Reducing Systematic Errors by Empirically Correcting Model Errors

FABIO D'ANDREA* AND ROBERT VAUTARD

Laboratoire de Météorologie dynamique
École Normale Supérieure
24, rue Lhomond, 75005 Paris, France

(January 2000, *Tellus*, 52A, 21-41)

*Corresponding Author; dandrea@lmd.ens.fr, tel. +33 1 44 32 22 32, fax +33 1 43 36 83 92

Abstract

A methodology for the correction of systematic errors in a simplified atmospheric general circulation model is proposed. First, a method for estimating initial tendency model errors is developed, based on a four dimensional variational assimilation of a long analysed dataset of observations in a simple quasi geostrophic baroclinic model. Then, a time variable potential vorticity source term is added as a forcing to the same model, in order to parameterize subgrid-scale processes and unrepresented physical phenomena. This forcing term consists in a (large scale) flow dependent parametrization of the initial tendency model error computed by the variational assimilation. The flow dependency is given by an analogues technique which relies on the analysis dataset. Such empirical driving causes a substantial improvement of the model climatology, reducing its systematic error and improving its high frequency variability. Low-frequency variability is also more realistic and the model shows a better reproduction of Euro-Atlantic weather regimes. A link between the large scale flow and the model error is found only in the Euro-Atlantic sector, other mechanisms being probably the origin of model error in other areas of the globe.

1 Introduction

1.1 *Systematic errors and model errors*

The first requirement a GCM is normally asked to fulfill is to have a realistic long term mean flow, or climatology. For this reason the simplest diagnostic performed on GCMs is to observe their systematic error, i.e. the difference between their long term mean flow and the observed climatology for a given variable.

Typical examples of systematic errors of state-of-the-art GCMs are the tendency to have a too zonal middle tropospheric flow and its erroneous northward displacement, especially in the Northern Hemisphere during winter, or a general too wet simulation of middle latitudes of both hemispheres, accompanied by an underestimation of total cloudiness, especially in the tropics. A review of overall performance of a group of state-of-the-art GCMs can be found in the works produced by the AMIP (Atmospheric Model Intercomparison Project), e.g. Gates et al (1998) and references therein.

Improving a model's climatology is not a simple task, for rarely systematic errors show explicitly the signature of the specific model deficiencies that generate them. On the contrary, it is often the case that specific model deficiencies actually generate compensating errors so to shade their effect on the long term mean. Normally, systematic error reductions are attained more by means of modelers intuition and *ad hoc* experimentation than by objective and quantitative analysis of the model formulation.

A natural way to tackle this problem is to examine the so called *model errors*, or initial tendency errors of the model, and try to deduce from them, for example, possible ways to improve the model parameterizations. Model errors are defined as R in the following equation

$$\frac{\partial X_0}{\partial t} = \frac{\partial X_M}{\partial t} + R \tag{1}$$

where X_M stands for any prognostic variable of the model and X_0 stands for the same variable as resulting from observation. Consequently, the model error R is the residual of the modeled evolution of X with respect to its observed evolution. R is also sometimes referred to as *model residual* or *forcing error*. In general R will encompass a mean or "systematic" part, constant in time, and a time-varying part. This transient part can, for the moment, be thought as being merely stochastic, but its nature will be further investigated in this paper.

Also the estimation of the forcing error can be a quite complicate matter. In the literature, the number of studies that propose objective methods of estimating model errors, and their use for improving the model performance, is comparatively small. Examples are given by Klinker and Sardeshmukh (1992) and Schubert and Chang (1995).

Klinker and Sardeshmukh (1992) proposed to use many one-step model integrations to obtain an average estimate of R. They performed these short integrations starting from initialized data, and compared diabatic and adiabatic terms in the equations by successively switching off all the individual parameterizations. In this way they could isolate the contribution of the different terms.

Schubert and Chang (1995), on the other hand, proposed to use short term model forecasts rather than local tendencies. They estimated model errors by the so called analysis increment, that is the difference between analysis and model first guess, calculated throughout the assimilation process in the model.

In both approaches, a fundamental point is that the model error is estimated *for the same model* for which the assimilation of observation is carried out. When such assimilated data is to be used to estimate forcing errors in other models, particular attention should be given to possible initialization problems (moisture spinup, generation of large amplitude gravity waves) and to the estimation of the observed tendency, which may require some high order differentiation scheme. In general, some form of re-assimilation of the analysis in the new model should be envisaged.

A recent work by Kaas et al. (1999) proposes an approach to this problem, based on a simplified assimilation scheme on a perfect model environment. By a newtonian relaxation (nudging) of a low resolution (T21) version of a GCM on the output of a T106 version of the same model, Kaas et al. (1999) could estimate empirical interaction functions for the horizontal diffusion. An integration of the T21 model with the accordingly corrected horizontal diffusion coefficient resulted in a reduction of its difference with the high resolution one.

A number of authors are currently working in a joint effort on the exploitation of model errors for parameterization tuning and systematic error reduction. A brief review of preliminary results can be found in Kaas et al. (1998)

1.2 *Estimation and parameterization of Model Errors*

This article has two main objectives. First, a methodology to estimate model errors is proposed, and second a parameterization is constructed in order to compensate their effect and improve the model's climatology.

The model used is a quasi-geostrophic (QG) baroclinic model, first developed by Marshall and Molteni (1993, hereafter MM93), a brief description of which is given below. This model was used by MM93 with a time-constant source (forcing) term in the potential vorticity equation, this term being nothing but the estimated average tendency error of the non-forced model. This formulation of the forcing term basically accounts for the constant part of R; a formulation of the time varying part of it is the purpose of the present study.

In order to estimate model errors, a variational procedure is proposed, designed to fit the ECMWF operational analysis with the forcing term R as the unknown. This method is compared with a direct finite difference estimation of the observed tendency term in equation (1).

Once the residuals are estimated, a flow dependent empirical parametrization of the model error is developed. The model error R is decomposed as

$$R = \bar{R} + R(X) + R', \tag{2}$$

where $\bar{R}$ is a time constant part, $R(X)$ a part that depends on a given model state variable X, and R' a time varying part, non dependent on X, that can be thought of as stochastic. The transfer function between X (in the case of the QG model X would be potential vorticity) and R is constructed empirically. It will be shown that the application of such time varying model residual as a forcing term in the equations not only reduces the systematic error of the model, but also substantially improves its high and low frequency variability in the Euro-Atlantic sector.

1.3 *Outline*

In section 2 a brief description of the model used in this work is given, with particular attention to the forcing term.

Section 3 describes the methodology developed for estimating model tendency errors (subsection 3.1) and for building the flow dependent parameterization of the forcing in the QG model (subsection 3.2)

Section 4 contains the results obtained by applying the forcing parameterization to the QG model. Systematic errors and standard deviations at high and low frequency are shown (subsection 4.1). To characterize the low frequency variability of the model, weather regimes are also defined by cluster analysis and compared in the different integrations and in the analysis (subsection 4.2).

A discussion and interpretation of the results is reported in section 5. The discussion mainly consists in a sensitivity study to the main parameters of the forcing parameterization, including the regional sensitivity of the model to forcing perturbation, which allows to make some intesting physical and dynamical remarks.

Section 6 concludes with a summary of the results, including some ideas for future developments.

2 The quasi geostrophic model

The model used in this study is spectral on the sphere with a triangular truncation at total wavenumber 21. Vertical discretization is performed in pressure coordinates at three levels (200, 500 and 800 hPa). For each level, the model consists in the quasi-geostrophic potential vorticity equation with dissipation and forcing:

$$\frac{\partial q}{\partial t} = -J(\psi, q) - D(\psi) + S \tag{3}$$

where J represents the jacobian operator and ψ is the QG streamfunction, linked to the QG potential vorticity q by a relation $q = \mathcal{L}\psi$. The linear dissipation term D includes in a simple way the Ekman dissipation (orography dependent) and a newtonian relaxation between the layers. The orographic contribution to the potential vorticity at the lower layer is included in the operator $\mathcal{L}$. For a more detailed description of the model the reader is referred to MM93, while some more attention will be devoted to the forcing term S in the following.

The forcing term is designed to include the sources of potential vorticity that result from processes not explicitly included in the equations. These phenomena are typically sea-atmosphere interactions, diabatic heat fluxes (linked for example to precipitations etc.), and the effect of the divergent flow. On top of this the forcing implicitly contains the effects of subgrid scale processes. The forcing term have been estimated empirically by MM93 as follows.

Having a long series of analysed states, $\hat{q}$ and $\hat{\psi}$, one can write, by averaging (3) in time,

$$\overline{S} = \overline{J(\hat{\psi}, \hat{q})} + D(\overline{\hat{\psi}}). \tag{4}$$

With this mean source term, computed from a long dataset of observed winter data, MM93 (and many other authors thereafter) performed long integrations of the model, which showed a satisfactory degree of realism. (4) can in turn be written as

$$\overline{S} = J(\overline{\hat{\psi}}, \overline{\hat{q}}) + D(\overline{\hat{\psi}}) + \overline{J(\hat{\psi}', \hat{q}')} \tag{5}$$

where overbar denotes time average and primes the deviation from this time average. The latter decomposition allows to intepret the first two terms of the r.h.s. of equation (5) as imposing a basic state, and the last term as adding the average contribution from transient eddies to the flow. A necessary condition for the model to reproduce the exact climatology is that its transients fluxes exactly match the contribution of $\overline{J(\hat{\psi}', \hat{q}')}$.

