TECHNICAL MEMORANDUM

473

Influence of stochastic physics on the frequency of occurrence of North Pacific weather regimes in the ECMWF model

T. Jung, T.N. Palmer and G.J. Shutts[1]

Research Department

[1]The MetOffice, Exeter, UK

Submitted to Geophysical Research Letters

July 2005

European Centre for Medium-Range Weather Forecasts
Europäisches Zentrum für mittelfristige Wettervorhersage
Centre européen pour les prévisions météorologiques à moyen terme

Series: ECMWF Technical Memoranda

A full list of ECMWF Publications can be found on our web site under:
http://www.ecmwf.int/publications/

Contact: library@ecmwf.int

European Centre for Medium-Range Weather Forecasts
Shinfield Park, Reading, RG2 9AX, England

Abstract

One common problem of many atmospheric circulation models is that the strength of the mean westerly winds is overestimated in the mid-latitude North Pacific. This westerly wind bias is also a prominent feature in a recent version of the ECMWF model. Here we use the ECMWF model in order to investigate whether the use of stochastic physics helps reducing this error using the concept of weather regimes. The focus is on the winter season when the atmospheric regime structure is most pronounced. It is shown that the operational version of the ECMWF stochastic physics has little impact on the frequency of occurrence of North Pacific weather regimes. A recently developed scheme, however, which is based on combining a cellular automaton with a stochastic backscatter component, leads to substantial improvements in the simulation of the frequency of occurrence of North Pacific weather regimes.

1 Introduction

During the last 50 years numerical models of the atmosphere became one of the most valuable tools in the atmospheric sciences and in climate research. Nowadays they are routinely used to carry out short-range and medium-range weather forecasts and, by coupling to other components of the climate system, they are used for monthly and seasonal forecasting as well as for assessing possible future climate change under increasing greenhouse gas concentrations. Despite substantial model improvements in recent decades even state-of-the-art models still have systematic errors all the way from the short-range into the extended-range (Jung, 2005; Jung et al., 2005). Another peculiarity is that different atmospheric models tend to show similar systematic errors as pointed out by Gates et al. (1999), even though different model formulations are used.

One possible explanation for the persistence of systematic model error has been provided by Palmer (2001), who pointed out that one source of forecast error might be the way the governing equations are approximated, that is, the equations of motion are truncated at some prescribed scale and the influence of unresolved scales is modeled by a set of deterministic bulk formulae. Palmer (2001) suggests that the effect of unresolved processes should be represented by relatively simple stochastic-dynamic systems coupled to the resolved system. In this way a common model problem, that is, the underestimation of kinetic energy at the truncation level (Shutts, 2005) could be addressed.

One way to understand how the use of stochastic-dynamic systems could reduce systematic errors has been be outlined by Molteni and Tibaldi (1990). Their argument is based on the notion that the distribution of the extratropical flow is non-Gaussian or even multi-modal (see Kimoto and Ghil, 1993; Smyth et al., 1999, for observational evidence). Suppose a dynamical system whose distribution is bi-modal and which is driven by a stochastic forcing. Further suppose that one of the peaks of the probability density function is much higher than the other, which corresponds to a potential well structure with two minima of different depth. If the stochastic forcing is too weak than the more stable regime will be overpopulated which will give rise to systematic model error. If the stochastic forcing is increased the two regimes become more evenly populated and the mean climate of the model changes accordingly due to a change in the frequency of noise induced transitions (for more details, see Molteni and Tibaldi, 1990; Palmer, 2001). From the above discussion it becomes clear that weather regimes provide a promising framework for studying the impact that stochastic physics schemes can have on the mean climate of models.

2 Data and methods

In order to address the questions posed in the Introduction, the atmospheric model component of the European Centre for Medium-Range Forecasts (ECMWF) Integrated Forecast System (IFS) is used. Specifically, the

Table 1: Abbreviations along with some of the basic characteristics of the datasets used within this study.

Abbreviation	Cycle	Resolution	Stochastic Physics Scheme	Period
ERA40	23r4	T_L159L60	none	1962–2001
CNTL	26r3	T_L95L60	none	1962–2001
SSP	26r3	T_L95L60	Buizza et al. (1999)	1962–2001
CASB	26r3	T_L95L60	Shutts (2005)	1962–2001

numerical experimentation is based on model cycle 26r3, which was used operationally at ECMWF from 7 October 2003 to 8 March 2004. A horizontal resolution of T_L95 (linear Gaussian grid, $\approx 1.875^o$) is used and 60 levels in the vertical are employed. Observed sea surface temperature fields are used as lower boundary condition. The performance of earlier model cycles in simulating the observed climate is described elsewhere (Brankovic and Molteni, 2004; Jung, 2005; Jung et al., 2005).

