Tellus (1995), 47A, 495–501
Printed in Belgium – all rights reserved

TELLUS
ISSN 0280–6495

Approximating dominant eigenvalues and eigenvectors of the local forecast error covariance matrix

By JAN BARKMEIJER, *Royal Netherlands Meteorological Institute, POBox 201, 3730 AE De Bilt, The Netherlands*

(Manuscript received 6 April 1994; in final form 12 September 1994)

ABSTRACT

Examining the dominant eigenvectors of a forecast error covariance matrix for Western Europe during a 607-day period, shows that these daily changing vectors remain in a low-dimensional space. The first few dominant eigenvectors of each day can almost completely be described by a fixed basis consisting of a relatively small number of elements. A simple method is presented that utilizes this property to determine the daily dominant eigenvectors and eigenvalues of the covariance matrix in an efficient manner. Results are given for a 2-day forecast period, but apply also for a forecast period of 3 days. Use of the method, instead of the Lanczos algorithm, in approximating the seven largest eigenvalues within a 1% accuracy level, resulted in a 30% reduction of the computational costs.

1. Introduction

The quality of forecasts has improved considerably during the last decades (Arpe, 1990; Kalnay et al., 1990). Still, the day-to-day variability is quite large. One particular predicted atmospheric circulation pattern must be considered as an arbitrary member of a probability distribution of circulation patterns. Ensemble prediction, in which the forecasting model is integrated from different, but equally likely, initial conditions, has been proposed for estimating this distribution, see, e.g., Leith (1974). When the distribution is narrow, the forecast must be accurate.

Information about the likely structure of the error patterns in the forecast may be quite useful for forecasters. These error patterns can, among others, be used to generate alternative forecasts. In this way one can get, for instance, information about the uncertainty in the position and intensity of circulation structures.

One way to determine these error patterns is to compute the covariance matrix of the forecast error. Error patterns are given by the eigenvectors of this symmetric matrix. The corresponding eigenvalue is related to the expectation value for the amplitude of that error pattern. Besides the difficulty of choosing a proper analysis error covariance matrix, the dimension of this matrix is generally too large to handle. Use of an adjoint model and a small target area for which the quality of the forecast is studied, actually makes the computation of a geographically local covariance matrix of the forecast error possible. In Barkmeijer et al. (1993), we presented a skill prediction method along these lines. As a predictor of the quality of forecasts we used the sum of the variances of the predicted error patterns in the target area. This predictor is given by the trace of the local covariance matrix or, equivalently, by the sum of its eigenvalues. The computation of this matrix was performed with a 3-level quasi-geostrophic model (T21QG) triangularly truncated at wavenumber 21 (Marshall and Molteni, 1993). The tangent linear and adjoint T21QG model were integrated along an orbit that closely followed the operational European Centre for Medium-Range Weather Forecasts (ECMWF) forecast orbit. It was found that skill forecasts starting from simple diagonal covariance matrices of the analysis error already performed well. Use of a feasible method by Bouttier (1993) to transform a diagonal matrix into a more realistic covariance matrix, may improve the usefulness of these skill forecasts even more.

Since the numerical costs of running such models are low, we were able to perform a large skill experiment involving 607 ECMWF forecasts with a forecast period up to 72 h. As target area we selected Western Europe. For later reference we refer to this experiment as the 607-day experiment. The number of orthogonal error patterns one can compute for each day equals the number of grid-points that is considered in describing the target area. For the 607-day experiment this number is 23. It appears that the seven most dominant eigenvectors already describe over 90% of the variance of the ECMWF forecast error for the target area. This implies that it is not necessary to determine the complete eigenstructure of the local covariance matrix. A method that yields the dominant eigenvectors in a more efficient manner already will suffice. A Lanczos algorithm can be used for this purpose, see, e.g., Buizza et al. (1993). In an iterative procedure applied to the operator, which corresponds with the covariance matrix of the local forecast error, this algorithm yields approximate eigenvalues and eigenvectors. The advantage of this approach is that it is not necessary to know the explicit form of the covariance matrix. The eigenvalues, thus obtained, converge to the eigenvalues of te covariance matrix in a monotonous way. For a more complete description of the algorithm we refer to Parlett (1980). In the 607-day experiment, at least 12 iterations are needed for the Lanczos algorithm in order to obtain 7 eigenvalues within a 1% accuracy level. Each iteration involves an integration of the equally expensive tangent linear and adjoint model. Given this computation time, one might as well determine the explicit form of the local covariance matrix, which takes 23 integrations of the adjoint model backwards in time.