The only difference between the model used by MM93 and the present one is the numerical time integration scheme; here a predictor-corrector (first order Adams-Bashforth-Mouton) is used, while MM93 used a Leap-frog scheme. The time step used is 1 hour. A climatology of 500hPa streamfunction of a long (10000 days) integration of the model is shown in Figure 1; it includes long term mean, systematic error and low and high frequency standard deviations. Hereafter this integration will be referred to as the *control integration.*

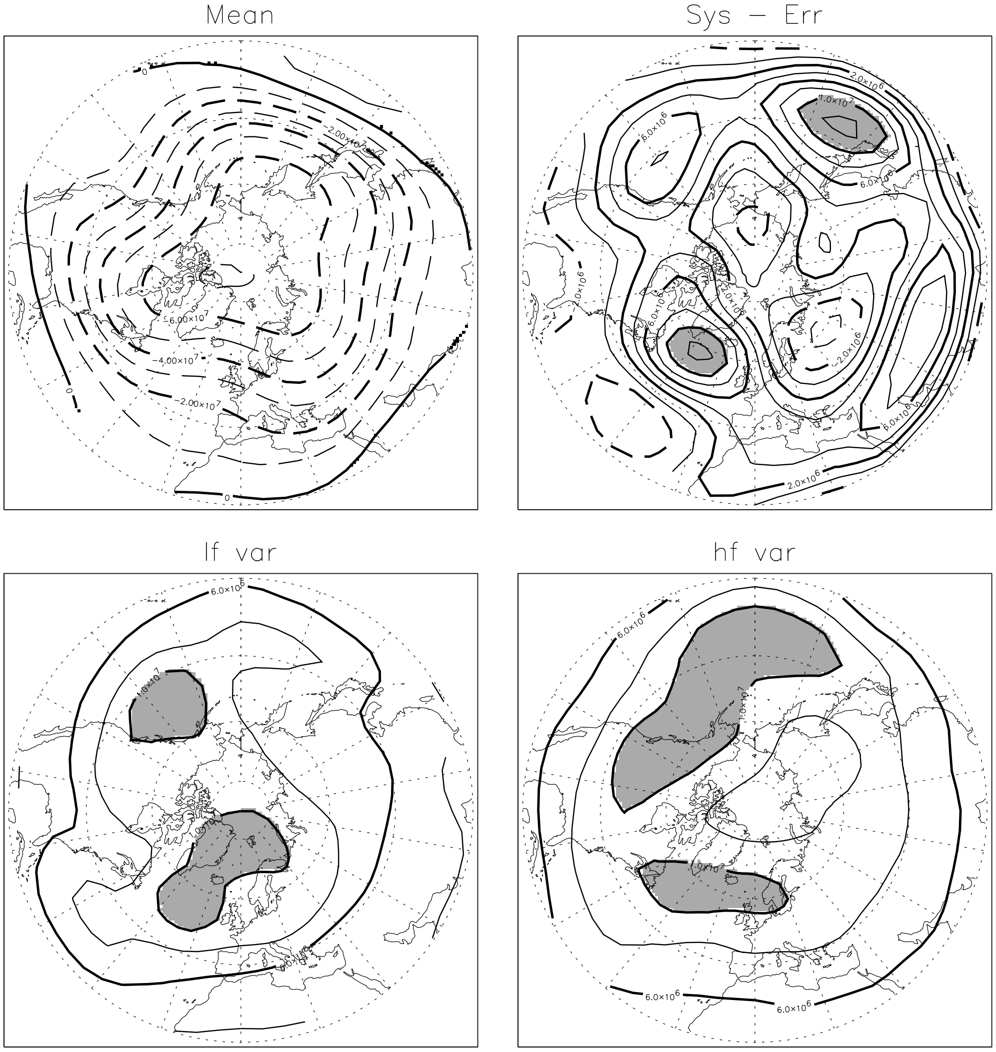

Figure 1: 500 hPa streamfunction climatology of the control integration. Top left: long term mean. Top right: Systematic error (i.e. difference of long term mean of model minus long term mean of analysis). Bottom left and right : Low and high frequency standard deviation (10 days square window filter). Contours every $10^7 m^2 s^{-1}$ in the long term mean field; every 2 10^6 $m^2 s^{-1}$ in the other three panels. Shading for amplitudes higher than 10^7 $m^2 s^{-1}$ negative or positive.

The forcing term used in the control integration was calculated by (4) using a twice daily ECMWF analysis dataset ranging from December 1984 to February 1994 for December-January-February (DJF), and consequently, the integration can be considered as a perpetual winter one. A climatology from the analysed dataset, similar to that in Fig.1 is shown in Fig.2 for comparison. Only the Northern Hemisphere middle latitudes are considered, for the QG approximation is only valid in the winter extratropics, and the forcing term S is computed from NH winter data. The SH climatology, with this forcing term, is in general less realistic and the model is integrated globally basically to avoid boundary conditions problems. A simulation of the SH could improve by the use of a forcing term computed from SH winter data, but this is not in the scope of the present work.

The two main areas of high systematic error of the control integration (Fig. 1) are over the North Atlantic and the Northwest Pacific. Two relative maxima are also found over the east Pacific and central Asia, giving on the whole a wavenumber 4 signature. The North-Atlantic positive error on the north side is accompanied by a negative anomaly to the south, revealing an excessive diffluence of the flow, clearly visible also in the mean field panel. The maximum value of the positive error in this region exceeds 1.2 10^7 $(m^2 s^{-1})$. For a measure more familiar to meteorologists this value can be approximately expressed in meters of geopotential height multiplying by $\frac{f_0}{g}$, where as usual f_0 is the Corolis parameter at $45^o N$ and g is the acceleration of gravity at the surface of the Earth. This transformation gives a peak error of 126 m, which is substantially higher than those typically found in state-of-the-art climate GCMs.

High and low frequency variabilities are computed by a simple 10-day square filter (the standard deviation of ten-days means and of the departure from ten-days means). In the control integration, the underestimation and north-eastward displacement of the atlantic stormtrack is evident, as well as a related northward displacement of the atlantic low frequency center of activity.

Notwithstanding the errors described above, the overall climatology of the model, given its simplicity, is rather satisfactory. The aim of this paper is to further reduce its systematic error and to improve the variability patterns.

3 Methodology

3.1 *Estimation of model residuals*

The easiest and most natural way to compute model residuals is to compute R_t directly from the definition (1) for every available analysed field $\hat{q}_t$. One is thus left with the estimation of the tendency $\frac{\partial \hat{q}_t}{\partial t}$, which can simply be calculated by finite differences. The residual at time t is therefore computed

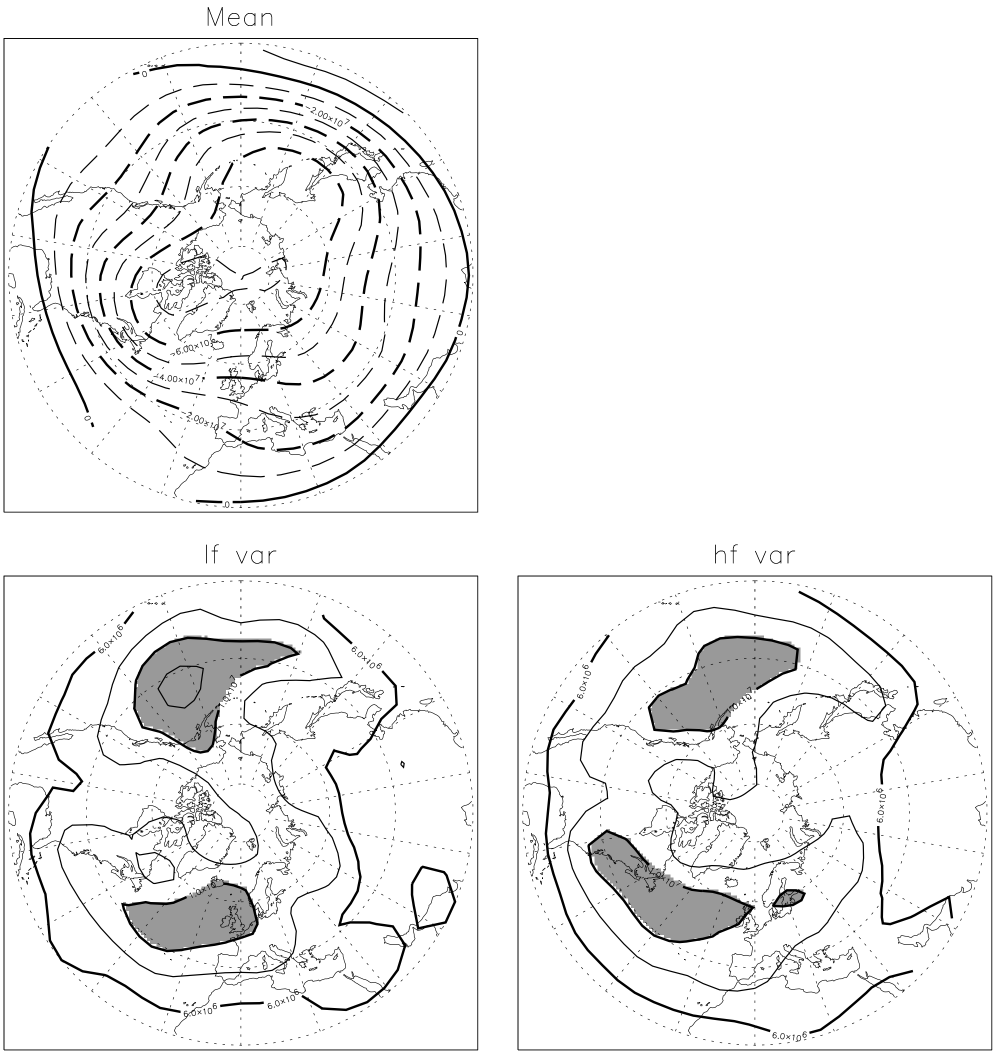

Figure 2: Like Fig.1 but for the ECMWF operational analysis for the period December 1983 - February 1994. DJF only. Systematic error is obviously missing.

as:

$$R_t = \frac{\hat{q}_{t+\Delta t} - \hat{q}_{t-\Delta t}}{2\Delta t} + J(\hat{\psi}_t, \hat{q}_t) + D(\hat{\psi}_t), \tag{6}$$

where Δt is the interval of time at which analyses are available, in this case $12h$. This definition, applied to the whole of the available analysed dataset, builds up a series of residuals (hereafter called "direct residuals") that approximate instantaneous forcing errors. Averaging all the R_t's simply provides the MM93 forcing term given by (4).

Residuals computed in this way are not completely satisfactory; their main shortcoming is that the long interval at which the analysis is available makes it difficult to estimate the actual tendencies. Moreover, if a more realistic primitive equation model was used, an additional problem would be given by the difficulty of distinguishing between "true" model tendency and tendencies related to initial gravity wave noise and moisture spin-up, which are generated mainly by the inconsistency between the model used for assimilation and the model for which the residuals are computed. In the case of the QG model, these initialization problems are inherently less important.

Another approach is here developed, which has the advantage of being of general applicability and not limited to QG models. As will be shown later, this new method also gives better results in terms of systematic error reduction. A re-assimilation of the analysed dataset has been carried out on the QG model in a procedure quite similar to 4D-var, with the difference that the control parameter for the variational problem is model forcing instead of initial condition.

More explicitly, an integration of the model is conducted following the analysis. At each time t when an analysed field is available, a forcing term is estimated so as to minimize the model forecast at time $t + \Delta t$, i.e. to minimize a cost function defined as $J(S) = \frac{1}{2}||q_{t+\Delta t} - \hat{q}_{t+\Delta t}||^2$, the hat indicating as usual the analysed fields, and Δt still being the interval at which the analysis is available. The minimization algorithm requires the knowledge of the gradient of J which is computed at given points by the use of the adjoint model, as in classical variational data assimilation (see for example LeDimet and Talagrand 1986). In appendix A the method is explained with more details. The norm chosen is the potential enstrophy norm. The minimization algorithm used in this study is a quasi-Newton method that has been developed at INRIA (Gilbert and Lemarechal, 1992).