In total, three experiments were conducted. The first experiment, the control integration (CNTL, herafter), is based on the original ECMWF model. In the second experiment the standard stochastic physics scheme (Buizza et al., 1999)(SSP, hereafter), used operationally at ECMWF in the ensemble prediction system, has been applied during the course of the integration. In the SSP scheme, the term εP is added to the model tendencies, where P denotes the parameterized tendency associated with sub-grid scale processes (e.g., convective heating) and ε is a random number drawn from a uniform distribution in the interval $[-0.5, 0.5]$. The same random numbers are used over a time range of 6h and a spatial domain of 10×10 latitude/longitude (for details, see Buizza et al., 1999). In the third experiment (CASB, hereafter), the cellular automaton/stochastic backscatter scheme by Shutts (2005) is used (see also Palmer et al., 2005). In the CASB scheme, the streamfunction tendency arising from unresolved or poorly resolved physical processes is proportional to the square root of the dissipation rate associated with numerical dissipation, mountain drag as well as deep convection; the spatial and temporal structure of streamfunction forcing is given by the cellular automaton.

For each of the above experiments 6-months long integrations were started on 1 October of each of the years 1962–2001, a total of 40 integrations for each experiment. The first two months of the integrations have been discarded in order to remove transient effects and to focus on the winter season (December through March). ERA-40 reanalysis data (Uppala et al., 2005) are used as “truth”. A summary of the experiments used in this study is given in Table 2.

The main method used to identify weather regimes in ERA-40 data and the three experiments is K-means clustering (Hartigan and Wong, 1979). The outcome of K-means clustering depends to some degree on the choice of initial clusters used. In order to reduce the sensitivity to initialization an ensemble of cluster analyses has been carried out, encompassing a total of 500 members. For further diagnosis we have chosen that member which is most similar to all other 499 members taking into account that the cluster ordering might be different for different members. Here similarity is measured in terms of root mean square differences (small values implying strong similarity). In this study we use $K = 3$ clusters, which has been found to be the best choice (J. Berner, personal communication) using the mixture model clustering technique described bySmyth et al. (1999).

K-means clustering has been applied to the leading ten non-normalized principal components (PCs) of North Pacific 500 hPa geopotential height (Z500) anomalies. By using non-normalized PCs (the PCs carry the units) it is ensured that the K-mean clustering algorithm gives more weight to the leading PCs. The North Pacific region encompasses the domain 20^o–80^oN and 100^oE–100^oW. Z500 anomalies were obtained as follows. First, non-overlapping ten-day averages have been computed and then the mean annual cycle (based on fitting a third-