In this study we present an alternative method to compute the dominant eigenvalues and eigenvectors of local covariance matrices. Important in this approach is that the space spanned by the dominant eigenvectors does not vary much in time. For the 607-day experiment this condition is satisfied. The first few dominant eigenvectors for the Western European area seem to rotate in a low-dimensional space during the 607-day period. A fixed basis, consisting of a relatively small number of elements, suffices to describe the first few dominant eigenvectors of all days to a large extent. The computational cost of determining the dominant eigenvalues and eigenvectors of the daily covariance matrix can then be reduced substantially. Use of the method in approximating the seven largest eigenvalues within a 1% accuracy level may result in a 30% decrease in computational costs when compared to the Lanczos algorithm.

In Section 2 we briefly recall how the local covariance matrix can be obtained. In Section 3 we investigate how the eigenvectors evolve in time and the alternative method to determine the eigenvalues and eigenvectors of the covariance matrix is introduced in Section 4. In Section 5 some numerical results are presented and finally in Section 6 we give some concluding remarks.

2. Local covariance matrix Q_{loc}

In defining the covariance matrix of the regional forecast error, we omit the contribution of model errors to the forecast error. We further assume that the error dynamics is linear. Errico et al. (1993) investigated the validity of this assumption for a mesoscale primitive equation model (the PSU/NCAR MM4) during two synoptic situations. They found that for perturbations magnitudes as large as typical current analysis errors, error growth is linear for forecast periods up to 72 h. Experiments performed by Lacarra and Talagrand (1988) with an f-plane shallow-water model show linear error growth up to 48 h for realistic perturbations.

We first explain how the local covariance matrix can be computed, assuming linear error growth, c.f., Barkmeijer et al. (1993). The evolution of an initial error $\varepsilon(0)$ to time T of the forecast is determined by integrating the tangent linear T21QG model. This defines a linear operator $R(0, T)$ such that

$$\varepsilon(T) = R(0, T)\,\varepsilon(0). \tag{1}$$

In the sequel we leave the time dependence of R, the propagator of the tangent linear model, from the notation. We point out that R also depends on the daily varying forecast orbit. The structure of the initial error is restricted by prescribing the analysis error covariance matrix $Q=\overline{\varepsilon(0)\,\varepsilon(0)^{t}}$. The superscript t states that the transpose has been taken and the bar denotes the ensemble mean of the set of possible initial errors. Note that the symmetric matrix Q can be written as $Q=BB^{t}$. The columns of B are the eigenvectors of Q multiplied with the square root of the corresponding eigenvalues. An ensemble member $\varepsilon(0)$ can be generated by putting $\varepsilon(0)=Bd$, where the elements d_i of d are $N(0, 1)$ distributed. In determining the covariance matrix Q_{loc} of the local forecast error, we only need to know the forecast error in the target area. This is handled by introducing a projection operator P. The operator P acting on $\varepsilon(T)$ may project onto any quantity of interest in the target area, e.g., geopotential height at 500 hPa as in the 607-day experiment. In the latter case, the elements of the vector $P\varepsilon(T)$ equal the geopotential height in each gridpoint that is considered. The covariance matrix Q_{loc} is given by

$$\begin{aligned} Q_{\text{loc}} &= \overline{P\varepsilon(T)(P\varepsilon(T))^{t}} \\ &= \overline{PR\varepsilon(0)(PR\varepsilon(0))^{t}} \\ &= PRQR^{t}P^{t} \\ &= PRBB^{t}R^{t}P^{t}. \end{aligned} \tag{2}$$

For a sufficiently small target area, consisting of say n gridpoints, the matrix Q_{loc} can be computed explicitly. Since $PRB=(B^{t}R^{t}P^{t})^{t}$, it suffices to determine $B^{t}R^{t}P^{t}g_i$ for $i=1, \ldots, n$. Here the vectors g_i, $i=1, \ldots, n$, form a basis for the n-dimensional space of patterns in the target area. The adjoint T21QG model enables us to compute the action of R^{t}. In case of the Euclidean inner product $\langle x, y\rangle=\sum x_i y_i$ we can write $R(0, T)^{t}=R^{*}(T, 0)$. The linear operator R^{*} is the adjoint of R with respect to the Euclidean inner product. When the Euclidean inner product is used in deriving the adjoint T21QG model, the operator R^{*} is also the propagator of the adjoint T21QG model. So, the numerical costs of computing the $n\times n$-matrix Q_{loc} is basically determined by integrating the adjoint model n times backwards in time.