The minimization yields very good results, reducing the forecast error at $12h$ well beneath the computer accuracy.

For different reasons none of the approaches is ideal. In the case of direct residuals, one faces large temporal discretization and possible unconsistency between the model used for assimilation of analysis and the model for which the residuals are computed. In the case of variational estimation, although the residuals obtained (hereafter called "optimal residuals") are optimal in a least square sense on a period of time, their definition depends strongly on the assumption that forcing is constant over the period $[t, t+\Delta t]$. This is quite an *ad hoc* hypothesis due to data availability rather than a physically based one.

3.2 *Forcing parameterization*

In the previous subsection two different datasets of model residuals (the "direct" and the "optimal" residuals) have been produced. The scope of this subsection is to try and use these data in order to parameterize the model forcing at a given time t as a function of the model flow. In other words one needs to build a transfer function from the space of model states to the space of residuals. Since no information is available about the form of such a function, the most natural approach is the use of analogues.

The idea of an analogue parameterization is simple. At given times during the model integration, the model state is compared to all the analysed atmospheric states, and the nearest neighbors (by a metric to be defined) are chosen as its analogues. The residuals that are simultaneous to the closest K analogues are then averaged to provide the flow dependent forcing term.

Analogues have long been the object of research, essentially for empirical weather forecast purposes, following a seminal article by Lorenz (1969). In the case of prediction, the most relevant problem with this method is the lack of long enough series of historical data. Atmospheric states have a relatively high number of degrees of freedom, so that long time series are necessary to "explore" its phase space to produce analogues accurate enough to have a practical forecasting capacity (Van den Dool, 1994). This problem can somehow be reduced by wisely filtering the data timeseries in order to lower the dimensionality of its phase space, and therefore find more accurate analogues.

As stated before, one needs to define a metric in the phase space of model states, in order to define analogues. In this work, the model state is represented by the global potential vorticity field, but the analogues are sought only within the Euro-Atlantic region ($30.4^oN - 69^oN$ and $78.75^oW - 39.4^oE$) by an euclidean distance. The choice of limiting to a region of the globe is given by the necessity to reduce the dimensionality of the space where analogues are sought for, further discussion on the choice of the region is found in section 5.1 below. In order to further reduce this dimensionality, the distance is actually computed in the EOF (Empirical Orthogonal Functions) space truncated to the first T modes, this filter having also the effect of limiting the computation to the large scale fields. The distance is computed on the three vertical levels of the model, the EOFs of the different levels are computed separately, but the PCs (principal Components) are normalized by their variance in order not to weight one level more than the others.

In practice, the parameterization of forcing in the model follows the following steps:

1) at intervals of time τ the model state $q(t)$, in the Euro-Atlantic sector, is projected on the EOFs of the analyses.

2) The distance of $q(t)$ from each analysed field is computed on the leading T EOFs. The analysed fields are classified with respect to this distance.

3) The K model residuals simultaneous to the K closest analogues are averaged to give the forcing field for the following interval of time $[t, t+\tau]$.

Note that although the analogues are computed on a limited region, the associated residuals are global fields and, therefore, so is the forcing.

There are three tuning parameters for this procedure; the interval of time τ at which forcing is refreshed, the number T of PCs on which the distance is computed, and the number K of analogues considered. In section 5 a discussion about the effects of these parameters is presented. By trials and errors, the best combination found was $\tau = 12h$, $K = 40$ and $T = 15$. 15 modes explain about 70 to 80 percent of the total variance, depending on the vertical level. The dependence on the region chosen for the computation of analogues will also be discussed later.

4 Results

4.1 *systematic error and standard deviations*

Figures 3 and 5 are similar to Figures 1 and 2, but refer to long model integrations (10000 days) performed with the forcing parameterization described above and with the use, respectively, of optimal and direct residuals. These integrations can be hereafter called *empirical forcing integrations.* A comparison with Fig. 1 gives the measure of the improvement obtained.

Concentrating first on optimal residuals (Fig.3), it can be observed that the positive anomaly in the North Atlantic is reduced to 6 $10^6 m^2 s^{-1}$ which is equivalent to about 64 meters of geopotential height. Surprisingly enough, the improvement is found also over regions other than the North Atlantic, and the amplitude of the error is reduced almost everywhere in the Northern Hemisphere. The only exception is represented by the Arctic region that shows an increased, albeit very slightly, negative anomaly. Systematic errors of this amplitude at 500 hPa are comparable to those found in more complex GCMs used for climate simulations. See for example D'Andrea et al. (1998).

The variability maps also bear the sign of this improvement, the high frequency showing a better placed and extended atlantic stormtrack and a related displacement of the low frequency anomaly towards the correct location. The Pacific region this time does not show any improvement, the stormtrack being even more overestimated. The improvement of the atlantic stormtrack is far more evident at 200 hPa as can be seen in Fig. 4, where the high frequency

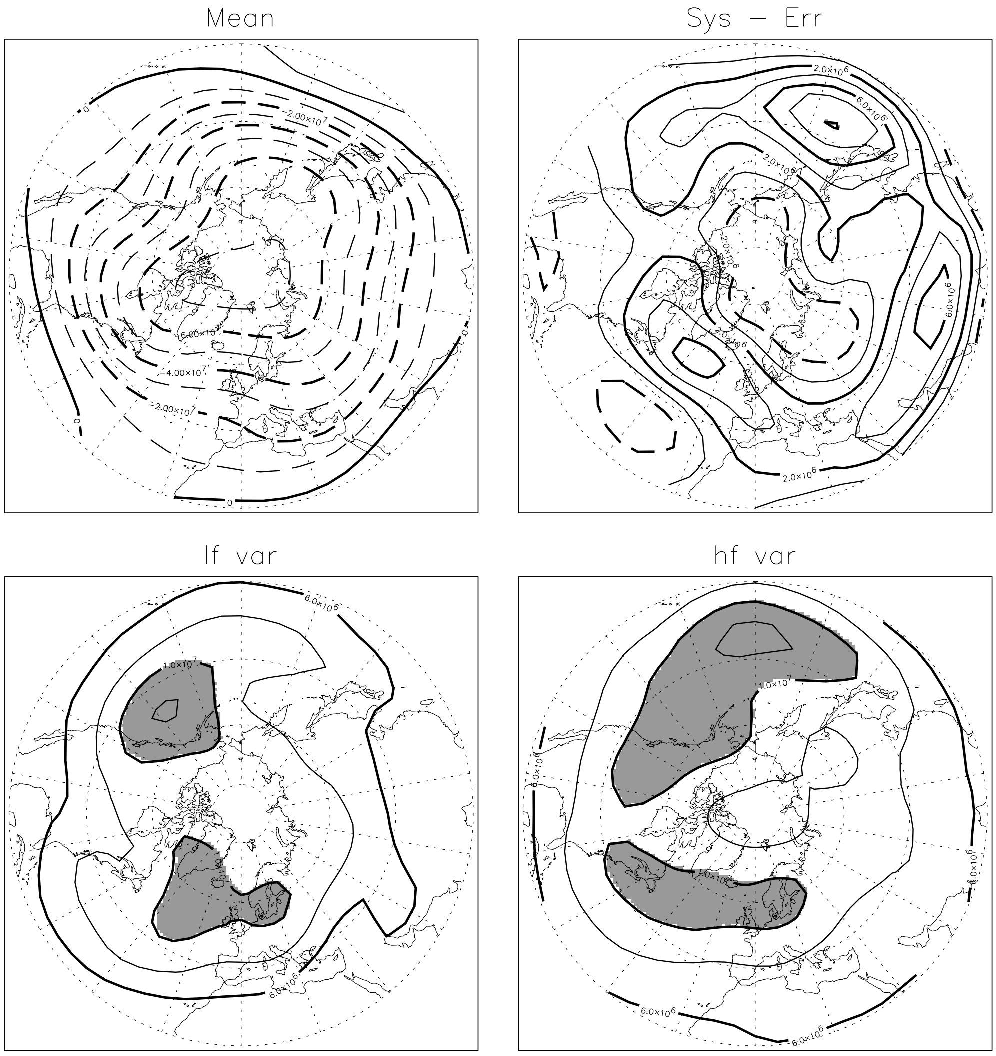

Figure 3: Like Fig.1 but for the empirical forcing integration. Optimal residuals. $\tau = 12h$, $K = 40$, $T = 15$. See text for details.

variability of the analysis, control and empirical forcing integration are shown. The atlantic stormtrack, almost missing in the control integration, has a correct position and amplitude in the empirical forcing one.

Focusing now on direct residuals (Fig.5), most of the features of Fig.3 are found. The main difference regards the North Atlantic, where the systematic error is less reduced than in the previous case. The maximum reaches the equivalent of 85 meters and the region of error is much more extended.

From now on, more attention will be devoted to the case of optimal residuals. Comparable results, although less accurate, are also found with direct residuals.

The robustness of these results was tested in two ways. First, to be sure of the statistical significance of the climatology, two other empirical forcing integrations were performed starting from different initial conditions. The same result was obtained (not shown). Second, the analogues method was tested against a simple stochastic variation of the forcing field. A cheap way to obtain a stochastic model of the time-varying forcing term is to perform an integration defining the forcing as the average of M residuals chosen randomly; all other parameters are left unchanged. The number M controls the amplitude of the forcing variance and was chosen in order to be approximatively the same as in the empirical forcing integration, this constraint gave $M = 12$[1].

The climatology resulting from this *stochastic forcing* integration is not different from the control integration, i.e. no reduction of the systematic error, with a general increase of variability. This last test positively validates the existence of a relation between the model flow and the forcing, and the ability of the empirical parameterization to reproduce it to some degree.

A way to measure the performance of the parameterization can be constructed computing the forcing terms relative to each analysed field and compare it to the optimal residuals. Linear correlation have been computed for every gridpoint of the timeseries of the so computed forcing terms and the timeseries of the residuals. The so obtained map (not shown) shows a region of maximum correlation over the North Atlantic with a maximum amplitude of 0.4 centered on southern Greenland. No significant correlation was obviously found in the case of stochastic focing terms. For this test, analogues have been sought only for analysed fields belonging to different years. Consequently, it can be said that at least 16% of the total variance of the forcing errors can be reproduced by the analogues.

[1] It is simply $M = \frac{\|{\sigma_R}^2\|}{\|{\sigma_{emp}}^2\|}$ where σ_R is the standard deviation of the residuals dataset and σ_{emp} that of the empirical forcing term. The norm is taken over the Northern Hemisphere.

