Pieter Leopold Houtekamer

Predictability in Models
of the Atmospheric Circulation

Proefschrift
ter verkrijging van de graad van doctor
in de landbouw- en milieuwetenschappen
op gezag van de rector magnificus,
dr. H.C. van der Plas,
in het openbaar te verdedigen
op woensdag 9 september 1992
des namiddags te vier uur in de Aula
van de Landbouwuniversiteit te Wageningen.

Promotor: dr. ir. J. Grasman
                 hoogleraar in de wiskunde, inclusief de numerieke
                 wiskunde
Co-promotor: dr. ir. J.D. Opsteegh
                 wetenschappelijk hoofdmedewerker aan het KNMI

Stellingen

behorende bij het proefschrift

Predictability in Models
of the Atmospheric Circulation

1. De onvoorspelbaarheid van het weer is ten dele voorspelbaar.

2. Voor het experimenteel vaststellen, met 95 % zekerheid, dat een perfect werkende betrouwbaarheidsverwachting voor de twee-daagse weersvoorspelling inderdaad enige kwaliteit heeft moet de fout in minstens ongeveer 150 voorspellingen vergeleken worden met de corresponderende betrouwbaarheidsverwachting.

3. Het toepassen van de in dit proefschrift behandelde methoden op modellen voor de foutengroei met een hogere resolutie of met een betere representatie van de fysica hoeft geen betere resultaten te geven.

4. De "Lagged Average Method" heeft geen wetenschappelijke basis en dient daarom verworpen te worden.

5. De beschrijving van modelfouten zoals gebruikt in hoofdstuk 4 van dit proefschrift kan verbeterd worden met de "maximum likelihood" methode.

6. Een goed uitgevoerde bifurcatieanalyse hoeft niet te leiden tot een volledig beeld van alle aantrekkers in de faseruimte.

   P. Houtekamer, An analysis of the Lorenz-1984 equations, Afd. Wiskunde, Rijks-universiteit Utrecht, Preprint 563, 1989

7. Een punt dat met homogene waarschijnlijkheid willekeurig geplaatst is in een n-dimensionale kubus ligt met de zeer kleine kans $2^{1-n}\pi^{n/2}/n\Gamma(n/2)$ in de grootst mogelijke door de kubus omsloten bol. Voor het homogeen willekeurig plaatsen van een punt in een bol is het daarom aan te bevelen eerst een willekeurige richting te bepalen met normaal verdeelde getallen.

8. Bij de beschrijving van interakties tussen vier zwaartegolven in een systeem met periodieke randvoorwaarden en met dispersie van de vorm:

$$\omega \propto (n^2 + m^2)^{1/4}$$

treedt het volgende stelsel vergelijkingen op:

$$(*)\begin{cases} r_1^{1/4} + r_2^{1/4} &=& r_3^{1/4} + r_4^{1/4} \\ r_3 &\neq& r_1 \neq r_4 \\ r_i &=& n_i^2 + m_i^2, \quad i = 1, \cdots, 4 \\ n_1 + n_2 &=& n_3 + n_4 \\ m_1 + m_2 &=& m_3 + m_4 \\ n_i, m_i &\in& \mathbf{N} \\ n_i, m_i &\leq& M \in \mathbf{N} \end{cases}$$

Het stelsel $(*)$ heeft de eerste oplossing voor $M = 495$:

$$\begin{aligned} n_1 &= 495, & n_2 &= 64, & n_3 &= 359, & n_4 &= 200 \\ m_1 &= 90, & m_2 &= 128, & m_3 &= 118, & m_4 &= 100 \end{aligned}$$

Deze oplossing is niet relevant voor het achterliggende fysische probleem.

Het stelsel $(*)$ heeft een benaderende oplossing:

$$\begin{aligned} n_1 &= 23, & n_2 &= 25, & n_3 &= 24, & n_4 &= 24 \\ m_1 &= 26, & m_2 &= 24, & m_3 &= 28, & m_4 &= 22 \end{aligned}$$

Zowel met wiskundige als met numerieke methoden kan aangetoond worden dat dit slechts een benaderende oplossing is. Wellicht heeft deze oplossing fysische relevantie.

9. Het aantal proefschriften dat gedrukt is op kringlooppapier is veel kleiner dan het aantal stellingen over het milieu.

10. Omdat niet iedere automobilist groot licht dimt voor fietsers verdient het aanbeveling ieder fietspad te voorzien van witte reflecterende strepen.

P. Houtekamer   Wageningen, 9 september 1992

Aan allen die mij tot hier gebracht hebben

Contents

Chapter 1

Introduction

Weather forecasts contain relevant information for a relatively short period of time only. After a few days the correspondence between the forecast and the observed weather phenomena decreases rapidly. The eventual failure of a forecast may point to a failure of atmospheric scientists to properly simulate the atmospheric evolution with their models. However, this is not a priori clear. It may very well be that the atmosphere is so sensitive to small and unavoidable differences in the description of the initial state of the atmospheric circulation that today the accuracy of the forecasts is close to the theoretical limits. Scientists are forced to ask the question: is it still possible to significantly improve the daily weather forecast or are we close to reaching perfection?

Models combine the scattered information from past and present observations into one complete picture of the atmosphere. The quality of this picture is so high that it is used as a faithful description of the large scale atmospheric circulation. This description is then used as if it were the actual state of the atmosphere. For example if one would be interested in the average temperature at 700 mb during winter, one would use the numerical weather prediction (NWP) archives containing the winter states. Nobody would consider obtaining these values from, for instance, radiosonde measurements.

The fundamental laws of motion which apply to the atmosphere are well known. It is in general not known how these laws can faithfully be represented in a model. It is for instance still not clear how to adequately represent the effect of mountains on the atmospheric circulation. Scientists

have to design some kind of parametrization instead. An improvement of the parametrization leads to an improvement of the model results. With forecasts of the quality of a forecast, so called skill forecasts, one has to deal with a similar situation. It is known how the evolution of error statistics can be described. However, a direct implementation of this knowledge causes insurmountable computational difficulties, such as the treatment of an infinite series of moments. Thus numerical methods have to be developed which approximate the knowledge of error growth processes in such a way that the skill forecast is both computationally feasible and of relevance to forecasters. The development of methods to forecasts the skill is the subject of this thesis.

In this thesis errors in the prediction are assumed to be caused only by errors in the initial state. Initial errors are unavoidable because it will never be possible to measure the atmosphere in all detail without errors. The evolution of the errors is described with the model that is assumed to be perfect. This approach may certainly be criticized. The atmosphere may have a sensitivity to errors which is different from the sensitivity of the model to errors and the model trajectory may systematically diverge from the atmosphere due to model errors. However, if models will continue to improve, the assumption of a perfect model will become less serious.

In this thesis attention is restricted to the quality of the forecast during the first four days of the forecast. The restriction to four days makes it possible to use linear theory for the evolution of initially small errors. A skill forecast is obtained using the forecast for the atmospheric circulation as a reference orbit in phase space around which a simple model is linearised.

The introduction starts with a description of fundamental limits to the predictability. It will appear that the present and past states of the atmosphere and its environment do not uniquely determine the state at all future times (Lorenz 1965). Less fundamental, but quite significant, are the limits imposed by the deficiencies of models. The introduction continues with an overview of current methods to produce error statistics. In this thesis adjoint models are used. A short overview of their applications is given. As a final point the assimilation of data is treated. Data

assimilation provides an estimate of the current state of the atmospheric circulation (the analysis) and its error statistics. With a skill forecast the time evolution of these error statistics is estimated.

1.1 Fundamental limits

The predictability of a system can be classified into three categories, according to the general evolution of initially small errors (Lorenz 1969):

1 *At all future times the error remains comparable to or smaller than the initial error. The error may be kept arbitrarily small by making the initial error sufficiently small.*

2 *The error eventually becomes much larger than the initial error. At any particular future time the error may be made arbitrarily small by making the initial error sufficiently small, but, no matter how small the initial error (if not zero), the error becomes large in the sufficiently distant future.*

3 *The error eventually becomes much larger than the initial error. For any particular future time there is a limit below which the error cannot be reduced, no matter how small the initial error (if not zero) is made.*

Systems in the first category in general show either periodic or constant behavior. Sofar nobody has produced evidence that the atmosphere is in the first category.

An example of a system in the second category is Lorenz's famous model for convection (Lorenz 1963). This model, which has three components, shows a non-periodic solution. Initially nearby solutions diverge until eventually their distance is comparable to the distance of two randomly chosen states. Many, if not all, of the current models for the atmospheric circulation fall into this second category. If the required calculations are performed with a sufficiently high precision and if the forecast error is small, then a reduction of the initial error appears to lead to a proportional reduction in the forecast error. It will then take at least a fixed finite time before the forecast error appears to reach its old level. The fixed finite time

is determined by the maximum rate of divergence of the model. Thus, for a system in the second category, it is useful to reduce the size of the initial error until eventually it is infinitesimally small.

This is not true for a system in the third category. Here the forecast quality is limited independent of the size (if not zero) of the initial error. A suggestion that the atmosphere behaves in this way comes from the following sequence of historic estimates of the error doubling time in ever more complex models. With a two-layer model for the northern hemisphere Smagorinsky estimated that errors double in about 8 days (Charney et al. 1966). A more advanced and global model of Mintz and Arakawa gave an error doubling time of 5 days (Charney et al. 1966). With a nine-level primitive equation model Smagorinsky (1969) obtained a 3 day doubling time. In 1982 Lorenz obtained a doubling time of 2.1 days with the then current version of the ECMWF operational model. In this sequence the doubling time refers to the initial growth rate of errors present at the synoptic scales. The growth rate of errors increases if one goes from a certain atmospheric model to a similar model with a higher resolution. One may fear that the atmosphere itself will display very short error doubling times at its smallest scales. Using turbulence theory, which extends the estimates to sub-synoptic scales, it is indeed argued that the atmosphere itself is a system in the third category. For small scales one may try to get estimates of the predictability with limited-area models with a very high spatial resolution. For the larger scales it is possible to use the Global Circulation Models (GCM's) which are the basis of the daily weather forecast. In this section the three different approaches are discussed and combined. The result is a statement on the performance of a hypothetical perfect weather forecasting model.

1.1.1 Turbulence theory

Lorenz (1969) uses a simple mathematical model to argue that the predictability of the atmosphere is limited in the sense of the third category above. An equation whose dependent variables are ensemble averages of the "error energy" in separate scales of motion is derived from the vorticity equation for two-dimensional flow. Solutions of the equation are

determined for cases where the horizontal extent and total energy of the system are comparable to those of the earth's atmosphere.

It is found that each scale of motion possesses an intrinsic finite range of predictability, provided that the total energy of the system does not fall off too rapidly with decreasing wavelength. For the smallest scale in the model, which is 40 meters, it appears that a non-zero initial error grows to saturation in 1.8 minutes or less. For a scale of 625 km this time is 0.65 days and for global errors it is 16.7 days. Thus, the atmosphere, which is formally a deterministic system, is observationally indistinguishable from an indeterministic system. Two states of the atmosphere differing initially by a small observational error will eventually evolve into two states differing as greatly as two randomly chosen states. This occurs within a fixed time no matter how small the non-zero initial error may be made. This is known as the "butterfly" effect. If at some part of the world a butterfly decides to follow a different course this may in a few weeks cause a completely different evolution of the global weather.

