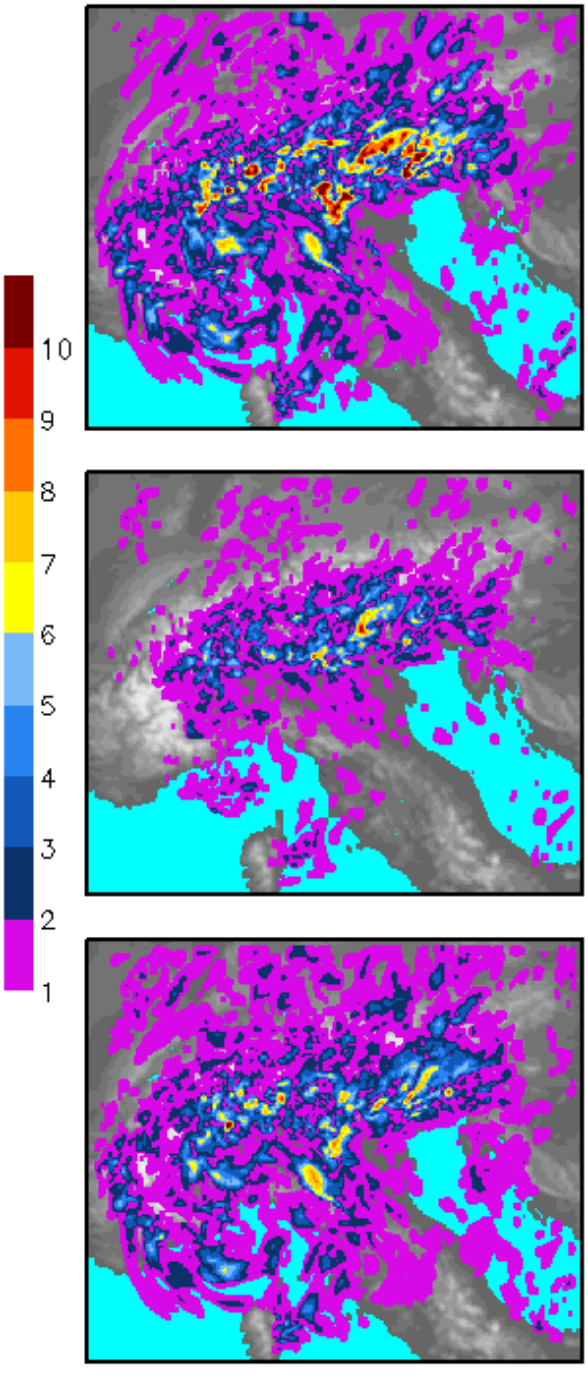

Fig. 7. Maps of error at 03:00 UTC on horizontal velocity (norm of 2-dimensional vector differences, $\mathrm{m\,s^{-1}}$). Top: error of control state 18H (trajectory started with large initial error). Middle: projection of the error on the 12 BV subspace. Bottom: difference between the two above fields, i. e. error of the restart state of Fig. 8.

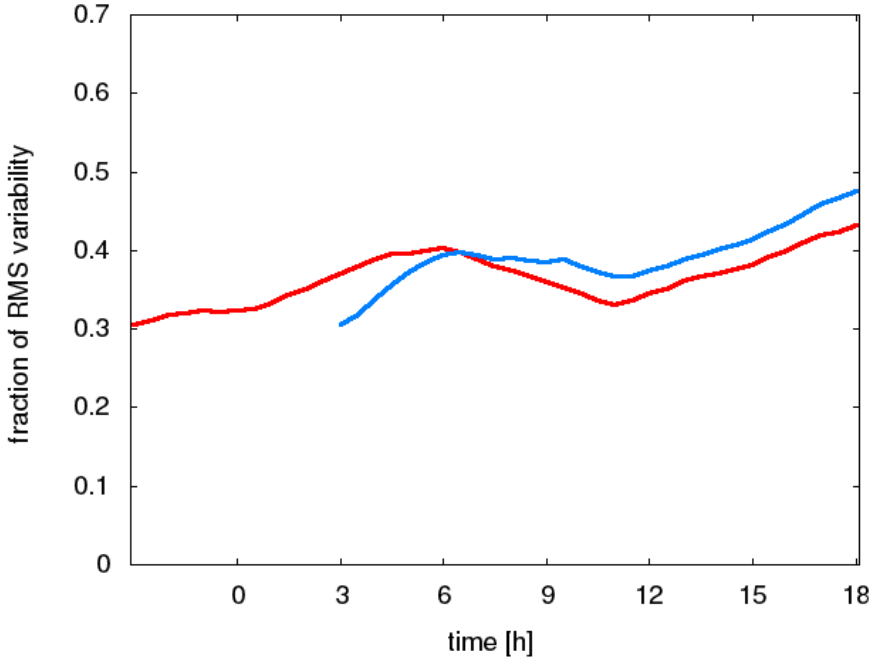

Fig. 8. Red: RMS error evolution of the 18H control trajectory, same as in Fig. 1. Blue: error of the trajectory restarted after subtraction of the orthogonal error component on the 12 bred vectors subspace. Times indicated on the x axis correspond to UTC times on 26 September 2006.