3. Daily variability of eigenvectors of Q_{loc}

From now on, we will deal with matrices Q_{loc} as obtained in the 607-day experiment. In this experiment, we studied the impact on regional predictability for three different covariance matrices Q of the analysis error. No significant difference in using these matrices was found. Here we select the case for which Q is equal to the identity matrix and the forecast period is 48 h. The results below also apply to the other two analysis error covariance matrices and a forecast period of 72 h. Choosing Q equal to the identity matrix means that, initially, the errors in the different spectral components are uncorrelated and that their squared amplitudes have the same expectation value. This constraint on $\varepsilon(0)$ can be interpreted as a white noise error distribution. A surface of equal probability density for $\varepsilon(0)$ is now defined by the sphere $\{\|\varepsilon(0)\|_0=\text{constant}\}$, where

$$\|\varepsilon\|_n^2=\iiint \nabla^n\varepsilon\cdot\nabla^n\varepsilon\,\mathrm{d}s. \tag{3}$$

In the 607-day experiment, we also considered a covariance matrix Q for which the initial errors in the different spectral modes have the same expectation value for their kinetic energy. A surface of equal probability density for $\varepsilon(0)$ is now defined by the sphere $\{\|\varepsilon(0)\|_1=\text{constant}\}$. As a third choice for Q, we used a covariance matrix as was determined in Houtekamer (1993). It assumes an equivalent barotropic structure of analysis errors with large amplitudes over the oceans.

The target area in the 607-day experiment is Western Europe (15°W–30°E and 40°N–65°N) and consists of 23 gridpoints. The 23 eigenvectors of Q_{loc} or forecast error patterns vary from day to day. Assume that the corresponding eigenvalues of Q_{loc} are ordered so that $\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_{23}$. Perturbations $R^{*}P^{*}g$, with g a dominant eigenvector of Q_{loc}, grow fast over the Western European area, see Barkmeijer (1993). Usually fast-growing perturbations are defined to yield large amplifications over a finite time interval with respect to some energy measure (Buizza et al., 1993). When integrated in time, it is not certain whether these perturbations or singular vectors will grow fast over a particular area of interest. Use of a projection operator P makes a better control of this possible. For a forecast period of 48 to 72 h, zonal flows in the Western European area frequently yield large values of λ_1 while blocking situations are generally related to small values of λ_1.

Examination of the daily changing patterns in the 607-day experiment shows that the space spanned by the dominant eigenvectors of Q_{loc} is rather invariant during the whole period. We will see that a basis consisting of a relatively small number of elements is capable to describe the first few dominant eigenvectors of all days almost completely. We point out that during the 607-day period, both blocking situations and zonal flows occur. The property of the dominant eigenvectors, to remain in a fixed low-dimensional space, enables an efficient computation of the dominant eigenvalues and corresponding eigenvectors of Q_{loc}. More precisely, let ε_i^d be the ith eigenvector of Q_{loc} for a particular day d with corresponding eigenvalue λ_i^d, using the above ordering. We may assume that the eigenvectors are normalised so that $\langle \varepsilon_i^d, \varepsilon_j^d \rangle = \delta_{ij}$. As a basis for the error patterns we choose the eigenvectors ε_i^1, $i = 1, ..., 23$, for an arbitrary day 1. The eigenvectors for day d can be written as:

$$\varepsilon_i^d = \sum_{j=1}^{23} a_{ij}^d \varepsilon_j^1, \qquad i = 1, ..., 23. \tag{4}$$

During the 607-day period, the dominant eigenvectors of Q_{loc} remain in a space with a dimension considerably smaller than 23. This becomes apparent by computing the average length of the approximated eigenvectors when only $n$vec eigenvectors ε_j^1, where $j = 1, ..., n\mathrm{vec}$, are being used in the expansion (4). The average length l_i of the approximated ith eigenvector for day d is defined by

$$l_i = \frac{1}{N} \sum_{d=1}^{N} \sqrt{\sum_{j=1}^{n\mathrm{vec}} (a_{ij}^d)^2}, \tag{5}$$

where $N = 607$.