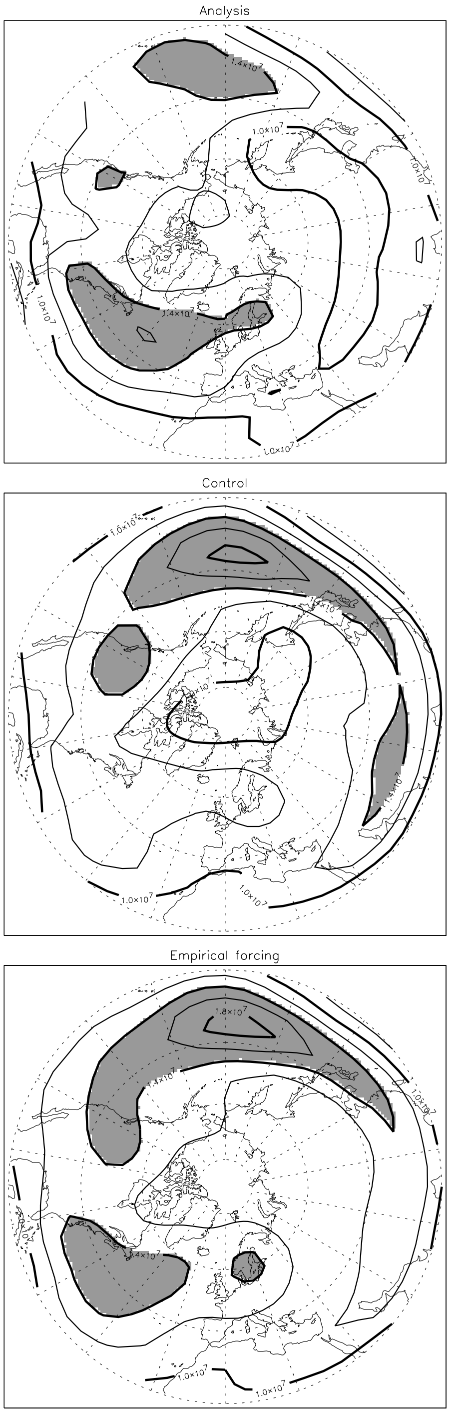

Figure 4: High frequency variability at 200 hPa, computed as in Fig.1, for the analysis, control and empirical forcing integrations. Contours as in Fig.1, Areas of amplitude higher than 1.4 10^7 m^2s^{-1} are shaded.

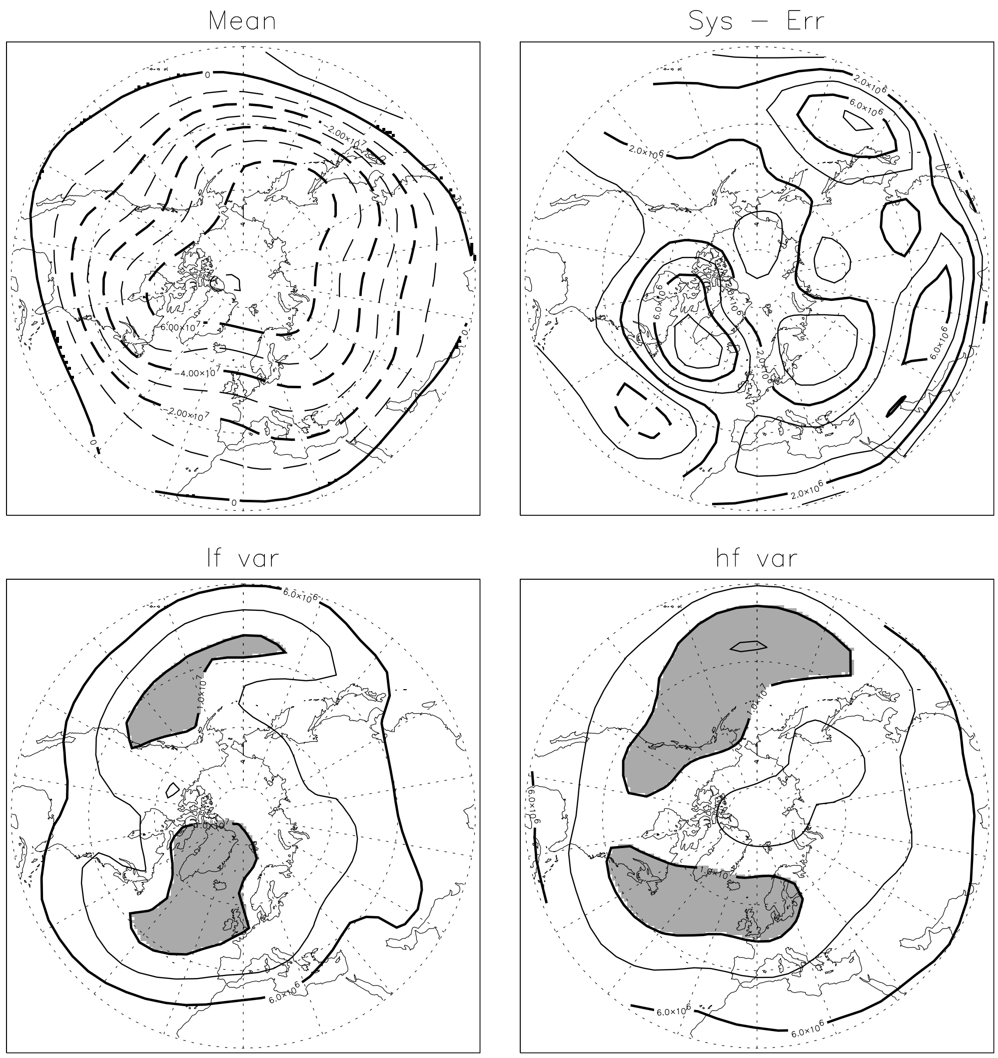

Figure 5: Like Fig.1 but for the empirical forcing integration. Direct residuals. $\tau = 1h$, $K = 40$, $T = 15$. See text for details.

4.2 *Low frequency variability: weather regimes*

In order to further examine possible improvements of the model due to the flow dependent forcing parametrization, the simulated low frequency variability is investigated by looking at weather regimes and their modifications in the control and in the empirical-forcing integrations.

The concept of weather regimes originates from the idea, borrowed from simple model studies, that the midlatitude large scale atmospheric circulation could be described by a low number of quasistationary states and transitions between them. Recent observational works, using various statistical and dynamical methods, have more or less supported the existence of these states, which are called regimes. It is not in the scope of this paper to describe or review the regimes theories and definitions. The reader is reported to the review included in the paper by Michelangeli et al. (1995, MVL95 hereafter), the methodology of which is also used in the present paper. MVL95 used a non-hierarchical clustering algorithm that finds an optimal partition of the the dataset into a given number of clusters, so to minimize variances within these clusters (the "dynamical clusters" method). The centroids of the obtained clusters are indicative of the most recurrent states of the system.

In practice, almost the same procedure of MVL95 is followed here. The algorithm is applied to daily 500 hPa streamfunction fields, filtered on the 8 leading EOFs (explaining around 82% of the total variance). Since the number of clusters (classes) must be given *a priori* to the algorithm, a criterion must be devised to chose it. MVL95 employed a statistical argument relying on a montecarlo test of the significancy of a partition in a given number of classes, and found four significant cluster on analysed data. Michelangeli (1996) performed a similar test on the same QG "control" model used in this study and also found a significant classification in four classes. Consequently, in this paper a classification into four classes is also performed, and the synoptic patterns of the clusters centers is compared in the different integrations. As long as this article is concerned, weather regimes will basically be used as a diagnostic tool.

Figures 6, 7 and 8 present the MVL95 cluster centroids for the Euro-Atlantic region of the ECMWF analysis, control and empirical forcing integrations respectively.

In Fig.6 the well known observed Euro-Atlantic Regimes anomalies can be seen. In panel 1 is the *zonal* regime, which displays an enhancement of the jet; in panel 2 is the signature of the *blocking* regime; in panel 3 the high pressure in the central Atlantic basin represents the *atlantic ridge*; and finally in panel 4 a strong *Greenland ridge* is found. These regimes are in good agreement with MVL95 despite the different data used.

Turning now the attention to the control integration (Fig.7), it can be seen how the climatology error is reflected by the regimes. The zonal regime shows

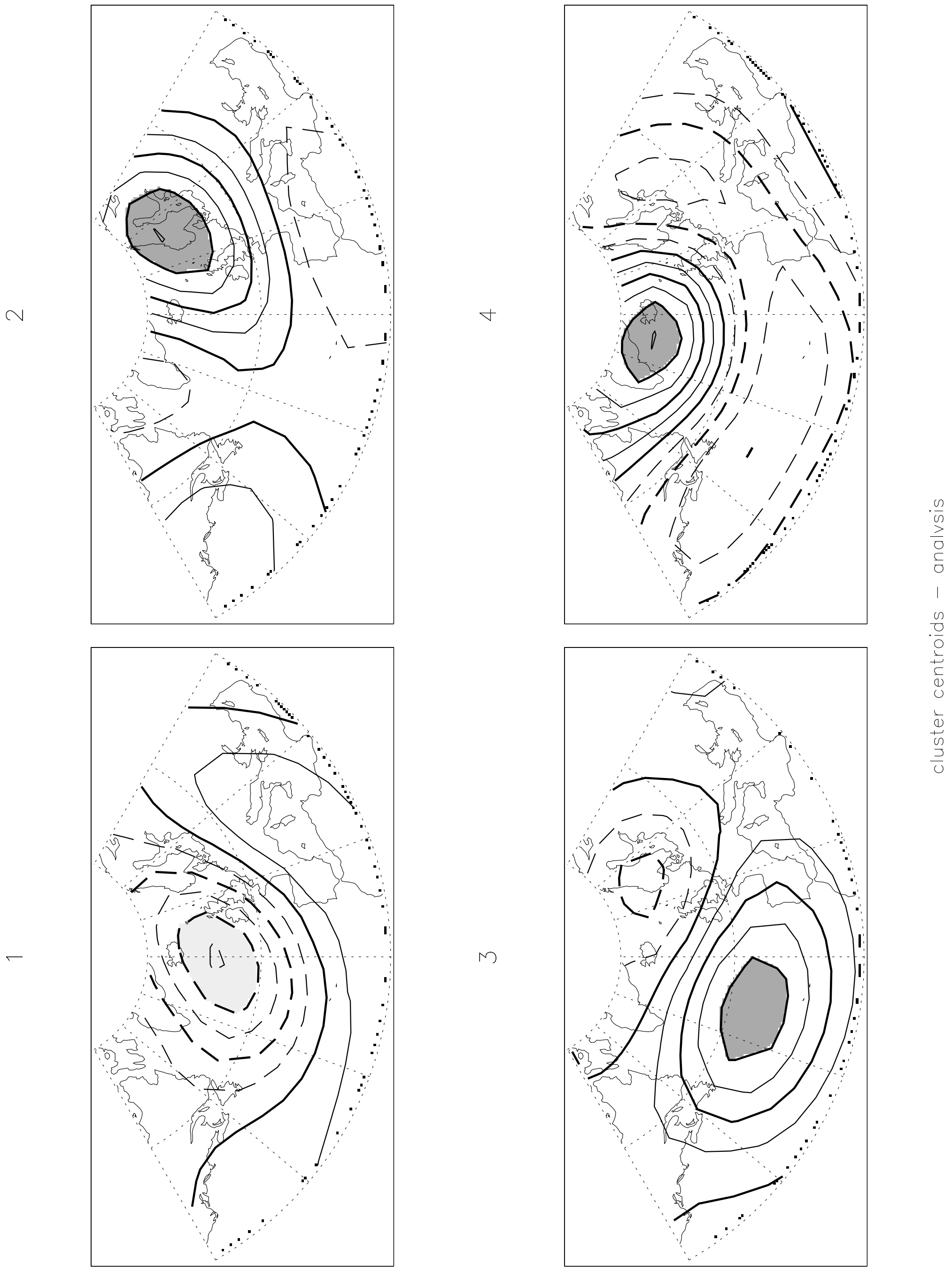

Figure 6: Cluster analysis centroids for the ECMWF analysis of streamfunction fields at 500 hPa, same dataset as Fig.2. Contour every 3 $10^6 m^2 s^{-1}$. Areas of amplitude higher than 1.2 10^7 $m^2 s^{-1}$, negative or positive, are shaded.