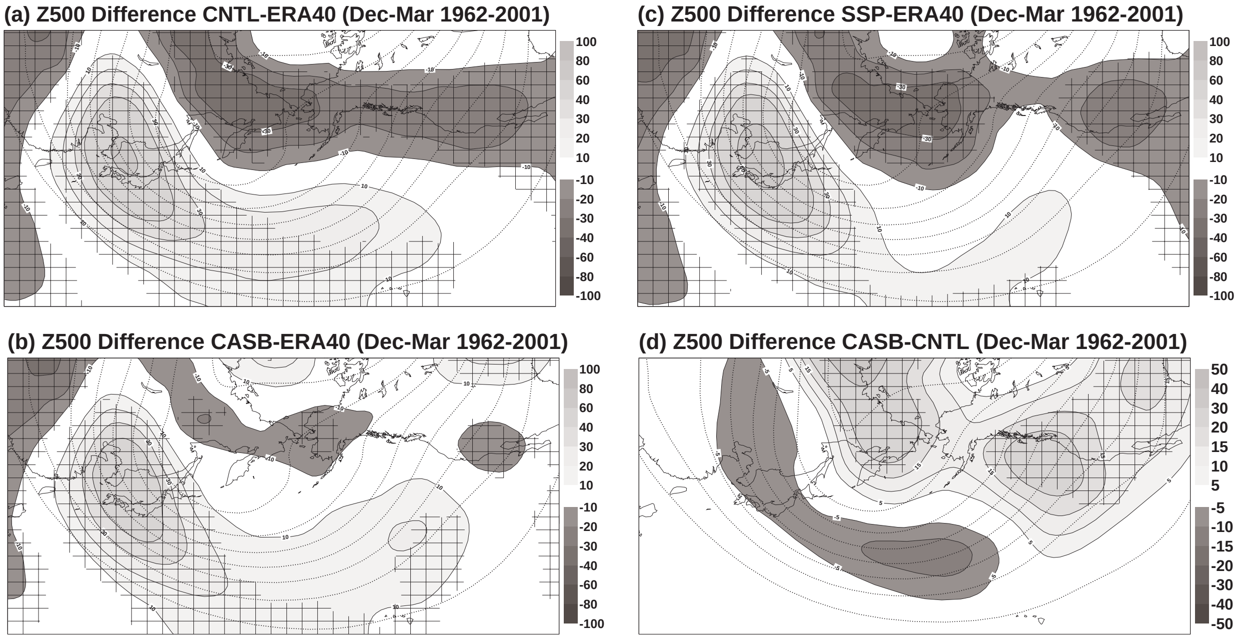

Figure 1: Mean differences of 500 hPa geopotential height fields (shading in m) for winters of the period 1962–2001: (a) CNTL−ERA40, (b) CASB−ERA40, (c) SSP−ERA40, and (d) CASB−CNTL. Statistically significant differences (at the 95% confidence level) are hatched

order polynomial to the data) has been removed. The mean annual cycle has been separately determined for each dataset. Notice that by using ten-day averages the total sample size is 480 for each of the datasets used in this study.

3 Results

To start with let us consider the impact the two stochastic physics schemes have on the mean wintertime circulation. Systematic Z500 errors for CNTL are shown in Fig. 1a. The abovementioned tendency of the model to produce too strong westerly winds (westerly wind bias) in the central North Pacific is clearly evident. The North Pacific westerly wind bias is significantly reduced in the CASB experiment (Fig. 1b,d). This is in contrast to the SSP experiment (Fig. 1c), which shows a very similar systematic error structure to CNTL throughout the Northern Hemisphere. Evidently the CASB scheme has a much stronger impact on the mean circulation than the SSP scheme, and this impact is clearly beneficial in terms of reducing the westerly wind bias.

The three clusters obtained from ERA-40 reanalysis data and the control integration are shown in Fig.2. The first and third cluster obtained from ERA-40 resemble the negative and positive phase of the Pacific North America (PNA) pattern, respectively. The second cluster closely resembles one of the regimes identified by Kimoto and Ghil (1993) and describes latitudinal shifts of the Aleutian low pressure system. A comparison between observed and simulated clusters reveals that the ECMWF model is capable of accurately simulating the spatial structure of the observed regimes. However, there are substantial differences in the observed and simulated frequencies of occurrence of the clusters. The least populated cluster (31.3%) in the ERA-40 data, for example, becomes the most populated cluster (40.8%) in the control integration. The second cluster, which is associated with reduced zonal flow in the northern North Pacific, on the other hand, is substantially underpopulated in the model. In summary the westerly wind bias of the ECMWF model can be explained in terms