In Table 1, we present this length l_i for some values of $n$vec. The length of the approximated ith eigenvector of day d does not deviate much from l_i. In case $n\mathrm{vec} = 9$, the maximal deviation from l_i, $i \leqslant 7$, is 0.03. So, it appears that nine eigenvectors of one particular day already describe the first 7 dominant eigenvectors of any other day to a large extent. The first seven eigenvectors in turn explain over 90% of the variance of the ECMWF forecast error in the target area. In connection with Table 1, we also investigated the impact of different choices for day 1. Selecting days related to, for instance, blocking situations or zonal flows, gave no significant change in Table 1.

Table 1. *The average length l_i of the approximated ith eigenvector is given for a fixed basis of nvec elements*

$n$vec	l_1	l_2	l_3	l_4	l_5	l_6	l_7	l_8	l_9
7	0.98	0.97	0.96	0.94	0.91	0.82	0.78	0.66	0.60
8	0.98	0.98	0.98	0.97	0.96	0.93	0.88	0.77	0.74
9	0.99	0.99	0.99	0.98	0.98	0.97	0.96	0.95	0.91

4. Approximating the eigenstructure of Q_{loc}

The property of the dominant eigenvectors to remain in a fixed space can be exploited to determine the largest eigenvalues and corresponding eigenvectors of the covariance matrix Q_{loc} in an efficient manner. Suppose that the first r eigenvectors $\varepsilon_1^1, ..., \varepsilon_r^1$ of an arbitrary day 1 suffice to describe the first m eigenvectors $\varepsilon_1^d, ..., \varepsilon_m^d$ of day d completely. So, by using a transformation matrix (a_{ij}^d), we have:

$$\varepsilon_i^d = \sum_{j=1}^{r} a_{ij}^d \varepsilon_j^1, \qquad i = 1, ..., m. \tag{6}$$

The vector ε_i^d is eigenvector of the daily varying operator $\mathscr{L}^d = PRQR^*P^*$ with eigenvalue λ_i^d and thus we can write, while omitting the index for day d:

$$\mathscr{L}\varepsilon_i = \lambda_i \varepsilon_i = \lambda_i \sum_{j=1}^{r} a_{ij} \varepsilon_j^1, \qquad i = 1, ..., m. \tag{7}$$

The second identity follows directly from (6).

An alternative expression for $\mathscr{L}\varepsilon_i$ can be obtained by applying $\mathscr{L}$ to both sides of (6). This yields

$$\begin{aligned}\mathscr{L}\varepsilon_i &= \sum_{j=1}^{r} a_{ij} \mathscr{L}\varepsilon_j^1 \\ &= \sum_{j=1}^{r} a_{ij} \sum_{k=1}^{23} c_{jk} \varepsilon_k^1 \\ &= \sum_{k=1}^{23} \left(\sum_{j=1}^{r} a_{ij} c_{jk} \right) \varepsilon_k^1, \qquad i = 1, ..., m. \end{aligned} \tag{8}$$

Notice that we expressed $\mathscr{L}$ in matrix form (c_{ij}) with respect to the basis, consisting of vectors ε_k^1,

$k=1, ..., 23$. The vectors ε_k^1 are orthonormal to each other for the Euclidean innerproduct $\langle\ ,\ \rangle$ and the entries c_{ij} of the matrix are given by

$$c_{ij} = \langle \mathscr{L}\varepsilon_i^1, \varepsilon_j^1 \rangle = \langle B^*R^*P^*\varepsilon_i^1, B^*R^*P^*\varepsilon_j^1 \rangle. \quad (9)$$

By combining (7) and (8) and taking Euclidean innerproducts with ε_k, $k=1, ..., r$, we obtain the following expression:

$$\sum_{j=1}^{r} a_{ij} c_{jk} = \lambda_i a_{ik}, \qquad k=1, ..., r$$

$$\text{and} \quad i=1, ..., m. \quad (10)$$

From (10) we conclude that the ith column of the $m \times r$-transformation matrix (a_{jk}), as was used in (6), is an eigenvector of the symmetric matrix $C_r = (c_{jk})_{1 \leqslant j,k \leqslant r}$ with eigenvalue λ_i.