Table 1: Anomaly correlation coefficients of modelled regimes centroids with the correspondent regimes centroids of the analysis

	zonal	blocking	atl. ridge	gre. ridge
control	0.662	0.712	0.756	0.761
empirical forcing	0.870	0.864	0.856	0.884

a slightly too low trough and is displaced to the west (compare panels 1 of Fig. 6 and 7). The blocking regime is slightly too low and displaced to the west (compare panels 2). The atlantic ridge is heavily underestimated (panels 3). The Greenland ridge is slightly overestimated and extends too much to the south (panels 4). The incorrect simulation of the regimes by the control model is not simply due to a projection of the systematic error on the different regimes. This can be proven simply "rebiasing" every simulated field to the correct climatology (by subtracting the systematic error) before performing the cluster analysis. The cluster centroids obtained in this way (not shown) are not at all better than those in Fig.7; on the contrary they show the sign of the systematic error anomaly, changed in sign. The regimes depend on the whole structure of the low frequency variability of the model and its interaction with the high frequency transients. For example, the reduced stormtrack causes a displacement north-west of the low frequency variability area (see Fig.1) and a related displacement westward of the zonal and blocked regimes.

The simulation of regimes substantially improves in the empirical forcing integration as can be seen from Fig.8. The low pressure associated to the zonal regime (compare panels 1 of figures 6, 7 and 8) is more correctly placed, and so does the blocking regime which also exhibits a more realistic amplitude (compare panels 2). The displacement of these two features eastward to their correct position reflects the better simulation of the atlantic stormtrack. The atlantic ridge is more recognizable, being well placed and having a more realistic amplitude (panels 3). Last, the Greenland ridge has a more realistic extension, although it remains slightly overestimated (panels 4).

The improvement in weather regimes reproduction is summarized and quantified by Table 1 which reports the anomaly correlation coefficient of the regimes of the two integrations with respect to the corresponding ones of the analysis: in the empirical forcing integration the correlations are all between 0.85 and 0.90. The quantitatively most important change is found in the zonal regime, which reflects a better average location of the midlatitude jet off the north american east coast.

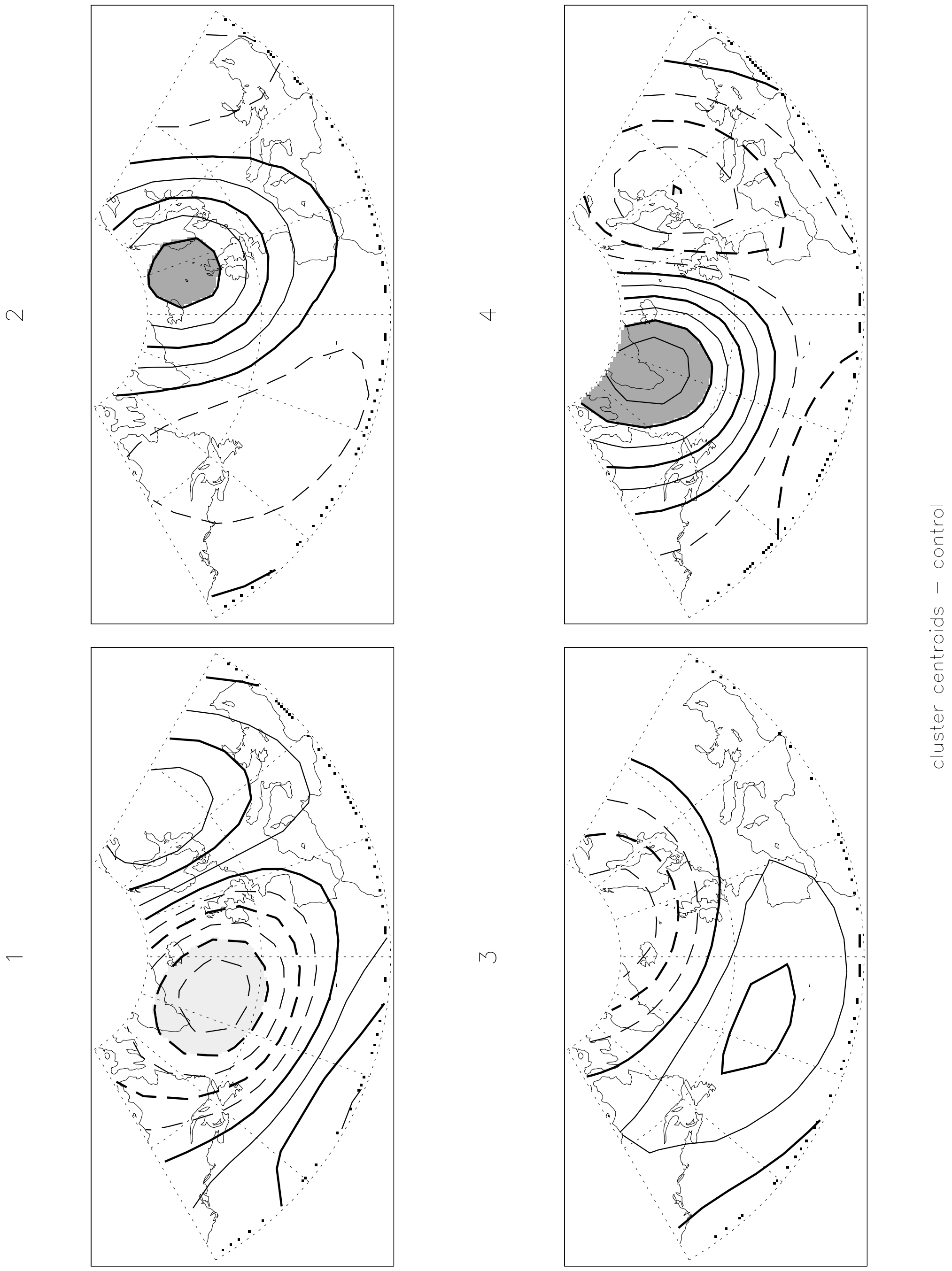

Figure 7: Same as Fig.5 but for the control integration.

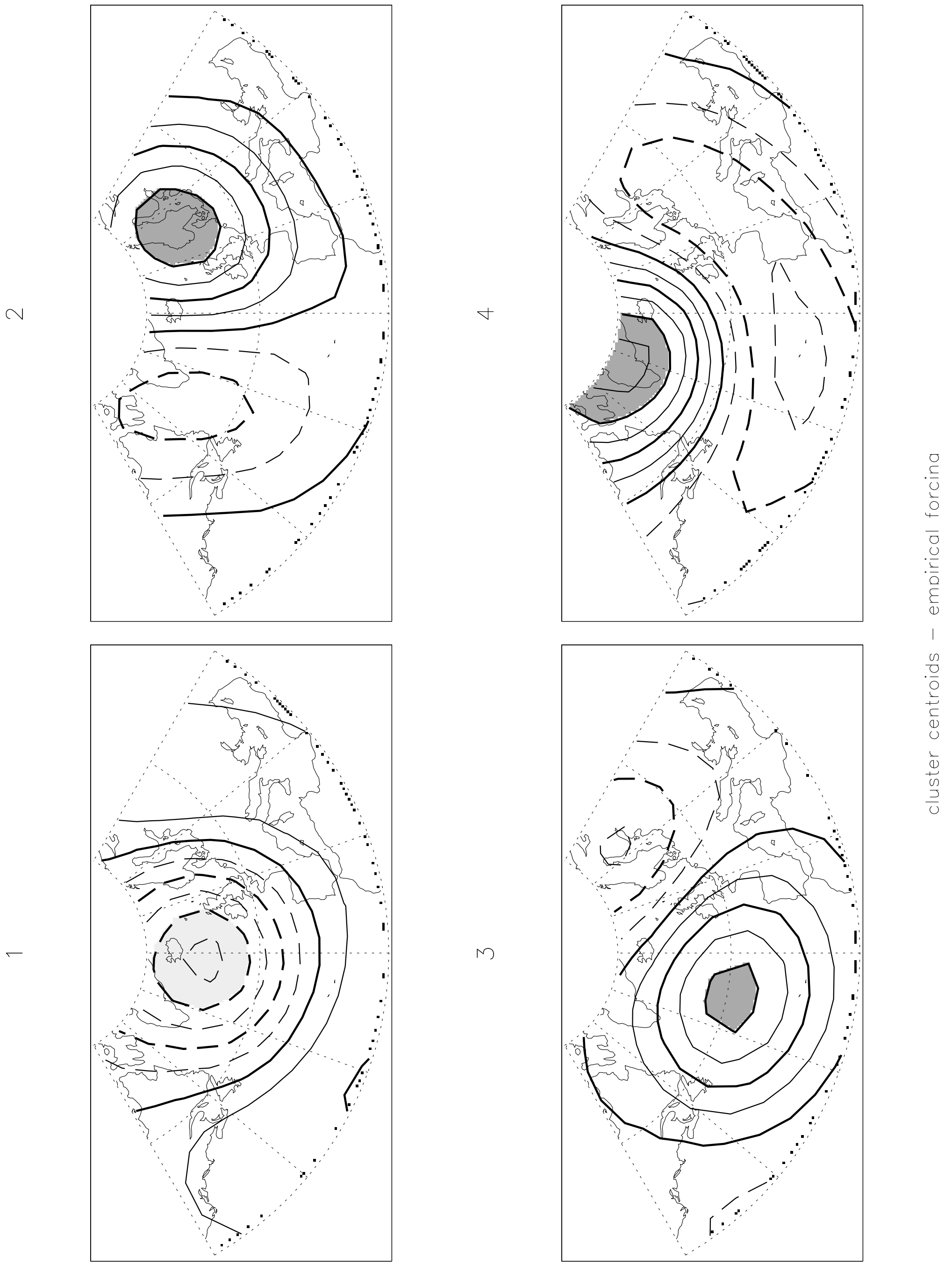

Figure 8: Same as Fig.5 but for the empirical forcing integration.