Lorenz based his computations on a rather speculative energy spectrum. Different estimates were obtained by Leith (1971) and Leith and Kraichnan (1972). Their energy spectra have less energy at the smallest scales. In addition they used an eddy-damped Markovian approximation for the closure problem while Lorenz used a modified quasi-normal approximation. Unfortunately a fundamental theory does not exist and the closure has to be made on phenomenological grounds. This makes it difficult to say which form of closure is the best one. Although qualitatively the results are similar the latter authors obtained predictability durations several times longer than Lorenz's. Lorenz's postulated energy spectrum has been confirmed by observations (e.g. Lilly and Peterson 1983). At large scales the atmosphere has an energy spectrum which follows the -3 power law of 2 dimensional turbulence. At small scales it has the -5/3 power law corresponding to 3 dimensional turbulence. The general applicability of turbulence theory to the predictability at small scales has been questioned because of the highly intermittent nature of small-scale weather events (e.g. Lilly 1985). Because of the uncertainty in Lorenz's results one may decide to follow an entirely different method. One may use

a weather model that has been explicitly designed to describe small-scale weather events.

1.1.2   Limited-area models

As an alternative, to the results from turbulence theory, it has been attempted (e.g. Anthes et al. 1985) to estimate predictability on small scales (100 to 1000 km) using a limited-area mesoscale model with prescribed boundaries. Anthes et al. found little or no error growth, as measured in a root mean square sense, during 72-hour forecast periods. This is distinctly different from what turbulence theory predicts. The explanation is to be found in artificial factors which are present in limited-area models. An example is the existence of a prescribed lateral boundary which is assumed to be accurate (e.g. Errico and Baumhefner 1987). If the lateral boundaries are identical for the control and the perturbed forecast, the growth of differences between the forecasts is artificially limited. In practice lateral boundary perturbations contaminate the solution in the interior within three to six hours. This suggests that medium-range forecasts with limited-area models will not be better than the same forecasts performed with global models. A review of predictability studies based on limited-area models is given by Vukićević and Errico (1990). They show that predictability of synoptic and large scale flow increases with a better representation of topography. This happens because surface inhomogeneities, if correctly incorporated in numerical models, may correctly induce small scale events (such as mountain waves). On the predictability at short scales (scales shorter than 200 km) they conclude that a more careful treatment of the short-scale dynamics than presently available in numerical weather prediction (NWP) models is needed.

In view of these problems with limited-area models the result from turbulence theory is accepted that, however accurate the initial state may be, the synoptic scales will always be contaminated by errors after about 0.65 days.

1.1.3 Global Circulation Models

For the predictability at larger scales it is possible to use results from Global Circulation Models (GCM's). One may obtain both a lower and an upper limit for the average atmospheric predictability. The lower limit is simply given by the present performance of the NWP models. The upper limit is directly obtained from the NWP models by computing the divergence of two initial states that are slightly different. This was first done by Lorenz (1982). At that time the model of the European Centre for Medium-Range Weather Forecasts (ECMWF) had obtained such a quality that the errors in the one day forecast could be considered as small perturbations to the true state one day later. In his study Lorenz assumes that the growth rate of errors in models is characteristic for the growth rate of errors in the real atmosphere. This is optimistic because in the atmosphere, which contains all scales, errors are likely to grow faster than in a model. Lorenz's estimates of the predictability were based on a hundred day dataset, starting at December 1 1980. It appears that, if a root mean square error of 60 meters for the geopotential height of the 500 mb level is taken as a limit to a useful forecast, this limit is reached after 3.5 days. Looking at the divergence between two successive model runs this limit is reached after 6 days. Thus in theory one might gain 2.5 days when a perfect model is available.

Recently Lorenz (1990) has repeated his analysis. This time data are for the 100-day period beginning on 1 December 1986. Internal divergence gives an error of 60 meters after 7 days. In reality the forecast becomes invalid after 4.5 days. Thus although the forecast has improved by as much as one day the potential for improvement is still 2.5 days. The explanation for this strange result is that the analysis error has decreased considerably in the same period. The data from 1987 and 1988 indicate that the potential for improvement now has decreased to 1.25 days (Lorenz 1990). If one accepts Lorenz's (1969) conclusion that a small but significant error in the synoptic scales will always be reached within about 0.65 days, one comes to the conclusion that the ultimate weather forecasting system, which starts from a vanishingly small but non-zero error in the smallest scales of motion and which perfectly represents all processes of relevance

for the atmosphere, will on average be valid for two days more than at
present.

1.2   Deficiencies of models

The quality of atmospheric models is, on average, increasing as a result of
modifications in their formulation. Nevertheless the deficiencies of present
models are such that they make the above discussion on fundamental limits
largely academic. Even if the analysis error could be made zero, model
deficiencies would introduce errors at the largest scales during the first
few days of the forecast (Boer 1984). This implies that skill predictions
should be based on a study of both internal errors and of model errors.

In practice it is not easy to distinguish model errors and internal errors
(Tribbia and Baumhefner 1988). From the point of view of model design
this is disturbing. If the model errors are much smaller than the internal
errors it may no longer be possible to distinguish them without having a
large dataset available. Likewise if the effect of internal errors becomes
small relative to the effect of model errors there is no point in further
improving the observing system. From this it appears logical that model
errors and internal errors are removed simultaneously where at each time
most attention is given to the then weakest part of the prediction system.

Studies by Lorenz (1982, 1990) indicate that between 1981 and 1987
the quality of the analysis and the quality of the model improved simul-
taneously. Between 1987 and 1989 the principal improvements were in
the model. One would expect this trend to continue. The quality of a
one day forecast has an upper limit, due to the rapid spectral cascade of
small scale errors towards the synoptic scales. Thus after one day there
is always a certain minimum error which starts to grow due to the inter-
nal dynamics. The superimposed effect of model errors becomes smaller
with every improvement of the model. At some point model errors, or
model improvements, will cease to have a detectable effect on the quality
of short- and medium-range forecasts.

One may distinguish a systematic and a random component of the
model error. The systematic model errors are studied mainly from a diag-

nostic point of view. That is one runs a model for some time, say 10 days, and one observes that the final flow is not atmospheric in some sense. So, with a certain atmospheric model, it might appear that the troposphere is always cooling down slowly. One may then adapt the model such that this problem no longer occurs. Because systematic model errors are difficult to detect in present models, research is now going into methods to quantify model performance with methods that are independent of observational errors (Lions 1990). Such methods can lead to optimal estimates of model parameters. In a skill forecast these estimates can be treated in the same way as estimates of the initial states. Methods developed to quantify the effect of the internal errors on the forecast can also be used to quantify the effect of model errors.

The model error may also have a random component. In Thiébaux and Morone (1990) it was demonstrated that the short-term forecast error has a fluctuating component with a time scale of about 10 days. Because they studied forecast errors with a range of 1 day or less, which is much less than the time scale of the detected random error, this component of the random error could be estimated and removed. The random forecast error may well have components with shorter time scales. One may even assume that successive forecasts have independent random model errors. Methods for estimating such random model errors are given by Bélanger (1974). A more efficient version of this method appeared in Dee et al. (1985a, 1985b). A possible cause of random model errors with a short time scale are truncation errors. These errors give random contributions during the forecast run. Parametrizations may cause random errors with a longer time scale. Depending on the flow situation a parameter value may be either too low or too high. Such random model errors will have the time scales of the flow features which cause them.

In this thesis model errors will be neglected because model errors can only be studied usefully in the context of operational models. One needs to compare actual predictions with actual observations to diagnose model errors. For the simple 30 component model, which is used in the first part of this thesis, a study of its model errors is not relevant.

1.3   Evolution of error statistics

Lorenz (1965) was the first to study the variability of atmospheric pre-
dictability. He used an atmospheric model with 28 variables and linearized
the equations for the evolution of errors around a reference solution. The
average forecast error was obtained from the spectrum of eigenvalues of
the covariance matrix of the forecast error. This matrix is symmetric and
can have only positive eigenvalues. These two properties make it easy to
study this matrix numerically.

He observed that the time required for small initial errors to grow to
intolerable size is strongly dependent upon the circulation pattern, and
varies from a few days to a few weeks. His model has a variable, related to
the position of the jet stream, with a much better predictability than other
variables. This is an indication that specific skill forecasts may be needed
for specific variables. It is not sufficient to give a probability distribution
for one global measure of the error.

The variability in the success of forecasts is now well known. It is
generally felt that we should try to predict the skill. Tennekes et al.
(1986) stress the need for objective methods. They suggest to give a skill
forecast in a geographically differentiated way, because predictability is a
local variable.

Lorenz (1965) noted that his methods are not immediately applicable
to more realistic models for computational reasons. He mentioned a prim-
itive equation model (Smagorinsky 1963) with 5184 variables which was
the current state of the art. A repetition of Lorenz's study for this model
would require 5184 separate numerical integrations, each with a set of
initial conditions differing from a basic set in only one variable. On the
then current computers this was out of the question. In a sense this prob-
lem has only become worse; the complexity of weather forecasting models
is keeping up with the rapidly increasing performance of computational
facilities.

The ideas of Lorenz (1965) have also appeared in dynamical system
theory (Goldhirsch et al. 1987). The eigenvalues of the covariance matrix
for the forecast error were shown to be related to so called time-dependent

Lyapunov exponents. In a finite dimensional system and in the limit of infinite time the time-dependent Lyapunov exponents become equal to the Lyapunov exponents. Lyapunov exponents are global properties of an attractor and are related to other global properties such as the dimension of the attractor (Kaplan and Yorke 1979). Unfortunately there is little evidence that the atmosphere behaves like a finite dimensional system in which Lyapunov exponents have a well defined meaning. Recently the concept of time-dependent Lyapunov exponents is being reconsidered in atmospheric sciences as well. This time the name of the method is optimal perturbation analysis (Farrell 1989, 1990) or finite-time Lyapunov stability analysis (Yoden et al. 1992). In this thesis it is discussed how optimal perturbations can efficiently be determined with methods which are based on the use of adjoint models.

1.3.1 Stochastic dynamic prediction

From a mathematical point of view the description of the evolution of the error statistics is straightforward. One can use a continuity equation for the evolution of the probability density in the same way as one uses a continuity equation for the evolution of the mass density (Epstein, 1969). Unfortunately the evolution equation for the first order moments includes terms of second order moments. The equation for the second order moments requires third order moments and so on. It was shown by Freiberger and Grenander (1965) that as long as the deterministic prognostic equations are nonlinear it is impossible to write a closed finite set of prognostic equations for the moments. This means that it is impossible to predict exactly the future behavior of even the mean of the distribution because for this all moments of the distribution must be predicted as well. Epstein concludes that the series of equations must be truncated after the second order terms. It is assumed that terms of third and higher order have only a small contribution. This assumption is reasonable for a short time if the initial error is sufficiently small. Unfortunately models have grown beyond the complexity foreseen by Epstein. Pitcher (1977) integrated the moment equations with second moment closure. According to Pitcher the computational cost is excessive. Fleming (1971) did experiments with

third moment closure. He remarks that the volume of computation is monumental but finite. Simplified schemes are proposed by Thompson (1985, 1986).

Thompson (1986) proposed a simple approximate method which is based on strong assumptions on the statistics of the initial errors. In particular he supposed that the ensemble of initial error fields is statistically isotropic and homogeneous, in the sense that the autocorrelation function of any particular velocity component is independent of the direction of that component. He showed that the local growth or decay of error variance depends primarily on the detailed structure of the vorticity field; in general, the most rapid error growth can be expected in concentrated regions of strong vorticity gradient. Although his approximate method requires only a marginal increase of the computations needed for a deterministic prediction, it has not yet been applied to actual short-range predictions.

1.3.2   Monte Carlo forecasts

Lorenz (1965) suggested to perform a Monte Carlo experiment in a model of reasonable complexity. In this an ensemble of error fields is added to the analysis. The error fields are chosen by a random process such that the likelihood of selecting any given error field is proportional to the initial probability density. Means and variances are determined for the ensemble. Leith (1974) states that a small ensemble with only about 10 members may be sufficient to get results comparable with a stochastic dynamic technique. The big advantage of the Monte Carlo method is that it gives unbiased results which converge to the true values when the number of ensemble members approaches infinity. Imperfections of the model can, if the corresponding statistics are known, also be sampled with a Monte Carlo process.