Thus the computation of C_r enables us to determine the dominant part of the eigenstructure of Q_{loc} for day d. In defining C_r, the vectors $B^*R^*P^*\varepsilon_i^1$, $i=1, ..., r$, are needed, see (9). The numerical costs to compute such a vector are mainly determined by an integration of the adjoint model backwards in time, along the forecast orbit starting from day d. The Euclidean innerproducts of pairs of these vectors provide then the entries of the matrix C_r.

In general (6) will not hold exactly. Yet the eigenvalues of C_r will converge to the eigenvalues of Q_{loc} in a monotonous way when r is increased and also the corresponding eigenvectors will approximate the eigenvectors of Q_{loc}. For let C_r and C_{r+1} have eigenvalues $\lambda_1, ..., \lambda_r$ and $\lambda'_1, ..., \lambda'_{r+1}$, respectively, and ordered with decreasing magnitude. Then a Sturmian separation theorem (Everitt, 1955) states that $\lambda_k \leqslant \lambda'_k \leqslant \lambda_{k-1}$, $k=1, ..., r$, with the convention that $\lambda_0 = \infty$. The eigenstructure of C_{23} and Q_{loc} are identical.

5. Numerical results

In order to investigate the feasibility of the method, we compute C_r during some period and compare the results with Q_{loc}. We will concentrate upon the approximation of the eigenvalues of Q_{loc} but the results also apply to the eigenvectors. The selected test period consists of 62 days: 15 November 1991 1200 UTC–15 January 1992 1200 UTC. For each day during this period, the matrix Q_{loc} is computed with a forecast period of 2 days. The eigenvectors of Q_{loc} for 15 November 1991 1200 UTC are used as a basis ε_j^1, $j=1, ..., 23$, in the expansion (4). The r eigenvalues of $C_r = (c_{ij})_{1 \leqslant i,j \leqslant r}$ approximate the eigenvalues of Q_{loc}. The maximal deviation (in percentages) for each ith eigenvalue, $i=1, ..., r$, from the actual eigenvalue of Q_{loc} is determined during the 62-day period. Thus we obtain r values which inform about the accuracy of the method in estimating the first r eigenvalues. These values will decrease monotonously when r is increased. Fig. 1 shows for some values of r the maximal deviation of the eigenvalues of C_r during this period. We conclude that approximating 9 eigenvalues of Q_{loc} within an accuracy of 1% takes at most 16 integrations with the adjoint model. The complete matrix Q_{loc} can be determined in 23 adjoint model integrations. Allowing the same computation time to the Lanczos algorithm one obtains, on average, 7 eigenvalues sufficiently accurate. The error made in approximating eigenvalue λ_7 of Q_{loc} is usually of the order of 1%. So when the first 7 eigenvalues and eigenvectors of Q_{loc} are needed within an accuracy level of 1%, use of this simple method prevails. A reduction of at least 30% in computational costs is possible. If, on the other hand, an accurate estimate of only the largest eigenvalue is desired, the Lanczos algorithm is more advantageous.

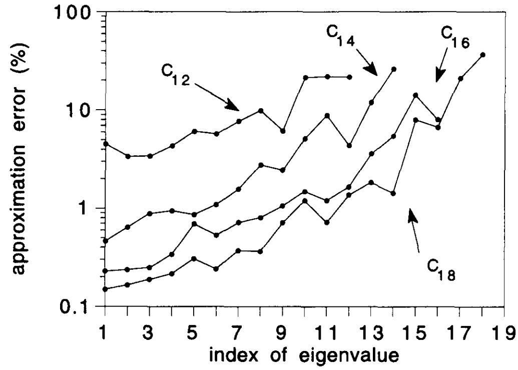

Fig. 1. For each of r eigenvalues of C_r, the maximal deviation (in %s) from the corresponding eigenvalue of Q_{loc} is presented during a 62-day period. The values r for which C_r is computed are indicated.

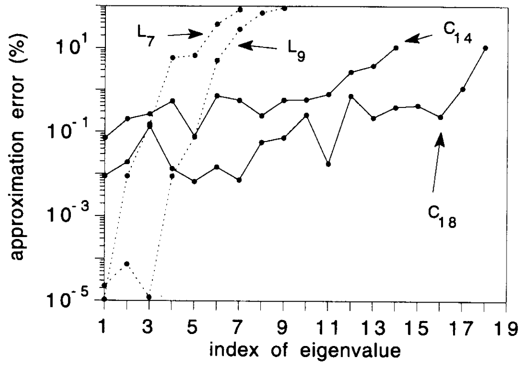

Fig. 2. The accuracy of the eigenvalues of C_{2r} in approximating the eigenvalues of $Q_{\rm loc}$ is compared with the performance of the Lanczos algorithm after r iterations (L_r).