5 Discussion

5.1 *Regional dependence*

The results shown in the previous sections refer to analogues defined in the Euro Atlantic region. Other regions have been tried, but no such results have been found. A discussion on this subject is presented here.

Restricting to a region of the globe when looking for analogues is a necessary step in order to reduce the dimensionality of the phase space for the analogue method. Some experiments were also conducted by searching analogues over the whole Northern Hemisphere, but no improvement was found on the model performance. Nevertheless, limiting oneself to a region on one side improves the quality of analogues, but on the other should obviously restrict the extent of the parameterization itself.

The relation between forcing and flow is provided by the analogues, but there is no indication of any physical explanation of it. The hypothesis can be made that such a relation is *local* in space; The distance between model gridpoints is much larger than the distance spanned in a model timestep (1 hour) by the fastest waves present in the atmosphere, that are likely to transport information. Consequently, the forcing given by analogues in a region will be "optimal" for that region only. Nevertheless, in the course of model integration, such forcing can have an indirect effect in other regions of the globe.

It is found that the model is more sensitive (in a sense to be defined below) to a change of forcing in some regions of the globe than in others. It is therefore important to optimize the forcing term for these "sensitive" regions.

A measure of the sensitivity to a change of forcing can be built making use of the *singular vectors* of forcing. If one linearizes the model (3) around a trajectory $q(t), t\in [t_0, t_N]$, the linear tangent model can be written in integral form as

$$\delta q_{n+1} = A_n \delta q_n + \delta S, \tag{7}$$

where A_n stands for the Jacobian matrix of the linearized model at time t_n (compare with annex A), δq is a small perturbation on q and δS a small perturbation of the forcing. Now, if $\delta q_0 = 0$, and the only perturbation at time t_0 is due to a change of forcing, (7) becomes (see the explicit form in annex A):

$$\delta q_N = L(t_0, t_N)\ \delta S \tag{8}$$

where the linear operator L is the so called *linear propagator* of (7) which depends on the nonlinear trajectory $q(t)$, but operates, in this case, only on δS. Consequently, the amplitude of the perturbation at time t_N is given by

$$\|\delta q_N\|^2 = < L\ \delta S, L\ \delta S >,$$

or also

$$\|\delta q_N\|^2 = < L^* L \ \delta S, \delta S > \tag{9}$$

where L^* is the adjoint of L and $< \cdot, \cdot >$ is the scalar product associated to a given norm $\| \cdot \|$. From (9) it becomes evident that the eigenvectors of $L^* L$ associated to the largest eigenvalues give the forcing that causes the maximum growth (in a linear sense) of the perturbation of q between times t_0 and t_N. Normally the interval of time $t_N - t_0$ is referred to as *optimization time*. The formalism of (9) is similar to the more classical initial condition singular vectors, used for ensemble forecasting and sensitivity studies, that derives from (7) in the case $\delta S = 0$ and $\delta q_0 \neq 0$ (see for example Buizza and Palmer, 1995).

A measure of model sensitivity to forcing has then been built as follows. Forcing singular vectors are computed at 500 hPa, starting every day of the months of January in the ECMWF analyis. The optimization time is taken here as $12h$, for sake of consistency with the calculation of optimal residuals. The norm used is potential enstrophy over the whole globe. The average of all the singular value spectra[2] is shown in Figure 9. It is evident that the first vectors have a much higher effect than the others. The areas of higher amplitude of the vectors correspond to areas of maximum sensitivity. Consequently, the root mean square of the first five singular vectors on the whole period, weighted by their singular value, is presented in Fig.10. (The map has also been normalized by its maximum value.)

The region of maximal sensitivity is found over the North Atlantic in correspondence of the atlantic stormtrack, and two other regions of high sensitivity are found over the east Pacific and in correspondence of the pacific stormtrack. This structure is very much resembling to a similar figure in the cited work by Buizza and Palmer (1995), the main difference being represented by the high sensitivity region in the east Pacific.

These three areas are consequently those were the forcing term must be sought for with particular care. Nevertheless, a high sensitivity does not by itself guarantee a good performance of the analogues in a region. In fact attempts to apply the empirical forcing parameterization by looking for analogues in the two regions other than the North Atlantic have been conducted, but without results. It seems that in these two regions, by contrast with the North Atlantic, no relation can be found between flow and forcing. On the other hand, some experiments have been made to test the stability of the results from the choice of the extension of the Euro-Atlantic region. Notably an attempt has been made computing the analogues in the small area of the

[2]the spectra have a very small dependence on the starting day. The average spectrum has a standard error of only 5%.

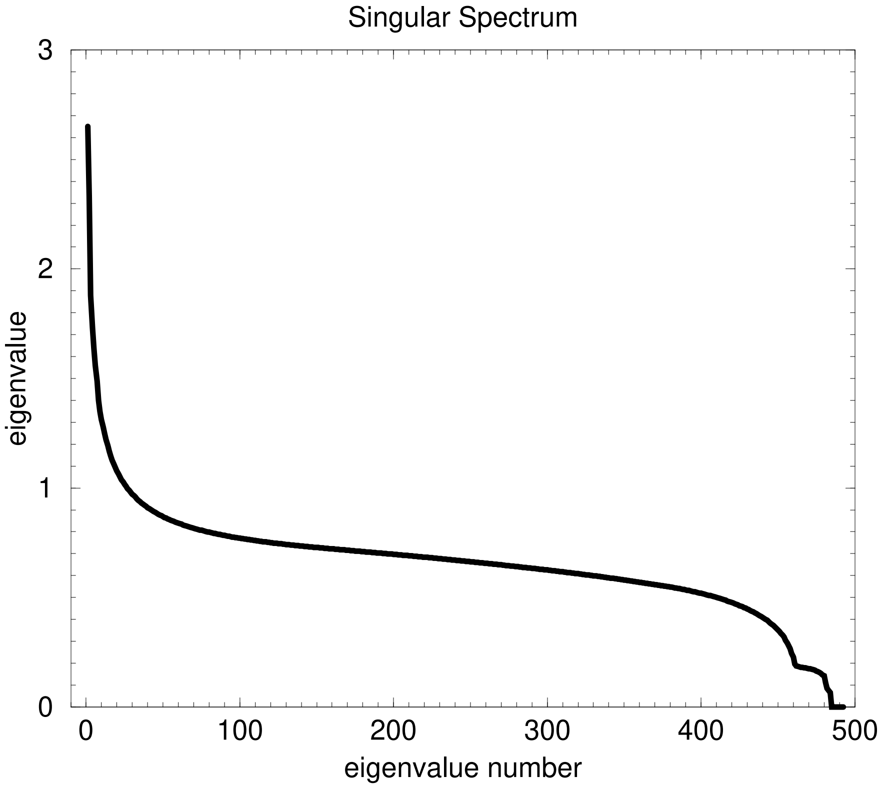

Figure 9: Average singular spectrum of the linear tangent model. 500 hPa only. Computed from 12h optimization time starting from all days of January of the Analysis. Eigenvalues in the y axis are multiplied by 10^{-10}.

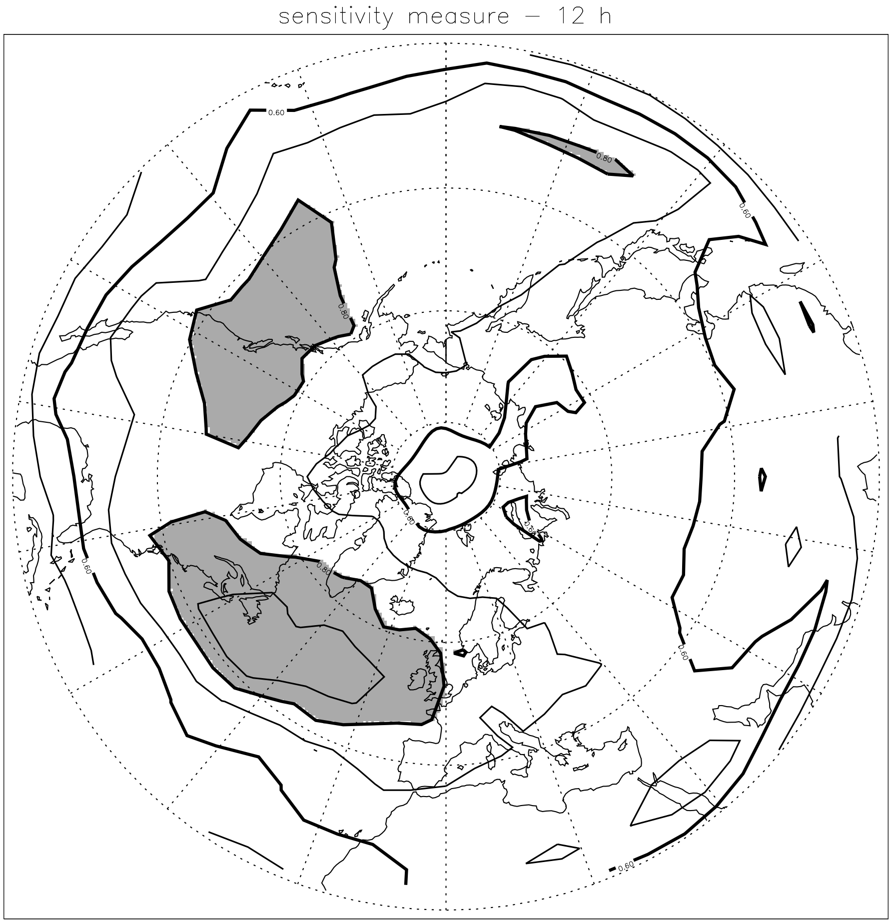

Figure 10: Average of first five singular vectors of the linear tangent model at 500 hPa. Computed from 12h optimization time starting from all days of January on the Analysis. The vectors are weighted by their corresponding singular value when averaging. The map is normalized by the *Sup* norm. Contour every 0.1, areas higher than 0.8 shaded.

North Atlantic with relative sensitivity higher than 0.9. No substantial change was found in the model performances with respect to Fig.3.