A technique to accelerate the convergence of the obtained statistics is stratified sampling (Kleijnen 1974; Balgovind et al. 1983). This sampling technique makes the ensemble more "representative" using a division of the initial state into equally likely classes. For instance if the initial state is defined by a scalar value between 0.0 and 1.0, one may generate a strati-

fied ensemble of 10 members by drawing the first random number between 0.0 and 0.1, the second random number between 0.1 and 0.2 and so on. A stratified sampling technique for random vectors of moderate dimension is Latin Hypercube sampling (McKay et al. 1979). Unfortunately these techniques are an improvement only when the ensemble size is large in comparison with the length of the random vector (Stein 1987). It is not a priori clear that in the case of atmospheric predictability the number of important initial error directions is small. If it is small then an equally small number of random numbers defines a member of the ensemble. A combination of stratified sampling methods with Lagged average methods, which are discussed later, would appear rational, because the Lagged average method is based on the assumption that the number of relevant initial error directions is very limited.

An example of a Monte Carlo experiment is the study by Kalnay and Dalcher (1987). For the period 7 January to 2 February 1979 14 forecasts were made. Both the sea-level pressure and the height of the 500 mb level were predicted. This gave a total of 28 global forecasts for which a Monte Carlo experiment could be performed. For each forecast a total of 5 ensemble members was generated using different sets of observations for the data assimilation. The first ensemble member used all observations. The second member did not use satellite data. The third, fourth and fifth member did not use temperatures, winds and cloud-tracked winds respectively. It is not clear that the removal of groups of measurements is a good strategy to simulate random observational errors, but it was assumed that the 5 ensemble members form a representative ensemble. It was attempted to predict the moment at which the forecast breaks down. This is the moment at which the correlation between the forecast and the subsequent analysis, computed relative to the climate, becomes less than 0.6. Usually this will happen after about 6 days. Because in this study the forecasts were only 5 days long, most of them remained skillful till the end of the forecast period. However about 20 % of the forecasts broke down before five days. A prediction of the quality of the forecasts was made on the basis of the spread of the ensemble. Contingency tables were constructed to validate the skill prediction. When the northern

hemisphere was used as a verification region, the prediction of the skill was rather poor. For the individually verified regions of North America, Europe, the North Atlantic and the North Pacific the skill forecast showed a highly significant performance. In view of this success it is unfortunate that no follow-up experiment has appeared in the literature.

1.3.3   Lagged average forecasts

As an alternative to Monte Carlo forecasting Hoffman and Kalnay (1983) introduced the Lagged Average Forecasting (LAF) technique. In a Lagged average forecast, just as in a Monte Carlo forecast, sample statistics are calculated from an ensemble of forecasts. Each ensemble member is an ordinary operational forecast, but they start from initial conditions that are for instance one day apart. So an ensemble at day five consists of the present forecast for day five and yesterday's forecast for day six and the seventh day forecast starting from the day before yesterday and so on. The forecasts are averaged at their proper verification times to obtain a lagged average forecast. If the spread in the ensemble is large the forecast is assumed to be poor. The LAF method is operationally feasible since the ensemble members are produced during the normal operational cycle. In fact the LAF method - the a priori prediction of forecast skill on a case by case basis - has been qualitatively used since the start of the numerical prediction of the weather. Forecasters have always had more confidence in a numerical prognosis when it was similar to older forecasts for the same verification time. A sceptic might argue that this tells us more about the quality of the old forecast than about the quality of the new forecast. In general one should not use a high quality forecast to predict the skill of a low quality forecast. It leads only to a good skill forecast for a low quality forecast which one does not intend to use (Palmer and Tibaldi 1988).

Recently the LAF method has been applied to operational models by Branković et al. (1990) and by Tracton et al. (1989). These studies aim at numerical weather prediction beyond the limit of instantaneous deterministic predictability. The peculiar construction of the LAF ensemble poses an upper limit to the number of ensemble members. For example Branković et al. use 30-day integrations which start from 9 consecutive

6-hourly analyses. It is not clear why this ensemble should be representative, but usually it is assumed that the ensemble properties become better with increasing forecast time. The general idea is that an extended range prediction can be issued if the spread within the ensemble is small. So far the obtained correlations between skill and ensemble spread have not been very high. This may be due to the working hypothesis of a perfect model, to the sampling strategy or also to the general difficulty of verifying a property of a probability distribution (spread) with a property of a single forecast (skill). Even if one computes correct probability distributions the subsequently observed correlation between spread and skill may be quite low. The conclusion that must be drawn is that the LAF method has not yet reached the quality which is necessary for extended range forecasts.

1.4 Adjoint models

A major development in meteorology is the introduction of adjoint models (Marchuk 1974). Using adjoint models one can compute gradients of functionals. Thus traditional trial and error methods can be replaced by mathematically more advanced schemes. A mathematical description of adjoint methods is given by Cacuci (1981). This paper contains references to the first uses of adjoints in nuclear physics in the 1940's.

The adjoint method starts by expressing a result as a functional, customarily called the response or the cost function, of the model variables. The sensitivity of this response is defined in terms of a functional derivative. A system of adjoint equations is developed from a differentiated form of the original equations. A single integration of the adjoint model suffices to obtain the sensitivity to all parameters of interest. An early example, in atmospheric science, is a sensitivity analysis of a radiative convective model by Hall et al. (1982). In this the sensitivity of the surface air temperature to all 312 parameters of the model is computed. A physical interpretation of their adjoint functions is given in Hall and Cacuci (1983). A slightly different application is the computation of the sensitivity of one aspect of the forecast to the uncertainties in the initial conditions. This is done for a short-range weather forecast by Errico and Vukićević (1992).

Le Dimet and Talagrand (1986) compare different variational algorithms for the assimilation of meteorological observations into a coherent estimate of the state of the atmosphere. Using a simple example they demonstrate the possible use of the adjoint method. A more detailed description is given in Talagrand and Courtier (1987). This paper is highly recommended as a general introduction to the methods based on the use of adjoint equations. In a subsequent paper Courtier and Talagrand (1987) show in a more realistic case how with adjoint models the information in the observations and the dynamics of the atmospheric models, can be combined to obtain an optimal initial state.

Yet another application of adjoint methods is the tuning of model parameters. In problems which are characterized by diverging trajectories and many approximately known parameters it is convenient to have a strategy to tune the model. Sentinel functions are functionals which, at first order, depend only on the parameters of the model. In particular they do not depend on the initial conditions. Their use is discussed and advocated in a rather technical paper by Lions (1990). This paper contains references to somewhat earlier papers which are mostly in French. A general description of the applications of adjoint equations in meteorology is to be found in Grasman and Houtekamer (1991).

In predictability research adjoint methods are applied because they provide an alternative to the computational problems (Lorenz 1965) of methods based on the brute force integration of error statistics. Lorenz had to integrate a model once for perturbations in all coordinates. As a result of the computation one gets all eigenvalues and eigenvectors which are necessary to describe the error at the time of a forecast. Lacarra and Talagrand (1988) proposed the use of adjoint methods for the description of the short-range evolution of forecast errors. They use a Lanczos algorithm (e.g. Parlett 1980) to find the few largest eigenvalues at a greatly reduced computational cost. Houtekamer (1991) proposed to use such an algorithm to actually produce a skill forecast. A problem with the Lanczos algorithm is that in the case of high dimensional space of the forecast error many iterations are needed to find all the relevant eigenvalues and eigenvectors. Houtekamer (1991) and Barkmeijer (1992) proposed using

adjoint methods for the description of geographically local errors. A limitation of the skill forecast to only some local aspects reduces the dimension of the errors.

The use of adjoint methods in general depends on the validity of the assumption of linear error growth. According to Lacarra and Talagrand (1988) this assumption is valid for two days. In three case-studies with a state of the art primitive equation limited-area model Vukićević (1991) showed that the linear theory gives a good approximation of the evolution of initial errors during the first 36 hours of a forecast. It appears from the experiments by Barkmeijer and Opsteegh (1992) that the adjoint method gives useful information on the local forecast error in the streamfunction at 500 mb during the first three days of a forecast. It is an entirely different problem to predict the skill of for instance the rainfall prediction. In the case of rainfall prediction the non-linearity will pose a much more serious problems also at a conceptual level. It is clear that adjoint methods break down when the errors become comparable in size to the size of the atmospheric attractor.

1.5   Data assimilation

The above text concentrates mainly on what happens during the forecast run. A skill forecast must clearly take account of this. However it is only part of the problem. A forecast starts from an analysis. The analysis is based on a numerical model of atmospheric processes, on a prior estimate of the state of the atmosphere and on recent observations. The process of computing an initial state is called data assimilation. To get error statitistics of the analysis one needs to study the sensitivity of the analysis to errors in its input. These statistics then serve as a starting point of the skill forecast. An excellent introduction to data assimilation is given in the book by Daley (1991).

If it is assumed that all the observations have Gaussian error statistics the most likely solution is given by a least squares fit of the model to the data. This well known method was developed by Gauss in the eighteenth century. Thacker (1989) explains that data assimilation with a

cost function, which is also called variational data assimilation, is in fact an application of the least squares method. From the Hessian matrix of second derivates of the cost function the statistics of the analysis error can be obtained. Work on variational data assimilation often assumes that a perfect model is available to combine measurements with prior knowledge. If this is not the case this assumption acts to distribute the model errors evenly over the entire domain, thus leading to considerable errors even in data-rich areas. A time dependent weighting of the observations only marginally improves the situation (Wergen 1992). A descent algorithm, which minimizes the cost function, must be used to find the best fit of the model to the data. Such algorithms are expensive and consequently much research is going into the development of minimization algorithms. Typical results show that at least 10 integrations with the adjoint model are necessary to obtain, with sufficient accuracy, the model trajectory which minimizes the costfunction. In some cases (Gauthier 1992) several local minima exist. In this case the assimilation may select any of these local minima. Though variational data assimilation is an area of active research it is not currently used by operational centers. The difficulties in coding the adjoint of a complete general circulation model, as well as the presence of significant model errors, have so far prohibited realistic competitive applications.

A slightly different approach is based on the Kalman filter (Kalman 1960; Kalman and Bucy 1961; Ghil et al. 1981). Theoretically the Kalman filter method gives the same results as the variational assimilation method. In general a Kalman filter allows for a more convenient treatment of model errors. In fact it is not at all clear how to incorporate the effect of model errors into the variational assimilation process (Derber 1989; Ghil et al. 1991 and Wergen 1992). The variational method is more convenient if it is decided not to integrate the error statistics. In this case it is a suboptimal version of the Kalman filter. The evolution of error statistics with the Kalman filter is not currently feasible with a general circulation model.

Though data assimilation methods have different origins and different proponents they can all be put in one common framework (Lorenc 1986). Most weather forecasting centers have adopted some kind of Optimal In-

terpolation (OI) technique (e.g: Kolmogorov 1941; Gandin 1963; Bergman 1979). Because the OI-technique uses simple rules for the evolution of error statistics it provides a feasible assimilation technique. The use of simple rules may well be preferable above the use of strong constraints, such as the perfect model assumption, which lack justification (Wergen 1992). In practice OI is rarely optimal and Daley (1991) suggested that statistical interpolation is a more appropriate term. From the point of view of a skill forecaster this use of OI is unfortunate. It is not clear to what extent OI-statistics are correct. This implies that a skill forecast has to start from statistics of unknown quality.

1.6 Outline of this thesis

It will be clear from the above discussions that skill forecasts are still in their infancy. Operational skill predictions do not exist. One is still struggling to prove that skill predictions, at any range, have any quality at all. It is not clear what the statistics of the analysis error are. The statistics of the model errors are not known and finally it is not clear how to efficiently evolve the error statistics to the time of the forecast.

In chapter 2 methods are developed to determine the variability of the predictability. The study is similar to the one by Lorenz (1965). The present atmospheric model, with 30 variables rather than 28, is only slightly larger than Lorenz's model. The main difference is in the use of methods. Adjoint models are used to find the most important error structures. These methods can be transported to state of the art models. Chapter 2 has appeared as a paper in Tellus (Houtekamer 1991).