For an arbitrary day, 22 November 1991, we compare the above approach with the Lanczos algorithm. After r Lanczos iterations r approximating eigenvalues of $Q_{\rm loc}$ are determined. Note that each Lanczos iteration involves the evaluation of $PRQR^*P^*$. Therefore r Lanczos iterarations (L_r) are as expensive as determining the matrix C_{2r}. In Fig. 2, the deviation from the corresponding eigenvalue of $Q_{\rm loc}$ is given for the two methods in case r equals 7 and 9. It shows that L_7 yields 3 eigenvalues within a 1% accuracy, while already 11 eigenvalues of C_{14} are accurate within this margin. Notice also that the accuracy of the approximating eigenvalues with a high index does not decrease so much when, instead of the Lanczos algorithm, the matrix C_r is used (see also Fig. 1).

6. Concluding remarks

We have presented a simple method which was able to compete with the Lanczos algorithm in computing the dominant eigenvalues and corresponding eigenvectors of local covariance matrices. For a given fixed analysis error covariance matrix, we considered the daily forecast error covariance matrix for the Western European area at 500 hPa. It appeared that the error space spanned by the dominant error patterns does not vary much during a long period. When different analysis error covariance matrices are used, almost the same error space resulted (not shown).

Instead of restricting the covariance matrix to the Western European area at 500 hPa, also some experiments have been performed in which the remaining model levels at 200 hPa and 800 hPa were included. Results (not shown) indicated that the dominant error patterns were not confined to such a low dimensional error space as in the one-level case. Enlarging the target area at 500 hPa to, for instance, the Northern Hemisphere will probably yield the same result. In computing the eigenstructure of the covariance matrix for the three-level case, the Lanczos algorithm showed, on average, a better performance than the above method.

7. Acknowledgements

I would like to thank Theo Opsteegh and Robert Mureau of the predictability section at KNMI for their careful reading of the manuscript. Also thanks are due to two reviewers for giving useful comments.

REFERENCES

Arpe, K. 1990. Impacts of changes in the ECMWF analysis-forecasting scheme on the systematic error of the model. Proceedings of the seminar on: *10 years of medium-range weather forecasting, 1989*. Reading, England, ECMWF, vol. I, 69–114.

Barkmeijer, J. 1993. Local skill prediction for the ECMWF model using adjoint techniques. *Mon. Wea. Rev.* **121**, 1262–1268.

Barkmeijer, J., Houtekamer, P. L. and Wang, X. 1993. Validation of a skill prediction method. *Tellus* **45A**, 424–434.

Bouttier, F. 1993. The dynamics of error covariances in a barotropic model. *Tellus* **45A**, 408–423.

Buizza, R., Tribbia, J., Molteni, F. and Palmer, T. 1993. Computations of unstable structures for a numerical weather prediction model. *Tellus* **45A**, 388–407.

Errico, R. M., Vukićević, T. and Reader, K. 1993. Examination of the accuracy of a tangent linear model. *Tellus* **45A**, 462–477.

Everitt, W. N. 1955. Two theorems in matrix theory. *American Math. Mon.* **69**, 856–859.

Houtekamer, P. L. 1993. Global and local skill forecasts. *Mon. Wea. Rev.* **121**, 1834–1846.

Kalnay, E., Kanamitsu, M. and Baker, W. E. 1990. Global numerical prediction at the National Meteorological Center. Proceedings of the seminar on:

10 years of medium-range weather forecasting, 1989. Reading, England, ECMWF, vol. II, 139–167.

Lacarra, J.-F. and Talagrand, O. 1988. Short-range evolution of small perturbations in a barotropic model. *Tellus* **40A**, 81–95.

Leith, C. 1974. Theoretical skill of Monte Carlo forecasts. *Mon. Wea. Rev.* **102**, 409–418.

Marshall, J. and Molteni, F. 1993. Towards a dynamical understanding of planetary-scale flow regimes. *J. Atmos. Sci.* **50**, 1792–1818.

Parlett, P. 1980. *The symmetric eigenvalue problem.* Series in computational mathematics. Prentice Hall, New Jersey, 348 pp.