So, over the North Pacific region, the flow doesn't seem to contain the information necessary to build a correction to the model error. In the literature, this region has often been proven to be rather sensitive to tropical diabatic forcing, which can generate potential vorticity sources in midlatitudes (e.g. Hoskins and Karoly, 1981). It can be inferred that eddy flux convergence is not the prevailing explanation of forcing error over the Pacific. A way to parameterize forcing errors in this area should probably include some representation of tropical convection and of the divergent flow.

5.2 *Sensitivity to tuning parameters*

As seen in section 3, three tuning parameters were adjusted in the formulation of the parameterization. The adjustment was done empirically, by conducting a number of experiments with different choices of parameters.

The time interval τ was set at $12h$ to be consistent with the fact that the optimal residuals are indeed optimized for the $12h$ forecast. Diminishing the interval has no effect on the model performance, while increasing it progressively worsens the performance and brings it closer to the control integration. However, the dependence on this paramenter is rather weak. In the case of direct residuals better performance is found with $\tau = 1h$. Direct residuals are not "optimized" on a given period of forecast, and rather seem to show a time-local nature.

The number K of analogues has a direct influence on the model's variability, and only in a much weaker way on the climatology. To illustrate this point, the decomposition (2) of the tendency error should be recalled: $S = \bar{S} + S(q) + S'$, that is, a time constant part (the MM93 forcing), a part that is flow-dependent, that is given by the analogues, and a stochastic noise part, that represent the part of the error that is not driven by the flow. Each of the K residuals, corresponding to the K closest analogues, is a realization of the stochastic process $S(q) + S'$, which will have a probability distribution around the mean S(q). So, one has to look for the conditional probability distribution of S, given the flow.

A correct parameterization of the source term, consequently, should involve an estimation of the expectation of this distribution, plus a random term, whose covariance should also be estimated. These mean and covariance could be estimated naturally by using a large enough set of analogs and the corresponding residuals, but unfortunately the lack of data excludes this approach for not enough "good" analogues can be found. Conversely, due to the limited number of accurate analogues, the expectation of the mean of the distribution will be affected by an error. This error has the effect of adding an extra-variability to the mean, that in a sense takes the place of the random term S'. In other words the unknown stochastic term S' is substituted by the unknown random error on the expectation of $S(q)$.

The amplitude of the error on the expectation, that is to say the variance of the extra-variability, is controlled by the number K. In order to chose this amplitude, the empirical technique has been adopted to tune the number K in order to have the exact amplitude on the variance of the model *response*. In particular, the number K has been chosen in order to give the observed amplitude and location of the atlantic stormtrack at 500 and 200 hPa (compare Fig. 3 and 4 above). Best results were obtained for K=30 or K=40 after performing simulations varying K between 10 and 60. The whole diagnostics presented in this paper was also carried out for K=30 providing results similar to K=40.

Finally, the truncation T performed on the EOF space when computing distances must be taken into account. By performing an array of simulations, it was found that, with a value of T from 15 to 20, best result were obtained, and 15 was eventually chosen for keeping the dimensionality of phase space as small as possible. Apparently, the dependence of the forcing from the flow relies also on relatively smaller scales, and hence a relatively high truncation is required. Increasing too much the truncation, though, progressively degrades the quality of analogues, and consequently the model performance gets worse.

The average distance of the 40 analogues from a given point can be compared to the distribution of distances of couples of points in the space of T EOFs of the Euro Atlantic region. It has been tested that the average distance of the 40^{th} analogue falls in the 5^{th} percentile of the distribution of all distances, and the average distance plus a standard deviation falls in the 8^{th} percentile[3].

5.3 *Constant and transient forcing*

The empirical parameterization adds a time-varying component on top of the control constant forcing (5). Indicating by $\tilde{S}$ the control forcing, the instantaneous forcing term for the empirical forcing integration can be written:

$$S = \tilde{S} + S(q) \tag{10}$$

where $S(q)$ depends on time through $q = q(t)$. In this section some consideration will be made on the nature of $S(q)$ and its effect on the model flow by analysing the time series of $S(q)$, stored every $24h$ during the empirical forcing integration with optimal residuals.

A first consideration can be done by computing the time average $\overline{S(q)}$, which is found to be different from zero. In other words, the empirical parameterization has an effect on the constant (systematic) part of S also. The effect

[3] For constructing these distributions, only couple of points belonging to different years were considered.

of this systematic correction alone can be shown by performing an integration of the model with forcing $S = \tilde{S} + \overline{S(q)}$. The 500 hPa climatology of this "unbiased forcing" integration (not shown) already shows some improvement with respect to the control, but some failures are still present; in general, the area of the North Atlantic still has a higher error than in the empirical forcing integration while the rest of the Hemisphere already presents a reduction of the error that is comparable with it. As for the variability maps, on the other hand, no improvement is found, indicating that at least part of the lack of variability in the MM93 model's response is due to a lack of variability in the forcing itself. Weather regimes were also analysed in this integration. With respect to the control, an improvement is found especially in the Greenland and atlantic ridges. The zonal and blocked regimes, on the contrary, appear severely displaced to the west, which can again be explained by the underestimation of the stormtrack.

Consequently, a systematic correction of the forcing is not sufficient for reducing the error, especially in the North Atlantic. The transient part of the forcing turns out to have an essential importance for the correction of the flow in this region. The main failure of the constant "unbiased forcing" integration, seems to be linked to the severe underestimation of the atlantic stormtrack. The mechanism that relates the baroclinic wave activity along the stormtrack and the mean flow is the eddy flux convergence, that is at the core of the formation of the midlatitude jet. A higher transient eddy activity, triggered by the variable forcing at the source of the stormtrack can therefore be at the origin of the enhanced zonality of the flow in the empirical forcing integration with respect to the control one. A more zonal jet reduces the North Atlantic systematic error and moves the blocked and zonal regimes maxima towards the east to the correct position.

A further test can be conducted, similar to that of section 4.1, in order to see whether this higher wave activity is driven by the model flow through the empirical parameterization or by a stochastic effect independent of the flow. A "stochastic" integration is thus conducted with a random choice of M optimal residuals, but by rebiasing them to the average value $\tilde{S} + \overline{S(q)}$. More precisely, the "rebiased" instantaneous stochastic forcing term can be defined as $S = \tilde{S} + \overline{S(q)} + S'$, where S' is the average of M random residuals as in section 4.1.

The 500 hPa systematic error for this "unbiased stochastic" integration (not shown) is quite different from the preceding ones but the highest anomaly is still in the North Atlantic and is stronger. Consequently, a stochastic variable forcing, yet on top of an unbiased constant forcing, is not sufficient for triggering the wave activity along the storm track with the right phase and amplitude so to establish a flux convergence mechanism. It can be concluded that it is really the flow-dependence of the parameterized forcing that makes it.

The flux convergence mechanism, nevertheless, may not be the only responsible of the climatology improvement. Zorita and Von Storch (1998) used an analogues technique to build a downscaling from the large scale Euro-Atlantic sea level pressure fields to local precipitation in a region of Europe. The large scale flow may therefore bear some information on local sources of potential vorticity due to heat fluxes (that are linked to precipitations).

6 Conclusions and perspectives

6.1 *Summary*

In this paper, a method is proposed to estimate tendency errors in a simple GCM, and then a flow-dependent empirical parameterization is constructed in order to correct their effect during the integration. In this way, an improvement of the model climatology is obtained.

Two ways of calculating model errors, or model residuals, are presented; a simple direct calculation by finite differencing and a more objective one. This new method consists in a re-assimilation of an analysed dataset of observations in the GCM, in order to estimate model parameters. The assimilation scheme is a reformulation of a four dimensional variational assimilation scheme, applied to model parameters (the model forcing in this case) rather than to initial model state.

Then, an empirical parameterization has been formulated in order to parameterize model errors. The parameterization consists in constructing a transfer function between the model flow and the model residual. This function is estimated throughout the integration by an analogue method, with focus on the Euro-Atlantic region.

Applying such a parameterization gives very encouraging results. The model systematic error is substantially reduced and its variability is more realistic. Tests were conducted on one side to check the statistical stability of the results, and on the other to confirm that this improvement actually stems from the flow dependence of the model error.

Such flow dependence was found in the Euro-Atlantic sector, while similar attempts to establish this relation in other sectors of the globe gave no results. In particular, two other regions of high sensitivity to forcing have been found using singular vectors of model forcing. These regions are the west coast of North America and the pacific stormtrack. It seems that the flow in regions other than the North Atlantic does not contain the information necessary to correct the error.

A further analysis demonstrated that the empirical model error correction is found to have an effect both on the systematic part of the model error and on the transient part. The transient part is found to have a specific importance in reducing the climatology error in the North Atlantic region, by enhancing the high frequency wave activity on the atlantic stormtrack, which results in a strengthening of the jet by flux convergence.

These findings confirm that the low frequency variability in the Euro-Atlantic is to a certain extent sustained by internal dynamics and relies less on the boundary conditions than in other regions of the Northern Hemisphere. This corroborates the work of Shutts (1983), and later Vautard and Legras (1988), who showed that low frequency variability can be sustained in a simple channel model by the sole influence of transients.

6.2 *Potential applications and developments*

Both objectives of this paper, the estimation of tendency errors of the model, and the empirical parameterization of their effect, bear interesting potential applications.

Tendency errors can be used in the case of more complex models, for diagnosing deficiencies and consequently improving or objectively tuning physical parameterizations. As discussed by Klinker and Sardeshmukh (1992), analysing local tendency errors allows the identification of model deficiencies (that end up creating the systematic error) in a situation in which the error on every model variable is decoupled, allowing for straightforward interpretation.

Also the possibility of constructing an empirical correction to the model in order to improve its climatology gives good hopes of applications. One example is the field of climate modelling; in this case the use of analogues obviously limitates to present climate studies and cannot be directly applied to climate change projections.

Another natural field of potential application is seasonal forecasting, in which case a correct model climatology is particularly important. Moreover, other empirical correction terms could be studied in this case, for example tendency erors in the parametrization of deep convection and its dependency on tropical SSTs.

But if on one side these methods could be applied to full-fledged climate models, on the other they can be interpreted as a step towards the construction of simplified models of the large scale flow.

It has been shown in literature, that the extratropical dynamics of the large scale flow of the atmosphere at a given level (typically 500 hPa) is relatively low dimensional. By different statistical methods, authors such as Lorenz (1969), Wallace et al. (1991) or Fraederich et al. (1995) estimated the spatial degrees of freedom of the wintertime Northern Hemisphere 500 hPa geopotential height as being comprised somewhere between 20 and 50. Consequently, it should in principle be possible to construct a model of the large scale flow with a number of independent variables equal to the number of degrees of freedom. All other variables would be slaved to these few free modes.

The parameterization of forcing establishes a link between prognostic variables of the model and forcing, which in turn is meant to represent all the non resolved scales and phenomena, as well as other errors in the model formulation. The key to establishing such link is the use of an empirical, statistical, approach.