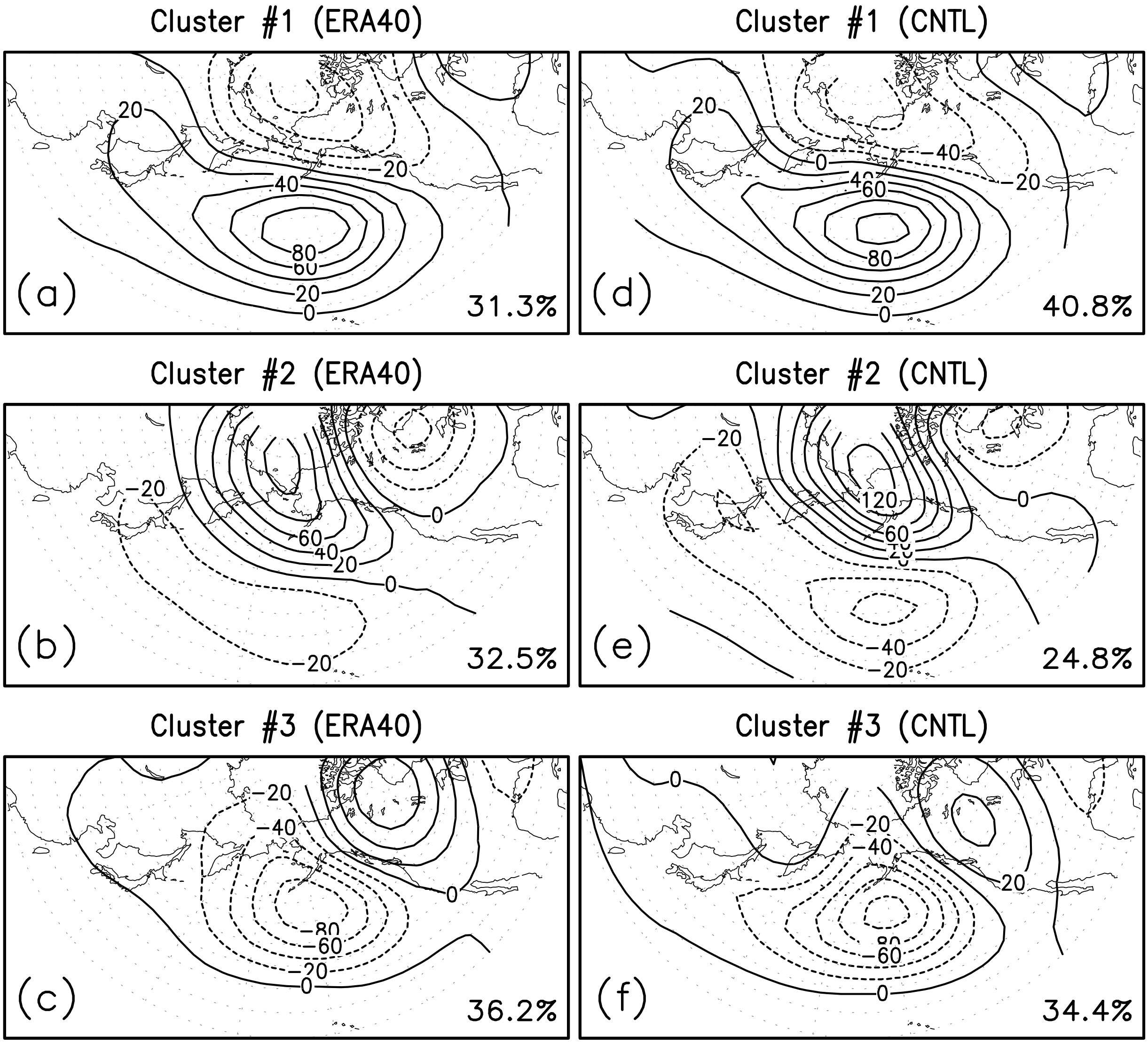

Figure 2: Three cluster centroids of wintertime Z500 anomalies (m): (a)–(c) ERA-40 reanalysis data and (d)–(f) control integration. The results are based on 40 seasonal integrations (1962–2001) of the ECMWF model (cycle 26r3). The percentage of days spent in each of the clusters is also given.

of deficits in accurately simulating the observed frequency of occurrence of North Pacific weather regimes.

The three regime centroids for the simulation with the SSP scheme is shown in Fig. 3a–c. As for CNTL, the spatial structure of the clusters is very similar to those obtained from ERA-40 reanalysis data. Moreover, the experiment with the SSP scheme shows similar deficits like CNTL in terms of simulating the observed regime frequencies. In fact the regimes frequencies are between the two experiments are hardly distinguishable.

Simulated North Pacific weather regimes in the experiment employing the CASB scheme (Fig. 3d–f) show the same spatial structure as for the other datasets. However, the simulated frequencies of occurrence of the individual regimes are much more realistic closely resembling those for the ERA-40 reanalysis data.