In chapter 3, the method is extended. A simple inhomogeneous observing network is used to obtain an inhomogeneous distribution for the analysis error. It is shown that ignoring this inhomogeneity will lead to a skill forecast of low quality. Thus skill forecasters have to use the error statistics which are obtained during the data assimilation process. If one uses an average distribution to describe the analysis error one may already obtain a reasonable skill forecast. Chapter 3 will appear in Monthly Weather Review (Houtekamer 1992).

In chapter 4 a much more advanced model is used. It has 1449 variables. It is used in conjunction with the state of the art ECMWF model. The usefulness of the methods developed in chapter 2 and 3 is tested in a realistic context. It appears that the global forecast error cannot efficiently be described with adjoint methods. Global forecast errors can better be predicted with a Monte Carlo method. Weather forecasts usually have a local nature. For the description of local forecast errors adjoint methods are feasible. It appears that the distribution of the analysis error is less variable as expected from chapter 3. The observing network, which is almost time independent, determines the main structures of the distribution of the analysis error. Because the properties of the analysis error are almost constant they need to be determined only once. This reduces the computational cost of a skill forecast enormously. This chapter is concluded with a discussion of the possible impact of a high quality skill forecast. It may increase or decrease the length of a forecast with about one day. This is significant compared to the effect of other possible improvements to the forecasting system. Chapter 4 has been submitted to Monthly Weather Review.

Acknowledgement

REFERENCES

Anthes, R.A., Y.H. Kuo, D.P. Baumhefner, R.M. Errico and T.W. Bettge, 1985: Predictability of mesoscale motions. *Advances in Geophysics*, Vol. 28, *Issues in atmospheric and oceanic modelling, Part B: Weather Dynamics, Academic Press*, 159-202.

Balgovind, R., A. Dalcher, M. Ghil and E. Kalnay, 1983: A stochastic-dynamic model for the spatial structure of forecast error statistics. *Mon. Wea. Rev., 111*, 701-722.

Barkmeijer, J., 1992: Local error growth in a barotropic model, *Tellus A.*

Barkmeijer, J. and J.D. Opsteegh, 1992: Local skill prediction with a simple model. *ECMWF workshop on predictability.*

Bélanger, P.R., 1974: Estimation of noise covariance matrices for a linear time-varying stochastic process. *Automatica, 10,* 267-275.

Bergman, K.H., 1979: Multivariate analysis of temperatures and winds using optimum interpolation. *Mon. Wea. Rev.,107,* 1423-1444.

Boer, G.J., 1984: A spectral analysis of predictability and error in an operational forecast system. *J. Atmos. Sci., 112,* 1183-1197.

Branković, č., T.N. Palmer, F. Molteni, S. Tibaldi and U. Cubasch, 1990: Extended-range predictions with ECMWF models: Time-lagged ensemble forecasting. *Quart. J. Roy. Meteor. Soc., 166,* 867-912.

Cacuci, D.G., 1981: Sensitivity theory for nonlinear systems. I. Nonlinear functional analysis approach. *J. Math. Phys., 22,* 2794-2802.

Charney, J., R. Fleagle, V. Lally, H. Riehl and D. Wark, 1966: The feasibility of a global observation and analysis experiment. *Bull. Am. Meteorol. Soc., 47,* 200-220.

Courtier, P., and O. Talagrand, 1987: Variational assimilation of meteorological observations with the adjoint vorticity equation. II: Numerical results. *Quart. J. Roy. Meteor. Soc., 113,* 1329-1347.

Daley, R., 1991: *Atmospheric Data Analysis,* Cambridge atmospheric and space science series, Cambridge University Press.

Dee, D.P., S.E. Cohn and M. Ghil, 1985a: Systematic estimation of forecast and observation error covariances in four dimensional data assimilation. Preprints, *Seventh conference on numerical weather prediction,* Boston, Amer. Meteor. Soc., pp 9-16.

Dee, D.P., S.E. Cohn, Dalcher, A. and M. Ghil, 1985b: An efficient algorithm for estimating noise covariances in distributed systems. *IEEE Trans. Automat. Contr., AC-30,* 1057-1065.

Derber, J.C., 1989: A variational continuous assimilation technique. *Mon. Wea. Rev., 117,* 2437-2446.

Epstein, E.S., 1969: Stochastic dynamic prediction. *Tellus*, *21*, 739-759.

Errico, R., and D. Baumhefner, 1987: Predictability experiments using a high-resolution limited-area model. *Mon. Wea. Rev.*, *115*, 488-504.

Errico, R., and T. Vukićević, 1992: Sensitivity analysis using an adjoint of the PSU/NCAR mesoscale model. *Mon. Wea. Rev.*, *120*, 1644-1660.

Farrell, B.F., 1989: Optimal excitation of baroclinic waves. *J. Atmos. Sci.*, *46*, 1193-1206.

Farrell, B.F., 1990: Small error dynamics and the predictability of atmospheric flows. *J. Atmos. Sci.*, *47*, 2409 - 2416.

Fleming, R., 1971: On stochastic-dynamic prediction. *Mon. Wea. Rev.*, *99*, 851-872.

Freiberger, W.F., and U. Grenander, 1965: On the formulation of statistical meteorology. *Review of the International Statistical Institute*, *33*, 59-86.

Gandin, L.S., 1963: Objective analysis of meteorological fields. *Gidrometeorologicheskoe Izdatelstvo* , Leningrad.   English translation by: Israel Program for Scientific Translations, Jerusalem, 1965, 242pp. [NTIS N6618047, Library of Congress QC 996.G3313].

Gauthier, P., 1992: Chaos and quadri-dimensional data assimilation: a study based on the Lorenz model. *Tellus*, *44A*, 2-17.

Ghil, M., S. Cohn, J. Tavantzis, K. Bube, and E. Isaacson, 1981: Applications of estimation theory to numerical weather prediction. in *Dynamic Meteorology: Data Assimilation Methods*, pp 139-224, Springer-Verlag, New York, 1981.

Ghil, M. and P. Malanotte-Rizzoli, 1991: Data assimilation in meteorology and oceanography. *Adv. Geophys.*, *33*, 141-266.

Goldhirsch, I., P. Sulem and S.A. Orszag, 1987: Stability and Lyapunov stability of dynamical systems: a differential approach and a numerical method. *Physica*, *27D*, 311-337.

Grasman, J., and P.L. Houtekamer, 1991: Methods for improving the prediction of dynamical processes with special reference to the atmo-

spheric circulation. Wageningen Agricultural University, *Technical Note 91-06*.

Hall, M.C.G., D.G. Cacuci and M.E. Schlesinger, 1982: Sensitivity analysis of a radiative convective model by the adjoint method. *J. Atmos. Sci.*, *39*, 2038-2050.

Hall, M.C.G., and D.G. Cacuci, 1983: Physical interpretation of the adjoint functions for sensitivity analysis of atmospheric models. *J. Atmos. Sci.*, *40*, 2537-2546.

Hoffman, R.N., and E. Kalnay, 1983: Lagged average forecasting, an alternative to Monte Carlo forecasting. *Tellus*, *35A*, 100-118.

Houtekamer, P., 1991: Variation of the predictability in a low order spectral model of the atmospheric circulation. *Tellus*, *43 A*, 177-190.

Houtekamer, P., 1992: The quality of skill forecasts for a low order spectral model. Accepted by *Mon. Wea. Rev.*

Kalman, R., 1960: A new approach to linear filtering and prediction problems. *Trans ASME, Ser. D, J. Basic Eng.*, *82*, 35-45.

Kalman, R., and R. Bucy, 1961: New results in linear filtering and prediction theory. *Trans ASME, Ser. D, J. Basic Eng.*, *83*, 95-108.

Kalnay, E., and A. Dalcher, 1987: Forecasting forecast skill. *Mon. Wea. Rev.*, *115*, 349-356.

Kaplan, J.L., and J.A. Yorke, 1979: Chaotic behavior of multidimensional difference equations. In: *Functional differential equations and approximations of fixed points*, (eds. H.O. Peitgen and H.O. Walther), Heidelberg- New York, Springer, 228-237.

Kleijnen, J.P.C., 1974: *Statistical techniques in simulation, Part I*, Marcel Dekker, NY, 285 pp.

Kolmogorov, A., 1941: Interpolated and extrapolated stationary random sequences. *Izvestia an SSSR, seriya mathematicheskaya 5(2)*, 85-95.

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

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

Leith, C.E., 1971: Atmospheric predictability and two-dimensional turbulence. *J. Atmos. Sci.*, *28*, 145-161.

Leith, C.E., and R.H. Kraichnan, 1972: Predictability of turbulent flows. *J. Atmos. Sci.*, *29*, 1041-1058.

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

Lilly, D.K., and E.L. Peterson, 1983: Aircraft measurements of atmospheric kinetic energy spectra. *Tellus*, *35A*, 379-382.

Lilly, D.K., 1985: Theoretical predictability of small-scale motions. In: *Turbulence and predictability in geophysical fluid dynamics and climate dynamics* (eds. M. Ghil, R. Benzi and G. Parisi). Amsterdam, North-Holland, 281-289.

Lions, J.L., 1990: Sentinels and stealthy perturbations, *Proceedings International Symposium on Assimilation of Observations in Meteorology and Oceanography*, Clermont-Ferrand, World Meteorological Organization, I3-18.

Lorenc, A.C., 1986: Analysis methods for numerical weather prediction. *Quart. J.R. Meteor. Soc.*, *112*, 1177-1194.

Lorenz, E.N., 1963: Deterministic nonperiodic flow. *J. Atmos. Sci.*, *20*, 130-141.

Lorenz, E.N., 1965: A study of the predictability of a 28-variable atmospheric model. *Tellus*, *17*, 321-333.

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

Lorenz, E.N., 1982: Atmospheric predictability experiments with a large numerical model. *Tellus*, *34*, 505-513.

Lorenz, E.N., 1990: Effects of analysis and model errors on routine weather forecasts, in ECMWF seminar proceedings: *Ten years of*

medium-range weather forecasting. 4-8 September 1989. Vol I, ECMWF, Reading, UK.

Marchuk, G.I., 1974: *The numerical solution of problems of atmospheric and oceanic dynamics* (in Russian), 387 pp., Gidrometeoizdat, Leningrad, USSR. (English translation, Rainbow Systems, Alexandria, Va.) (Russian version 1967).

McKay, M.D., W.J. Conover and R.J. Beckman, 1979: A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, *21*, 239-245.

Palmer, T.N., and S. Tibaldi, 1988: On the prediction of forecast skill, *Mon. Wea. Rev.. 116*, 2453-2480.

Parlett, P., 1980: *The symmetric eigenvalue problem.* Series in computational mathematics, Prentice Hall, Englewood Cliffs, New Jersey.

Pitcher, E.J., 1977: Application of stochastic dynamic prediction to real data. *J. Atmos. Sci.*, *34*, 1-21.

Smagorinsky, J., 1963: General circulation experiments with the primitive equations. I. The basic experiment. *Mon. Wea. Rev.*, *91*, pp 99-164.

Smagorinsky, J., 1969: Problems and promises of extended range forecasting. *Bull. Am. Meteorol. Soc.*, *50*, 286-311.

Stein, M., 1987: Large sample properties of simulations using latin hypercube sampling. *Technometrics*, *29*, 143-151.

Tennekes, H., A.P.M. Baede, and J.D. Opsteegh, 1987: Forecasting forecast skill. In *Workshop on predictability in the medium and extended range*, 17-19 March 1986, ECMWF, p 277-302.

Talagrand, O., and P. Courtier, 1987: Variational assimilation of meteorological observations with the adjoint vorticity equation I: theory. *Quart. J. Roy. Meteor. Soc.*, *113*, 1311-1328.