Hopefully, the statistical approach may not only be able to relate slaved and independent variables, but also may be the key to choose the optimal independent variables in the first place.

In literature, troughout the years, some authors has proposed the use of EOF of observed fields as an orthogonal basis for projecting model equations, as opposed to more classic spherical harmonics. Truncating the phase space of the model to a subspace of the first EOFs, Rinne and Karhila (1975), Schubert (1984) or Selten (1995), could with success reduce the dimensionality of simple barotropic models maintaining almost unchanged their dynamics. The EOFs, which are the modes that explain the greatest part of the variance of the model, could consequently also explain the most important part of its dynamics. Successive attempts (e.g. Selten 1997) to apply such basis to a baroclinic model revealed somewhat less interesting results. There are basically two key problems in these kind of studies: the first is the choice of the norm under which the EOFs are orthogonal, which can be chosen to have different physical meanings. The second is the formulation of a closure, that is to say, to express the effect of the modes neglected by the truncation in terms of the resolved ones.

The present paper represents the experimentation of such a closure in a spectral model. A natural future development will be to apply this closure to a model projected on some empirical basis in order to reduce its dimensionality.

Acknowledgements: Among the people that contributed to the article by discussion and exchange of ideas, the authors wish to thank Michel Déqué, Alain Joly, Eigil Kaas, François Lott, Franco Molteni and David Stephenson. The work was partially supported by the E.C. project POTENTIALS (ENV4-CT97-0497) and by the E.C. TMR grant ERBFMBI CT 96 1152.

A Variational estimation of forcing errors

The algorithm for estimating the forcing errors in the model is a special case of variational assimilation applied to the estimation of model parameters. The somewhat simplified problem solved in this paper is presented here, a more general formulation of the subject can be found, for example, in Stauffer and Bao (1993).

In discrete-time and in integral form, the model can be expressed as

$$q_{n+1} = M(q_n) + S \tag{A.1}$$

where q_n is the model variable (potential vorticity) at time n and S is the forcing. In the present application, one seeks the optimal forcing S which, over a certain number of timesteps N of (A.1) provides a forecast q_N that is closest to the corresponding analysis $\hat{q}_N$. The cost function $J(S) = \frac{1}{2}\|q_N - \hat{q}_N\|^2$, defined for a chosen norm, is therefore to be minimized. Most minimization algorithm (among which the quasi-Newton one used in this paper) require the knowledge of the gradient of the cost function at given points. This gradient, with respect to a change in the forcing, is defined by the relation:

$$\delta J = < \Delta_S J, \delta S > \tag{A.2}$$

Where $< \cdot, \cdot >$ is the scalar product associated to the chosen norm, δS is any perturbation of the forcing and δJ the resulting perturbation of the cost function.

Linearizing (A.1) around the full model trajectory starting from q_0, the linear tangent (LT) model can be defined:

$$\delta q_{n+1} = A_n \delta q_n + \delta S \tag{A.3}$$

where A_n is the Jacobian matrix $\frac{\partial M}{\partial q}(q_n)$.

It should be noted, that if $\delta q_0 = 0$, or no perturbation is assumed on the initial condition, the value of δq_{n+1} only depends from the variation of forcing δS through a linear operator $\delta q_{n+1} = L_n \delta S$, which can be expressed in terms of the jacobian A_n. Applying (A.3) explicitly yields:

$$\begin{aligned}
&\delta q_0 = 0 \\
&\delta q_1 = \delta S \\
&\delta q_2 = A_1 \delta S + \delta S \\
&\delta q_3 = A_2 A_1 \delta S + A_2 \delta S + \delta S \\
&\vdots \\
&\vdots \\
&\delta q_n = A_{n-1} \ldots A_1 \delta S + A_{n-1} \ldots A_2 \delta S + \;\ldots.. \; + A_{n-1} \delta S + \delta S.
\end{aligned}$$

The last row of the above formula gives a definition of the operator L_n like this:

$$\begin{aligned} L_n &= A_{n-1} \,\ldots\ldots\ldots\, A_1 \\ &+ A_{n-1} \,\ldots\ldots\ldots\, A_2 \\ &+ A_{n-1} \,\ldots\ldots\ldots\, A_3 \\ &\vdots \\ &\vdots \\ &+ A_{n-1} \\ &+ I. \end{aligned} \tag{A.4}$$

At this point, the differentiation of the cost function gives

$$\delta J =< q_N - \hat{q}_N, \delta q_N >=< q_N - \hat{q}_N, L_N \delta S >=$$

$$=< L_N^*(q_N - \hat{q}_N), \delta S >, \tag{A.5}$$

where L_N^* stands for the adjoint of L_N. Comparing (A.5) with (A.2) a useful expression for the gradient of the cost function is finally obtained.

$$\Delta_S J = L_N^*(q_N - \hat{q}_N). \tag{A.6}$$

The expression of L_N^* derives trivially from (A.4)

$$\begin{aligned} L_N^* &= A_1^* \,\ldots\ldots\ldots\, A_{N-1}^* \\ &+ A_2^* \,\ldots\ldots\ldots\, A_{N-1}^* \\ &+ A_3^* \,\ldots\ldots\ldots\, A_{N-1}^* \\ &\vdots \\ &\vdots \\ &+ A_{N-1}^* \\ &+ I. \end{aligned} \tag{A.7}$$

Given the form of the operator L_N^*, the r.h.s term of (A.6) can be cheaply computed by one only integration backward of the following system:

$$\begin{aligned} &z_N = (q_N - \hat{q}_N) \\ &u_N = 0 \\ & \\ &z_{n-1} = A_{n-1}^* z_n \\ &u_{n-1} = z_n + u_n \end{aligned} \tag{A.8}$$

where basically the value of z is cumulated at every step in u. In this way (A.6) becomes simply

$$\Delta_S J = u_0$$

In summary, the computation of gradient of the cost function at time t=0 requires: (i) one integration of the full model from t=0 to t=N, in order to evaluate the trajectory q_n, $n \in [1, N]$ that is necessary to estimate A_n^* at every n; and (ii) a single backward integration of the adjoint model starting from the initial condition $\delta q_N = (q_N - \hat{q}_N)$.

References

Buizza, R., T. N. Palmer, 1995: The Singular Vector Structure of the Atmospheric Circulation. *J. Atmos. Sci.*, **52**, 1435-1456.

D'Andrea, F. and Coauthors, 1998: Northern Hemisphere Atmospheric Blocking as simulated by 15 General Circulation Models in the period 1979-1988. *Clim. Dyn.*, **14**, 385-407.

Fraederich, K. C., C. Ziehmann and F. Sielmann, 1995: Estimate of Spatial Degrees of Freedom. *J. Climate*, **8**, 361-369.

Gates, W. L., and Coauthors, 1998: An Overview of the Results of the Atmospheric Model Intercomparison Project (AMIP). *PCMDI Report* No. 45, Lawrence Livermore National Laboratory, Livermore, CA, USA.

Hoskins, B. J. and D. J. Karoly, 1981: The steady linear response of a spherical atmosphere to thermal and orographic forcing *J. Atmos. Sci.*, **38**, 1179-1196.

Gilbert, J. C. and C. Lemarechal, 1992. The modules M1QN2, N1QN2, M1QN3 and N1QN3. *Technical report*, available from INRIA (Institut National de Recherches an Informatique et Automatique)

Kaas E., S. Yang, M. Déqué, F. D'Andrea, I. Kirchner and B. Machenhauer, 1998b: Project On Tendency Evaluations using New Techniques to Improve Atmospheric Long-term Simulations (POTENTIALS). Proceedings of the European Climate Science Conference, Vienna, 19-23 October 1998. (Available from the Danish Meteorological Center).

Kaas, E., A. Guldberg, W. May and M. Déqué, 1999: Using Tendency Errors to tune parameterization of unresolved dynamical scale interactions in atmospheric General Circulation Models. *Tellus*, **51A**.

Klinker, E. and P. D. Sardeshmukh, 1992: The diagnosis of Mechanical Dissipation in the Atmosphere from large-Scale Balance Requirements. *J. Atmos. Sci.*, **49** 608-627.

LeDimet, F.X. and O. Talagrand, 1986: Variational algorithms for analysis and assimilation of observations of meteorological observations: Theoretical aspects. *Tellus*, **38A** 97-110.

Lorenz, E.N., 1969: Atmospheric predictability as revealed by naturally occurring analogues. *Tellus*, **26** 636-646.

Marshall, J. and F. Molteni, 1993: Towards a Dynamical Understanding of Planetary-Scale Flow Regimes. *J. Atmos. Sci.*, **50**, 1792-1818.

Michelangeli, P.A., 1996: *Variabilité atmosphérique basse-fréquence observée et simulée aux latitudes tempérées.* (In French). PhD Thesis, University of Paris VI.

Michelangeli, P.A., Vautard, R. and Legras, B., 1995: Weather regimes: recurrence and quasi-stationarity. *J. Atmos. Sci.*, **52**, 1237-1256.

Rinne, J. and V. Karhila, 1975: A spectral barotropic model in horizontal empirical orthogonal functions. *Q.J.R.M.S.* **101** 365-382.

Schubert, S.D., 1985: A statistical-dynamical study of empirically determined modes of atmospheric variability. *J. Atmos. Sci.*, **42**, 3-17.

Selten, F.M., 1995: An efficient description of the dynamics of barotropic flow. *J. Atmos. Sci.*, **52**, 915-936.

Selten, F.M., 1997: Baroclinic Empirical Orthogonal Functions as basisfunctions in an atmosperic model. *J. Atmos. Sci.*, **54**, 2100-2114.

Shubert, S. and Y. Chang, 1995: An Objective method fo inferring Sources of Model error. *Mon. Wea. Rev.*, **124** 325-340.

Shutts, G.J., 1983: Propagation of eddies in diffluent jet streams: eddy vorticity forcing of blocking flow fields. *Q.J.R.M.S.* **109(462)** 737-761.

Stauffer, D. R. and J.-W. Bao, 1993, Optimal determination of nudging coefficients using the adjoint equations. *Tellus*, **45A** 358-369.

Van den Dool, H. M., 1994: Searching for analogues, how long must we wait? *Tellus*, **46A**, 314-324.

Vautard R. and B. Legras, 1988: On the sources of midlatitude low-frequency variability. Part II: Nonlinear equilibration of weather regimes. *J. Atmos. Sci.*, **45**, 2845-2867.

Wallace, J.M., X. Cheng and D. Sun, 1991: Does low-frequency variability exhibit regime-like behavior? *Tellus*, **43AB** 16-26.

Zorita, E. and H. von Storch, 1998, The analog method as a simple statistical downscaling technique: comparison with more complicated methods. *J. Clim.*, in press.