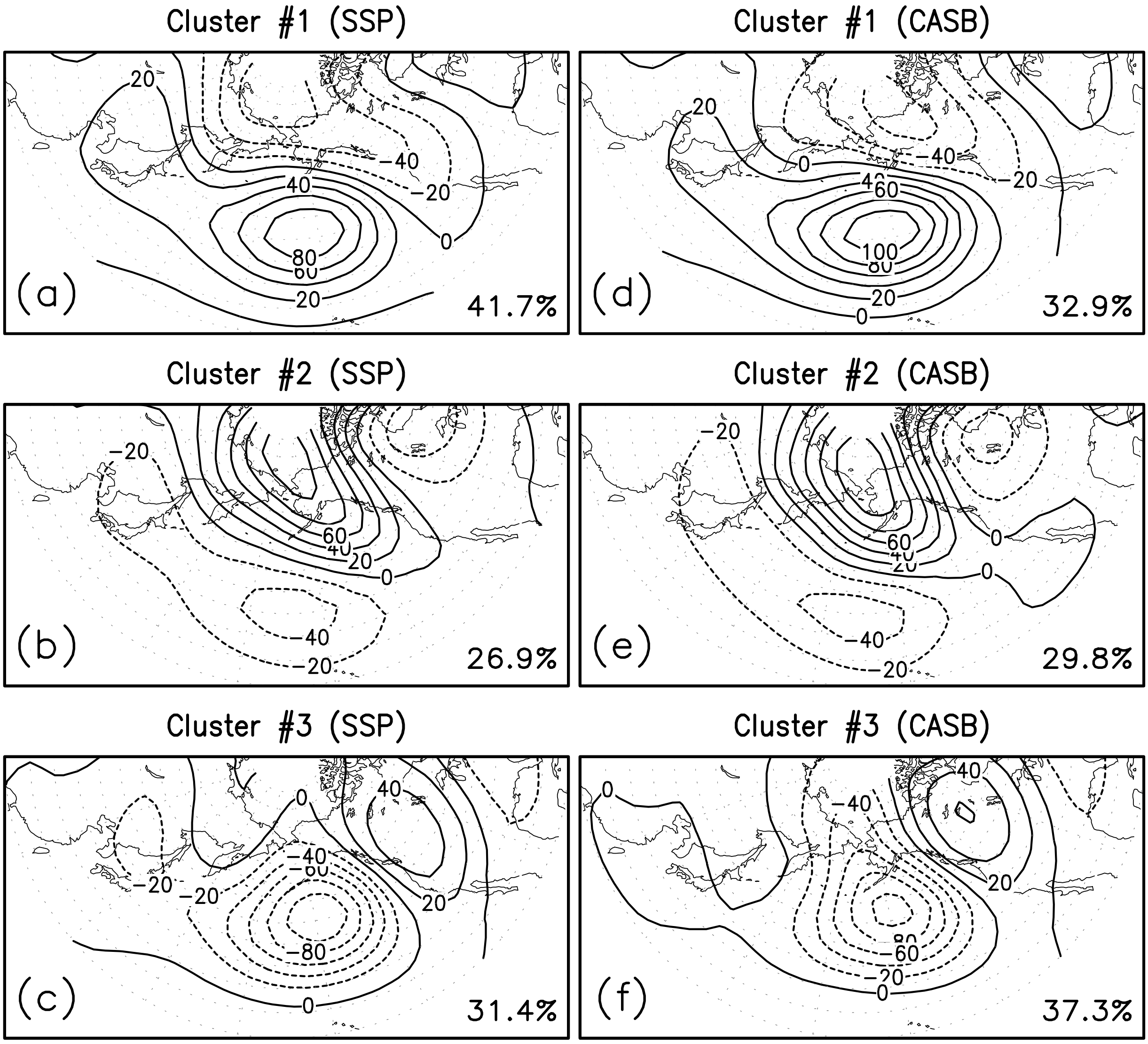

Figure 3: As in Fig. 2, except for (a)–(c) the experiment with the standard stochastic physics scheme (SSP) and (d)–(f) the new stochastics physics scheme (CASB).

4 Summary and discussion

The performance of the ECMWF model in simulating the observed North Pacific regime structure has been investigated. The control integration reveals realistic regimes structures; the frequency of occurrence of the regimes, however, has been found to be considerably misrepresented compared to those obtained from ERA-40 reanalysis data. Sensitivity experiments have been carried out in order to answer the question whether the use of stochastic physics schemes is beneficial in reducing the model problems. It was found that the SSP scheme has no beneficial impact. A recently developed scheme (CASB), however, improves the simulated regime frequencies dramatically.

We believe that the CASB scheme is more efficent because it forces the streamfunction field directly (i.e., the convective dissipation is used to force momentum fields). This conjecture is further supported by the relatively strong impact that the CASB scheme has on other fields such as precipitation (not shown). Furthermore, our results are supported by the study of Palmer et al. (2005) in which it is shown that the use of the CASB scheme significantly reduces the underestimation of the frequency of occurrence of simulated North Pacific blocking events (reduced westerly wind bias).