Thacker, W.C., 1989: The role of the Hessian matrix in fitting models to measurements. *J. Geophys. Res.*, *94*, 6177-6196.

Thiébaux, H.J., and L.L. Morone, 1990: Short-term systematic errors in global forecasts: their estimation and removal. *Tellus*, *42A*, 209-229.

Thompson, P.D., 1985: Prediction of the probable errors of predictions. *Mon. Wea. Rev.*, *113*, 248-259.

Thompson, P.D., 1986: A simple approximate method of stochastic-dynamic prediction for small initial errors and short range. *Mon. Wea. Rev.*, *114*, 1709-1715.

Tracton, M.S., K. Mo, W. Chen, E. Kalnay, R. Kistler and G. White, 1989: Dynamical extended range forecasting (DERF) at the National Meteorological Center. *Mon. Wea. Rev.*, *117*, 1604-1635.

Tribbia, J.J., and D.P. Baumhefner, 1988: The reliability of improvements in deterministic short-range forecasts in the presence of initial state and modeling deficiencies. *Mon. Wea. Rev.*, *116*, 2276-2288.

Vukićević, T., and R.M. Errico, 1990: The influence of artificial and physical factors upon predictability estimates using a complex limited-area model. *Mon. Wea. Rev.*, *118*, 1460-1482.

Vukićević, T., 1991: Nonlinear and linear evolution of initial forecast errors. *Mon. Wea. Rev.*, *119*, 1602-1611.

Wergen, W., 1992: The effect of model errors in variational assimilation, To appear in *Tellus*.

Yoden, S., and M. Nomura, 1992: Finite-time Lyapunov stability analysis and its application to atmospheric predictability, submitted to *J. Atmos. Sci.*

Chapter 2

Variation of the predictability

Abstract

Temporal variations in the predictability of a simple atmospheric model are studied. It is a spectral two-layer quasi-geostrophic hemispheric model and it is truncated at T5. The model has chaotic properties which is illustrated by four positive Lyapunov exponents and a smooth power spectrum of the first Empirical Orthogonal Function (EOF). This chaotic behaviour is responsible for the eventual separation of initially nearby phase points.

The adjoint of the tangent linear equations is used to obtain the few directions in which errors grow most rapidly. The distribution function for the errors can be estimated with the growth rates in these directions. Comparison with a Monte Carlo method shows satisfactory agreement. With a 100-day run it is shown how the predictability varies over the attractor.

It is discussed how, and under what conditions, the method can be used in a Global Circulation Model (GCM). An estimate for the dimension of error growth is obtained from a dimension derived from an EOF analysis. This estimate implies that information on the probability distribution of errors can be obtained at about 150 times the computational cost of a single model run.

2.1   Introduction

Weather forecasts tend to go wrong after some time. There is a consensus
that detailed deterministic predictions are limited to a range of about 10
days. One reason is the presence of systematic errors that are introduced
by imperfections of the forecast model. Another more fundamental reason
is the intrinsic deterministic growth of errors. The growth of systematic
and intrinsic errors makes the predicted weather gradually diverge from
the subsequently observed weather. In general systematic errors can be
reduced by improving the physical description which is used for the model.
Systematic errors with a low frequency can be reduced by post-processing
of model output (Thiébaux and Morone 1990). Intrinsic errors are caused
by the unavoidable sensitivity of the model to small errors in the ini-
tial conditions. Their possible importance was first discussed by Lorenz
(1963). This chapter deals with intrinsic errors.

Some predictions have a higher quality than others. To make effective
use of high quality forecasts and to be able to warn against low quality
forecasts a measure for the temporal predictability is needed. Stochastic
dynamic predictions (Epstein 1969) require the integration of the mean
values of the model variables and of the corresponding covariance ma-
trix. Because in addition to the N equations for the mean field, there
are $N(N + 1)/2$ equations for the evolution of the second moments, this
method may not be feasible in practice. Because only the first and second
moments are considered the method can only be applied to the short range
evolution of initially small errors. A probably less expensive method to
study variations in the predictability of the atmosphere would be to take a
sufficiently large ensemble with perturbed initial conditions (Leith 1974).
The divergence of the ensemble is a measure of the quality of the forecast.
This method should work, even in the regime of non-linear error growth.
A review of methods to estimate the predictability of forecasts in the 5 to
10 days range, where non-linearity is important, is given by Palmer and
Tibaldi (1988).

It is shown in a study by Lacarra and Talagrand (1988) with an f-
plane shallow-water model that the linear approximation for error growth

can be valid for ranges up to about 48 hours in meteorologically realistic situations. The linear approximation greatly simplifies the analysis. If the evolution of the errors is linear it is possible to make use of the adjoint of the tangent linear equations. Le Dimet and Talagrand (1986) have given a rather general presentation of the use of adjoint equations in the context of data assimilation. An example of a minimization of a distance function using the adjoint is given by Courtier and Talagrand (1987). Their model contains 231 independent degrees of freedom and is integrated for 24 hours. Their distance function is a measure for the agreement between a model-run and a set of verifying observations. They find strong indications that their distance function varies quadratically with respect to the initial conditions, implying that the tangent linear equations of their model are indeed sufficient to describe the 24-hour evolution of forecast errors. Consequently the minimization leads to the trajectory which shows the best agreement with the available observations. Thus, the adjoint equations can be used both for data assimilation and for error growth studies.

The present goal is to determine confidence limits for short range numerical forecasts. The probability distribution of the errors in an ensemble is obtained with a Monte Carlo method. This brute-force experiment is done in a 30-dimensional baroclinic model. It is attempted to reproduce the probability distribution with less time consuming methods. The error growth is described with the tangent linear equations. Because the model is small all growth rates can be computed rigorously. The growth rates have a very steep spectrum. This makes it possible to approximate the probability distribution for the errors with only the few dominating growth rates. The dominating growth rates correspond to the largest eigenvalues of a symmetric matrix. A Lanczos algorithm (see e.g. Parlett 1980) is used to estimate these eigenvalues. A Lanczos algorithm requires only a few matrix vector multiplications. Individual multiplications are done using both the tangent linear equations and the adjoint of the tangent linear equations. The applicability of the Lanczos algorithm depends on the number of relevant eigenvalues or, putting it in other words, on the dimensionality of the errors. A relationship between the dimensionality

of the instantaneous circulation and the dimensionality of the errors is
given. If the same relationship holds for a GCM then the algorithm can
be applied to it.

2.2   Description of the model

The model is a highly truncated two-layer hemispheric spectral model.
Both layers have 3 zonal and 6 wave modes. It is comparable with the
highly truncated two-layer spectral channel model of Reinhold and Pier-
rehumbert (1982) and the standard two layer model discussed by Holton
(1979). The notation introduced by Lorenz (1960) is used: ψ_1 is the 250-
mb streamfunction, ψ_3 is the 750-mb streamfunction, $\psi = (\psi_1 + \psi_3)/2$ is
the interpolated 500-mb streamfunction and $\tau = (\psi_1 - \psi_3)/2$ is the 250-750
mb thickness. The set of equations is:

$$\frac{\partial}{\partial t}\Delta\psi_1 = -J(\psi_1, \Delta\psi_1 + f) + \frac{f_0}{\Delta p}\omega_2 \tag{2.1}$$

$$\frac{\partial}{\partial t}\Delta\psi_3 = -J(\psi_3, \Delta\psi_3 + f) - \frac{f_0}{\Delta p}\omega_2$$
$$\qquad -f_0 J(\psi_3, h) - C\Delta(\psi_3 - \psi_3^*) \tag{2.2}$$

$$\frac{\partial}{\partial t}2\tau = -J(\psi, 2\tau) + \frac{\sigma\Delta p}{f_0}\omega_2 + 2Q(\tau^* - \tau) \tag{2.3}$$

Where $J(A, B)$ denotes the jacobian operator:

$$J(A, B) = \frac{1}{a^2}\left(\frac{\partial A}{\partial \lambda}\frac{\partial B}{\partial \mu} - \frac{\partial A}{\partial \mu}\frac{\partial B}{\partial \lambda}\right) \tag{2.4}$$

Here λ is the geographic longitude, μ is the sine of the geographic latitude
and a is the radius of the earth. The Eqs. (2.1)-(2.4) are nondimension-
alized using the radius a as unit of length and the inverse of the angular
speed of rotation of the earth as unit of time. The Coriolis parameter is
f. Its value at a latitude of 45^0 is f_0. The Ekman damping coefficient is
C. The Ekman damping is towards the 750 mb-streamfunction ψ_3^*, which
corresponds to a zonal westerly wind in the lower layer. A 250-750 mb
thickness is enforced with the function τ^*. The function τ^* corresponds to
a meridional temperature gradient, and Q is the coefficient for the cooling

towards this gradient. The mountain height is given by the function h. The static stability parameter is σ. The pressure difference Δp between the two layers is 500 mb. The vertical motion dp/dt at the 500 mb-level is ω_2. Eqs. (2.1) and (2.2) are the vorticity equations for respectively the 250-mb and the 750-mb layer. Eq. (2.3) is the thermodynamic energy equation. The radius of deformation Λ^{-1} is defined by:

$$\Lambda^2 = \frac{f_0^2}{\sigma(\Delta p)^2} \tag{2.5}$$

After the elimination of ω_2 from the Eqs (2.1), (2.2) and (2.3) one obtains:

$$\frac{\partial}{\partial t}\Delta\psi = -J(\psi, \Delta\psi) - J(\tau, \Delta\tau) - J(\psi, f) +$$
$$f_0 J(\tau - \psi, h)/2 + C\Delta(\tau - \psi + \psi_3^*)/2 \tag{2.6}$$

$$\frac{\partial}{\partial t}(\Delta - 2\Lambda^2)\tau = -J(\psi, \Delta\tau) - J(\tau, \Delta\psi) - J(\tau, f) +$$
$$f_0 J(\psi - \tau, h)/2 + C\Delta(\psi - \tau - \psi_3^*)/2 +$$
$$2\Lambda^2 J(\psi, \tau) - 2\Lambda^2 Q(\tau^* - \tau) \tag{2.7}$$

The streamfunctions are projected on a basis of eigenfunctions of the Laplace operator Δ. These eigenfunctions are:

$$Y_{m,n}(\lambda, \mu) = P_{|m|,n}(\mu)e^{im\lambda} \tag{2.8}$$

Here $|\cdot|$ denotes the absolute value and $P_{m,n}(\mu)$ denotes associated Legendre functions of the first kind, which are defined by:

$$P_{m,n}(\mu) = \left(\frac{(2n+1)(n-m)!}{(n+m)!}\right)^{1/2} \frac{(1-\mu^2)^{m/2}}{2^n n!} \frac{d^{m+n}}{d\mu^{m+n}}(\mu^2 - 1)^n \tag{2.9}$$

The streamfunction is approximated as follows:

$$\psi(\lambda, \mu, t) = \sum_{n=1}^{5} \sum_{\substack{m = -n \\ m + n = \text{odd}}}^{+n} \psi_{m,n}(t)Y_{m,n}(\lambda, \mu) \tag{2.10}$$

The restriction to modes with $m + n$ odd excludes currents across the equator. It makes the model hemispheric. The truncation at $n = 5$ is

called a triangular T5 truncation. The three modes with m equal to zero are the zonal modes. They are a function of only the geographic latitude. The coefficient $\psi_{0,n}(t)$ have only a real part. The remaining six wave-modes have a real and an imaginary part. Consequently 15 coefficients are needed to describe the 500-mb streamfunction and another 15 coefficients to describe the 250 - 750 mb shear streamfunction, making the total system 30 dimensional. The moderate number of modes is a compromise between a completely unrealistic low dimensional model, which can be handled by any computational method, and a higher dimensional model in which rigorous testing becomes computationally expensive.

A Galerkin projection of Eqs. (2.6) and (2.7) on these eigenfunctions gives a system of 30 coupled non-linear ordinary differential equations. The right hand side of this equation is evaluated using the interaction coefficient method (Platzman 1960). The interaction coefficients are computed with the analytical expressions given by Silberman (1954). An alternative method is the transform method (Orszag 1970). This method becomes more efficient if the truncation is made less severe (Bourke et al. 1977).

The Ekman damping coefficient C and the cooling coefficient Q are given a value of $2\pi/100$ a day. Thus the e-folding time for both effects is about 16 days. These unrealistically low values of C and Q make the circulation attractor high dimensional. If they are chosen equal to zero the circulation is no longer attracted towards realistic patterns and spreads out in the 30 dimensional phase space.

The forcings are in the $Y_{0,1}$ mode only. The strength of τ^* corresponds to a westerly thermal wind of 28 m/s and ψ_3^* corresponds to a lower layer westerly wind of 5 m/s. The forcing ψ_3^* can be considered to reproduce the effect of surface baroclinicity. Without this forcing the 750-mb westerlies are very weak. Due to the effects of the topography and the baroclinicity the enforced westerly winds become unstable.

The radius of deformation Λ^{-1} is 667 km. The model is not very sensitive to the exact value of this parameter.

The topography is described with the $Y_{2,3}$ mode. It stands for the influence of the Rocky Mountains and the Tibetan plateau. The topography
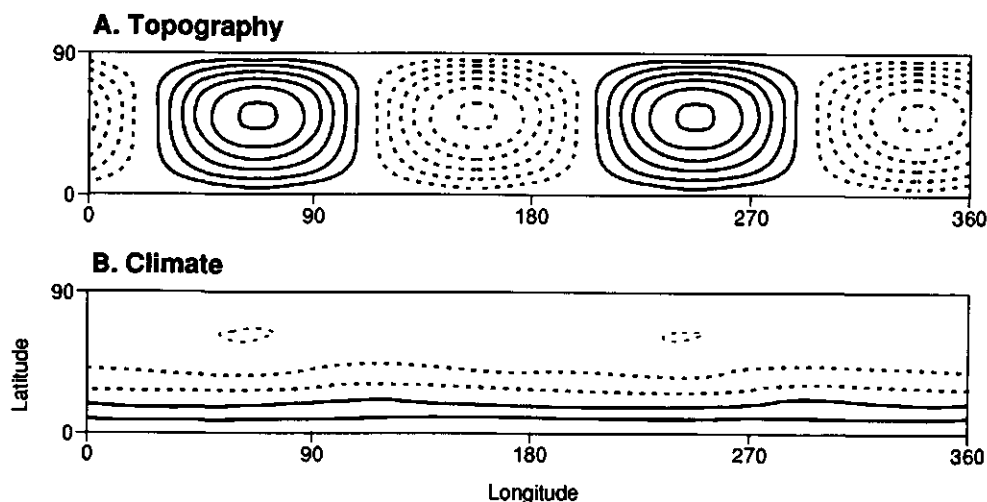
A. Topography

B. Climate

Figure 2.1: (a) Contours of the topography. The projection is area preserving. The vertical coordinate has been scaled with the sine of the geographic latitude. The horizontal coordinate is the geographic longitude. Negative heights are indicated with dashed lines. The contour interval is 250 meters. (b) Contours of the geopotential height at 500mb for the 30-year climate. The contour interval is 300 meters. Projection as in (a).

serves both to destabilize the enforced zonal flow and to locate preferential circulation patterns at fixed geographical positions. The amplitude of the mountain height is 1.56 km. In figure 2.1 both the topography and the barotropic component of the climate are shown. The climate is computed with a 30-year run. Since the thermal forcing is in the $Y_{0,1}$ mode and the topographic forcing is in the $Y_{2,3}$ mode it is possible, as is observed in figure 2.1, to have a model climate which is π-periodic in longitude. Such a climate is restricted to the $2 * (3 + 2 * 3) = 18$ dimensional subspace (m even, $m + n$ odd) of the $2 * (3 + 2 * 6) = 30$ dimensional space ($m + n$ odd) of the model. Because the symmetry is broken in the specification of the initial state the dynamics themselves are not confined to the 18 dimensional subspace.

The most significant component of the model circulation is an eastwards moving $Y_{4,5}$ wave. The wave is accelerated and decelerated by the varying intensity of the zonal wind. During short time intervals it may move westwards. All other waves continuously move westwards. The influence of the zonal wind is insufficient to keep them at fixed geographical longitude.

Figure 2.2 shows the power spectrum of the first EOF. In section 2.5 it is explained how the EOFs are determined. A power spectrum shows how the energy of a signal is distributed as a function of frequency. A periodic signal shows only isolated peaks in its power spectrum. A chaotic signal shows a much smoother spectrum (Schuster 1988). Figure 2.2 was obtained from four 2048-day integrations. The power spectrum contains most energy on periods between 9 and 30 days. This is a first indication that the model shows chaotic behaviour. The second EOF has an almost identical power spectrum. The first and second EOF approximately describe the travelling $Y_{4,5}$ wave.

A chaotic trajectory is, by definition, not periodic. It is verified numerically that solutions do not return negligibly close to previously visited phase points within a short time. This behaviour suggests that the asymptotic orbit defines a strange attractor. A strange attractor is characterized by a number of growth directions for errors, at least one neutral direction in which errors neither grow nor shrink and a number of directions with negative growth.

All 30 Lyapunov exponents are computed with the algorithm given by Wolf et al (1985). A basis of 30 orthonormal vectors is integrated forward with the tangent linear equations. After every 1.6 days a Gram-Schmidt orthonormalisation is applied. The change in length of the vectors fixes the Lyapunov exponents. After a model integration of 3.5 years four positive Lyapunov exponents, one exponent zero and 25 negative exponents have appeared.

Kaplan and Yorke (1979) conjectured the following general formula for the dimension of arbitrary strange attractors:

$$D_{KY} = j + \frac{\sum_{i=1}^{j} \alpha_j}{|\alpha_{j+1}|} \tag{2.11}$$

Power spectrum of the first EOF
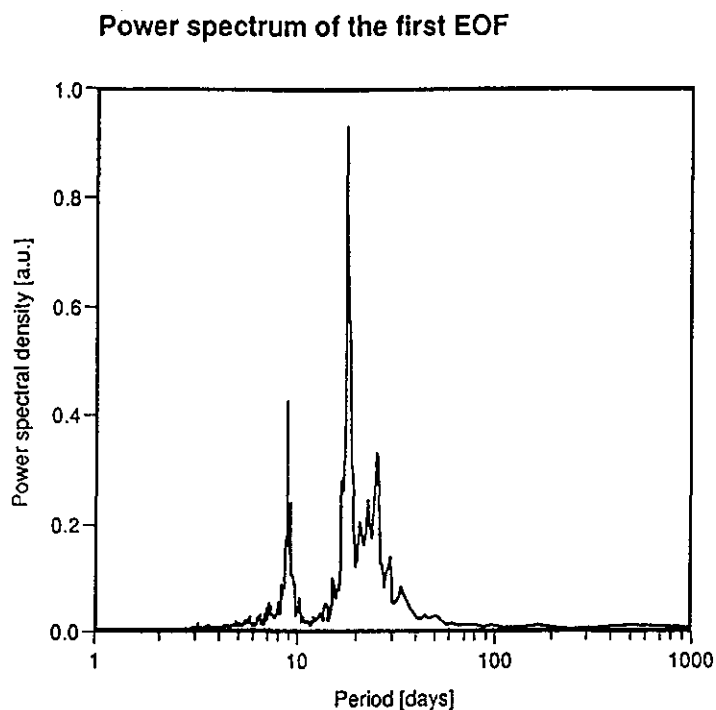
Figure 2.2: Powerspectrum of the first EOF. Vertical units are arbitrary. The area under the curve is a measure of the amount of energy contained in that part of the spectrum.

Here the Lyapunov exponents are ordered $\alpha_1 > \alpha_2 > \ldots > \alpha_{30}$, j is the largest integer for which $\sum_{i=1}^{j} \alpha_j > 0$ and D_{KY} is, in almost all cases, equal to the information dimension. In some cases D_{KY} is indeed shown to be equal to the information dimension within the precision of the estimate. In other non-generic cases the conjecture does not hold. The dimesion D_{KY} is called the Kaplan-Yorke dimension. For the present model a Kaplan-Yorke dimension of 10.7 ± 0.2 is found. Due to these observations it is assumed that the asymptotic dynamics of the system is characterized by a strange attractor and consequently complex unpredictable behaviour is expected to appear.

2.3   Determination of growth rates

It is assumed that errors in the initial conditions are small. As a consequence the initial evolution of errors can be described by a linearization of Eqs. (2.6) and (2.7) with respect to the unperturbed solution. Two methods for the computation of the growth rates of the errors are given. The first method uses the tangent linear equations and gives all 30 growth rates. The second method uses both the tangent linear equations and their adjoint. It gives only the largest growth rates, but requires less computations. The latter method can also be applied to a larger model. The tangent linear equations can be written as:

$$\frac{\partial}{\partial t}\begin{pmatrix} \delta\Delta\psi \\ \delta\Delta\tau \end{pmatrix} = S^{-1}L\begin{pmatrix} \delta\Delta\psi \\ \delta\Delta\tau \end{pmatrix} \tag{2.12}$$

The operators S^{-1} and L are:

$$S^{-1} = \begin{pmatrix} I & 0 \\ 0 & \Delta(\Delta - 2\Lambda^2)^{-1} \end{pmatrix} \tag{2.13}$$

$$L\begin{pmatrix} \delta\Delta\psi \\ \delta\Delta\tau \end{pmatrix} = \begin{pmatrix} -J(\delta\psi,\Delta\psi) - J(\psi,\delta\Delta\psi) - J(\delta\tau,\Delta\tau) \\ -J(\delta\psi,\Delta\tau) - J(\psi,\delta\Delta\tau) - J(\delta\tau,\Delta\psi) \end{pmatrix}$$

$$+ \begin{pmatrix} -J(\tau,\delta\Delta\tau) - J(\delta\psi,f) \\ -J(\tau,\delta\Delta\psi) - J(\delta\tau,f) \end{pmatrix}$$

$$+ \begin{pmatrix} -(f_0/2)J(\delta\psi - \delta\tau,h) + (C/2)(\delta\Delta\tau - \delta\Delta\psi) \\ +(f_0/2)J(\delta\psi - \delta\tau,h) - (C/2)(\delta\Delta\tau - \delta\Delta\psi) \end{pmatrix}$$

$$+2\Lambda^2\begin{pmatrix} 0 \\ J(\delta\psi,\tau) - J(\delta\tau,\psi) + Q\delta\tau \end{pmatrix} \tag{2.14}$$

For the development of the adjoint model distances are computed with a kinetic energy norm. The kinetic energy of a field is given by:

$$\begin{aligned} K &= \int_\Sigma (\nabla\psi \cdot \nabla\psi + \nabla\tau \cdot \nabla\tau)d\Sigma \\ &= \sum_{n=1}^{5} \sum_{\substack{m=-n \\ m+n = \text{odd}}}^{+n} n(n+1)\left\{|\psi_{m,n}(t)|^2 + |\tau_{m,n}(t)|^2\right\} \end{aligned}$$

The factor $n(n+1)$ in this expression is removed with the rescaling:

$$\tilde{\psi}_{m,n}(t) = (n(n+1))^{1/2}\psi_{m,n}(t) \tag{2.15}$$

$$\tilde{\tau}_{m,n}(t) = (n(n+1))^{1/2}\tau_{m,n}(t) \tag{2.16}$$

The real and imaginary parts of $\tilde{\psi}_{m,n}$ and $\tilde{\tau}_{m,n}$ form one single real valued vector q (of length 30). The vector q contains the coeffients q_i of the state vector with respect to a 30-dimensional basis $\{e_1, e_2, \ldots, e_{30}\}$ which is orthonormal with respect to the kinetic energy inner product:

$$(e_i, e_j) = \delta_{ij} \tag{2.17}$$

The kinetic energy or length squared associated with the vector q becomes:

$$R^2 = \sum_{i=1}^{30} q_i^2 = (q, q) \tag{2.18}$$

In the rest of this chapter the kinetic energy inner product is used to define lengths or error-magnitudes.