As mentioned in the Introduction the conceptual model of Molteni and Tibaldi (1990) predicts that a lack of stochastic forcing leads to an overpopulation (underpopulation) of the more (less) stable regimes. Increasing the level of stochastic forcing leads to a more evenly population of weather regimes. It is worth pointing out that this is exactly what is happening in the experiment in which the recently developed cellular automaton/stochastic backscatter scheme has been used to increase the kinetic energy of the model at the truncation level.

Acknowledgements The authors are grateful to Judith Berner for useful discussions.

References

Brankovic, C. and F. Molteni, 2004: Seasonal climate and variability in the ECMWF ERA-40 model. *Clim. Dyn.*, **22**, 139–155.

Buizza, R., M. Miller, and T. N. Palmer, 1999: Stochastic representation of model uncertainities in the ECWMF Ensemble Prediction System. *Quart. J. Roy. Meteor. Soc.*, **125**, 2887–2908.

Gates, W. L., J. Boyle, C. Covey, C. Dease, C. Doutriaux, R. Drach, M. Fiorino, P. Gleckler, J. Hnilo, S. Marlais, T. Phillips, G. Potter, B. Santer, K. Sperber, K. Taylor, and D. Williams, 1999: An overview of the results of the Atmospheric Model Intercomparison Project (AMIP I). *Bull. Amer. Meteor. Soc.*, **80**, 29–55.

Hartigan, J. A. and M. A. Wong, 1979: Algorithm AS136: A K-means clustering algorithm. *Appl. Statist.*, **28**, 100–108.

Jung, T., 2005: Systematic errors of the atmospheric circulation in the ECMWF forecasting system. *Quart. J. Roy. Meteor. Soc.*, **131**, 1045–1073.

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

Kimoto, M. and M. Ghil, 1993: Multiple flow regimes in the Northern Hemisphere winter. Part II: Sectorial regimes and preferred transitions. *J. Atmos. Sci.*, **50**, 2645–2673.

Molteni, F. and S. Tibaldi, 1990: Regimes in the wintertime circulation over the northern extratropics. II: Consequences for dynamical predictability. *Quart. J. Roy. Meteor. Soc.*, **116**, 1263–1288.

Palmer, T. N., 2001: A nonlinear dynamical perspective on model error: A proposal for non-local stochastic-dynamic parameterization in weather and climate prediction models. *Quart. J. Roy. Meteor. Soc.*, **127**, 279–304.

Palmer, T. N., G. J. Shutts, R. Hagedorn, F. Doblas-Reyes, T. Jung, and M. Leutbecher, 2005: Representing model uncertainty in weather and climate prediction. *Ann. Rev. Earth Planet. Sci.*, **33**, 163–193.

Shutts, G. J., 2005: A kinetic energy backscatter algorithm for use in ensemble prediction systems. *Quart. J. Roy. Meteor. Soc.*. Accepted.

Smyth, T. M., K. Ide, and M. Ghil, 1999: Multiple regimes in Northern Hemisphere height fields via mixture model clustering. *J. Atmos. Sci.*, **56**, 3704–3723.

Uppala, S., P. W. Kallberg, A. J. Simmons, U. Andrae, V. Da Costa Bechtold, M. Fiorino, J. K. Gibson, J. Haseler, A. Hernandez, G. A. Kelly, X. Li, K. Onogi, S. Saarinen, N. Sokka, R. P. Allan, E. Andersson, K. Arpe, M. A. Balmaseda, A. C. M. Beljaars, L. van de Berg, J. Bidlot, N. Bormann, S. Caires, F. Chevallier, A. Dethof, M. Dragosavac, M. Fisher, M. Fuentes, S. Hagemann, E. Holm, B. J. Hoskins, L. Isaksen, P. A. E. M. Janssen, R. Jenne, A. P. McNally, J.-F. Mahfouf, J.-J. Morcrette, N. A. Rayner, R. W. Saunders, P. Simon, A. Sterl, K. E. Trenberth, A. Untch, D. Vasiljevic, P. Viterbo, and J. Woollen, 2005: The ERA-40 reanalysis. *Quart. J. Roy. Meteor. Soc.*. Submitted.