An initial error vector u(0) is integrated forward with the tangent linear equations to give a vector u(t). If an ensemble with more than 30 initial vectors is to be integrated then it is more efficient to compute the 30 × 30 matrix A defined by:

$$u(t) = Au(0) \tag{2.19}$$

The matrix A is time dependent. For convenience of notation A is written rather than A(t). The columns of A follow from the forward integration of the unit-vectors of the basis. To compute column i one starts with a unit vector e_i. This vector is transformed to streamfunction coefficients using Eqs. (2.15), (2.16), (2.17) and (2.18). It is integrated with Eq. (2.12) and finally it is transformed back to the basis $\{e_1, e_2, \ldots, e_{30}\}$. The kinetic energy of the error at time t is given by:

$$(u(t), u(t)) = (Au(0), Au(0)) = (A^*Au(0), u(0)) \tag{2.20}$$

A^* is the adjoint matrix of A, which is just the transpose matrix:

$$A^* = A^T \tag{2.21}$$

The eigenvalues of A^*A are all positive. In order of decreasing magnitude they are called: $\lambda_1^2, \lambda_2^2, \ldots, \lambda_{30}^2$. The eigenvectors of A^*A are orthogonal. Once the eigenvalues and eigenvectors of A^*A have been obtained the statistics for the error at time t can be determined. If for instance, one starts with an ensemble of random initial error vectors of unit length, the average error kinetic energy after time t is given by (Lorenz 1985):

$$\overline{(u(t), u(t))} = \sum_{i=1}^{30} \frac{\lambda_i^2}{30} \tag{2.22}$$

The model analysed in the present chapter is sufficiently small so that the above formulation can be applied. This is not the case for models with a larger number of degrees of freedom. In the following paragraphs an alternative method is sketched, which does not require the computation of the entire matrix A and which makes use of the adjoint equations. Basic facts about adjoint operators and a specific example with the adjoint of the vorticity equation can be found in Talagrand and Courtier (1987). Because the same inner product is used the present adjoint model is comparable to theirs.

The adjoint of the tangent linear equation (Eq. (2.12)) can be written as:

$$\frac{\partial}{\partial t} \begin{pmatrix} \delta\Delta\psi' \\ \delta\Delta\tau' \end{pmatrix} = -L^*(S^{-1})^* \begin{pmatrix} \delta\Delta\psi' \\ \delta\Delta\tau' \end{pmatrix} \tag{2.23}$$

$(S^{-1})^*$ and L^* are the adjoints of the operators S^{-1} and L given in Eqs. (2.13) and (2.14). The adjoint operators read:

$$\begin{pmatrix} \delta\Delta\psi'' \\ \delta\Delta\tau'' \end{pmatrix} = (S^{-1})^* \begin{pmatrix} \delta\Delta\psi' \\ \delta\Delta\tau' \end{pmatrix}$$

$$= \begin{pmatrix} I & 0 \\ 0 & (\Delta - 2\Lambda^2)^{-1}\Delta \end{pmatrix} \begin{pmatrix} \delta\Delta\psi' \\ \delta\Delta\tau' \end{pmatrix} \tag{2.24}$$

$$L^* \begin{pmatrix} \delta\Delta\psi'' \\ \delta\Delta\tau'' \end{pmatrix} = 2\Lambda^2 \begin{pmatrix} J(\tau, \delta\tau'') \\ -J(\delta\psi, \tau'') + Q\delta\tau'' \end{pmatrix}$$

$$+ \begin{pmatrix} -J(\Delta\psi, \delta\psi'') + \Delta J(\psi, \delta\psi'') - J(\Delta\tau, \delta\tau'') \\ -J(\Delta\tau, \delta\psi'') + \Delta J(\tau, \delta\psi'') - J(\Delta\psi, \delta\tau'') \end{pmatrix}$$

$$+ \left(\begin{array}{c} \Delta J(\tau, \delta\tau'') - J(f, \delta\psi'') \\ \Delta J(\psi, \delta\tau'') - J(f, \delta\tau'') \end{array} \right)$$

$$+ \left(\begin{array}{c} -(f_0/2)J(h, \delta\psi'' - \delta\tau'') + (C/2)(\delta\Delta\tau'' - \delta\Delta\psi'') \\ (f_0/2)J(h, \delta\psi'' - \delta\tau'') - (C/2)(\delta\Delta\tau'' - \delta\Delta\psi'') \end{array} \right) \tag{2.25}$$

It can be seen from Eqs. (2.24) and (2.25) that the complexity of the adjoint model is comparable to the complexity of the tangent linear model (Eqs. (2.13) and (2.14)). Backward integration with the adjoint model replaces the multiplication with the matrix A^* in the same way as forward integration with the tangent linear equations replaces the multiplication with the matrix A. For a large model this provides a practical tool for obtaining the largest eigenvalues of the matrix A^*A.

A Lanczos algorithm with orthogonalization (see e.g. Parlett 1980) is used to find the largest eigenvalues of A^*A. A description of the Lanczos algorithm is given below. In a series of j iterative steps a basis $\{\Delta_1, \Delta_2, \ldots, \Delta_j\}$ is produced. This basis approximates the basis of eigenvectors $\{v_1, v_2, \ldots, v_j\}$. The iteration starts with a random direction vector Δ_1 of unit length:

$$\Delta_1 = \sum_{k=1}^{30} b_k v_k = \sum_{k=1}^{30} a_k e_k \tag{2.26}$$

$$(\Delta_1, \Delta_1) = 1$$

Notice that at this point the coefficients b_k in Eq. (2.26) are still not known. The coeffients a_k give the random direction. If the operator A^*A is applied to the vector Δ_1 the coefficient for the k'th vector increases with λ_k^2:

$$A^*A\Delta_1 = \sum_{k=1}^{30} b_k \lambda_k^2 v_k \tag{2.27}$$

This property suggests using the matrix A^*A for getting the new direction Δ_2 and subsequent new directions Δ_{i+1}. A Gram-Schmidt orthonormalisation is used to make a new direction Δ_{i+1} orthonormal to the previous directions $\Delta_1, \Delta_2, \ldots, \Delta_i$:

$$\Delta_{i+1}' = A^*A\Delta_i - \sum_{k=1}^{i} (A^*A\Delta_i, \Delta_k)\Delta_k \tag{2.28}$$

$$\Delta_{i+1} \;=\; \frac{\Delta'_{i+1}}{(\Delta'_{i+1},\Delta'_{i+1})^{1/2}} \tag{2.29}$$

After every iteration step the eigenvalues of A^*A can be estimated with the eigenvalues of the tridiagonal j x j-matrix $E_j^T E_j$.

$$E_j = [A\Delta_1, A\Delta_2, \ldots, A\Delta_j] \tag{2.30}$$

The estimate is exact if the eigenvalues $\lambda_j^2, \lambda_{j+1}^2, \ldots, \lambda_{30}^2$ are equal to zero. The largest eigenvalue is estimated most accurately. In a typical case 4 iterations are needed to find the largest eigenvalue. Three additional iterations suffice to find the second largest eigenvalue. Successive iterations give one more accurate estimate per iteration. The maximum number of 30 iteration steps gives all 30 eigenvalues.

The only time consuming step in the iteration procedure is the computation of $A^*A\Delta_i$. The computation of $A\Delta_i$ is just as expensive as one model integration. The subsequent computation of $A^*(A\Delta_i)$ is roughly twice as expensive as one model integration because differentiation doubles the number of quadratic terms. This makes one iteration step three times as expensive as a single model run. It follows that the explicit computation of A, which requires 30 model integrations, is as expensive as 10 iteration steps. So the computation of the matrices $E_{10}, E_{11}, \ldots, E_{30}$ is no longer an efficient strategy for the computation of the largest eigenvalues. If such a large number of eigenvalues is needed it is cheaper to compute A explicitly and to compute the full set of 30 eigenvalues.

The method using the adjoint may be applied to a model with many degrees of freedom. In the remaining sections it is explained how the growth rates can be used to compute the probability distribution for the error and why a limited number of iterations should also be sufficient in a realistic model.

2.4   Probability distribution of errors

In this section the above iterative procedure is used to estimate the largest eigenvalues and it is demonstrated how these eigenvalues can be used to establish an estimate of temporal predictability. The integration time is 4

days rather than 2 days as suggested in the introduction, because the error growth in the model is slower than the error growth in the atmosphere. In the model errors grow with a factor of 2.5 in 4 days.

A Monte Carlo method is used to determine the temporal probability distribution of the errors. An ensemble of 40000 random errors is integrated with the tangent linear equations. All errors initially have length one. All errors are added to the same starting point on the attractor. After four days their length R is given by:

$$R = (\mathrm{u}(t = 4), \mathrm{u}(t = 4))^{1/2} \tag{2.31}$$

The probability distribution of this radius is considered to be a perfect measure of the temporal predictability. It will be shown how this distribution can be approximated using the largest eigenvalues of A^*A.

The matrix $E_5^T E_5$ is computed in order to estimate the 5 dominant eigenvalues. In the following paragraphs the method to derive an error distribution function on the basis of a number of eigenvalues is given. Then this method is applied to the example with the 5 eigenvalues. Comparison with the Monte Carlo result indicates that 3 more iteration steps are needed.

First a reference distribution function f_{ref} is needed for the square root of the fraction of the kinetic energy of a unit-vector with random direction which projects on a coordinate axis. This distribution is identical for every possible coordinate axis. A Monte Carlo experiment is performed to obtain this distribution function. The function f_{ref} can be used for the projection of an initial error vector of unit length on vector v_i. The distribution function f_i for the square root of the error energy, which after an integration of four days, projects on the coordinate axis spanned by Av_i is:

$$f_i(r) = f_{ref}(\frac{r}{\lambda_i})\frac{1}{\lambda_i} \tag{2.32}$$

In other words, the reference distribution function f_{ref} is stretched by a factor λ_i. Consequently it must also be divided by λ_i, as the integrated probability must remain 1. Taking successive convolutions of the individual functions f_i with $i = 1, 2, \ldots, j$ the distribution function for the error becomes:

$$f_{1,\Sigma}(r) \;=\; f_1(r) \tag{2.33}$$

$$f_{i+1,\Sigma}(r) \;=\; \int_{\sigma=0}^{r} f_{i,\Sigma}((r^2 - \sigma^2)^{\frac{1}{2}}) f_{i+1}(\sigma) \left(\frac{r^2}{r^2 - \sigma^2}\right)^{\frac{1}{2}} d\sigma \tag{2.34}$$

$$f(R = r) \;\approx\; f_{30,\Sigma}(r) \approx f_{j,\Sigma}(r) \tag{2.35}$$

The functions $f_{i,\Sigma}$ give the distribution for the contribution of the first i eigenvalues to the error. So, Eq. (2.33) is satisfied by definition. The function $f_{i+1,\Sigma}$ is computed with the approximation that the distribution of the energy on axis Av_{i+1} is independent of the distribution of the energy on the axes $\{Av_1, Av_2, \ldots, Av_i\}$. The integral is over all possible states with energy r^2 on the first $i + 1$ axes. Differentiation of the argument of $f_{i,\Sigma}$ gives the geometry factor in Eq. (2.34):

$$\frac{d}{dr}(r^2 - \sigma^2)^{1/2} = \left(\frac{r^2}{r^2 - \sigma^2}\right)^{1/2} \tag{2.36}$$

The real error distribution is given by $f(R)$. The first approximation sign in Eq. (2.35) is due to the assumed independence of the functions f_i. This approximation is reasonable for the first few convolutions because a rather large number of 30 dimensions is used. At the last possible convolution to obtain $f_{30,\Sigma}$ from $f_{29,\Sigma}$ and f_{30} the approximation would break down completely. However, because λ_{30}^2 is negligibly small this convolution does not need to be performed. In fact it will appear that it is not necessary to go beyond the convolution of $f_{7,\Sigma}$ and f_8. As long as j is much smaller than 30 the most significant error is the one indicated with the second approximation sign, which is caused by taking only j distribution functions.

Figure 2.3a compares the distribution $f_{5,\Sigma}$ which is based on 5 eigenvalues with the Monte Carlo result. The tail of the distribution is estimated accurately. At lower lengths there is a systematic error. Here the error growth is significantly underestimated because 25 eigenvalues are assumed to be zero. One might correct for this statistically. An alternative is to do a few more iterations. Figure 2.3b compares the distribution $f_{8,\Sigma}$ which is based on 8 eigenvalues with the Monte Carlo result. It is concluded

that the temporal distribution function for the forecast error can be reproduced reasonably accurate on the basis of 8 eigenvalues. This does not mean that it can be predicted whether one particular forecast will have a large error or not. Indeed, from the Monte Carlo result it is observed that the error in the forecast depends heavily on the stochastic direction of the initial perturbation. It does mean that one can say how a large ensemble of equally likely initial perturbations will spread out in space.

To analyse the variability of the temporal predictability over the attractor a series of error distribution functions for 4-day integrations starting at 100 consecutive days is computed. Thus consecutive integrations have 3 days in common. The probabilities $P(R < 2)$, $P(R < 3)$ and $P(R < 4)$ are computed for every distribution. The first probability gives the probability of a high quality forecast. The second one gives the probability that a forecast is at least of moderate quality. $1 - P(R < 4)$ is the probability of a low quality forecast. The three probabilities are plotted in figure 2.4. One observes that the probability that a forecast is of low quality may vary between 10 and 25 %. It is concluded that there is a considerable variation in the predictability as a function of the position on the attractor.

2.5 Dimensionality of the error growth

In this section the dimensionality derived from the error growth analysis is compared with the dimensionality which can be derived from an EOF analysis. EOF analysis is a well known technique (see e.g. Preisendorfer 1988) that is often applied to meteorological data. It is first applied to a 30 year model data set to estimate the dimension of the model attractor. Then it is studied to what extent this information on the dimension of the model can be used to estimate the dimension of the error growth. The assumption is that the dimension of error growth will be of the order of the dimension of the model attractor. If this appears to be true the dimension of error growth in the real atmosphere can be estimated with the help of an EOF analysis of the atmospheric circulation. This then gives information on the number of eigenvalues of A^*A that have to be computed in a realistic atmospheric model with many degrees of freedom.

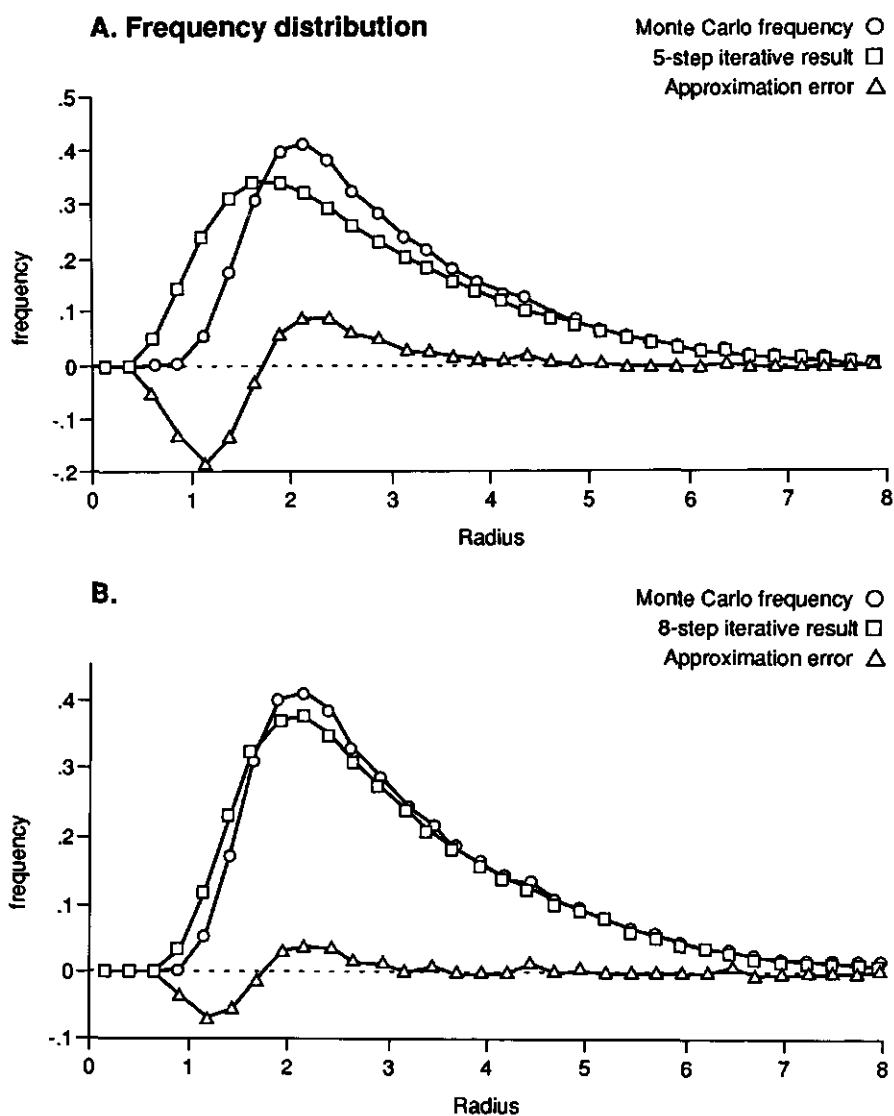Figure 2.3: (a) Comparison of the error size distribution obtained from 40000 Monte Carlo integrations and the profile based on E_5. The iteration result is marked with squares; the Monte Carlo curve with circles; the difference between the two curves with triangles. (b) Comparison of the error size distribution with the profile based on E_8 instead of E_5. Otherwise it is the same as (a)
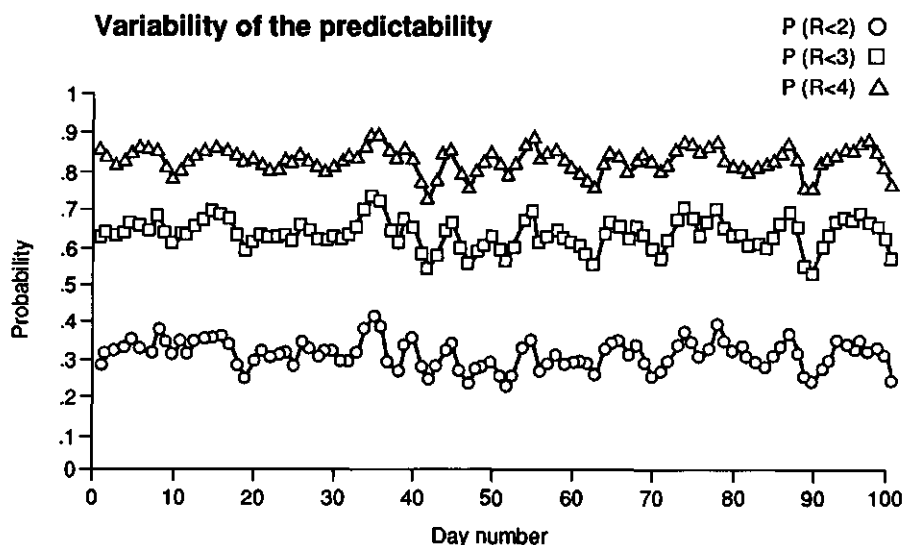
Figure 2.4: Variations in time of the probabilities $P(R < 2), P(R < 3)$ and $P(R < 4)$ during a 100 day integration. $P(R < 2)$ is indicated with circles; $P(R < 3)$ with squares; $P(R < 4)$ with triangles.

The argument for making the assumption on the dimension of error growth is that a model with a perturbed initial condition will rapidly converge to the model attractor in the first few hours of the integration. So after some time both the perturbed and the reference run can be described with the same EOF basis. As time progresses, the EOF basis will then also become an appropriate basis for the errors.

A 30 x 30 symmetric scatter matrix F is defined by:

$$F = \sum_{t=1}^{n} q(t)q^{T}(t) \tag{2.37}$$

Here the vector $q(t)$ is the state vector at time t. It is defined by Eqs. (2.15), (2.16), (2.17) an (2.18). The summation is over a 30-year run. Successive values of t are one day apart. The matrix F has a set of 30 orthonormal eigenvectors (i.e. orthonormal with respect to the kinetic energy inner product). These eigenvectors are the "Empirical Orthogonal Functions (EOFs)". The eigenvalues l_i of F are in descending order. They

Figure 2.5: Logarithmic eigenvalue graph for the instantaneous circulation and for the errors. The squares give the variance of the instantaneous circulation described with the EOFs. The crosses give the variance of the errors described with the basis-vectors for the errors.

are scaled such that their sum equals 100:

$$l_1 > l_2 > \ldots > l_{30} \geq 0$$
$$\sum_{i=1}^{30} l_i = 100$$

The individual l_i give the percentage of the variance described by their corresponding eigenvector. If the l_i decrease sharply towards 0, as a function of index number i, then only a few EOFs are needed to describe most of the variance of the system. The system is then low dimensional. The logarithms of the eigenvalues l_i are plotted against index number in figure 2.5. With the eigenvectors $1, 2, \ldots, 9$ 97 % of the variance can be explained. To explain the entire 100 % of the variance the maximum number of 30 EOFs is necessary.

Something similar can be done for the error variance with the eigenvalues of A^*A. The i-th eigenvalue λ_i^2 gives the average error energy projecting

on Av_i after four days of integration. For a total of 100 different initial positions on the attractor the eigenvalues $\lambda_1^2, \lambda_2^2, \ldots, \lambda_{30}^2$ are computed. After averaging and scaling a set of numbers m_i is obtained, which is comparable to the l_i for the EOF analysis:

$$m_i = 100 \frac{\overline{\lambda_i^2}}{\sum_{j=1}^{30} \overline{\lambda_j^2}}$$
$$m_1 > m_2 > \ldots > m_{30} \geq 0$$
$$\sum_{i=1}^{30} m_i = 100$$

The logarithms of the m_i are shown in figure 2.5. The first nine vectors again explain 97 % of the variance of the errors.

Comparing the eigenvalues of the scatter matrix with the eigenvalues of A^*A it is observed that the errors project a higher percentage of their energy on the first vector. This difference is compensated with lower percentages at the next three vectors. The remaining 26 eigenvalues have comparable percentages. It is concluded that the dimensionality of the error vectors is only slightly lower than the dimensionality of the instantaneous circulation patterns. So, the a priori assumption has now been validated for the 30-dimensional model.

To verify the argument for the assumption the same set of 100 initial positions on the attractor is used and from each point 100 integrations are done with a random initial error. According to the argument the EOF basis will become an appropriate basis for these errors as time progresses. In figure 2.6 the errors after four days of integration are projected on the EOF basis. As a reference the information on the logarithms of the l_i from the previous figure is included. It is clear that the EOF basis is more efficient for the description of the state vectors than for the error vectors. However, considering the random initial error direction, it is observed that the tangent linear equations are quite efficient in pushing a perturbed initial point towards the attractor. Combining this with the fact that the basis of A^*A is optimum for the errors, it is concluded that, in general, one may expect the dimensionality of the errors to be comparable to the dimensionality of instantaneous circulation patterns. The same reasoning will now be applied to a GCM.
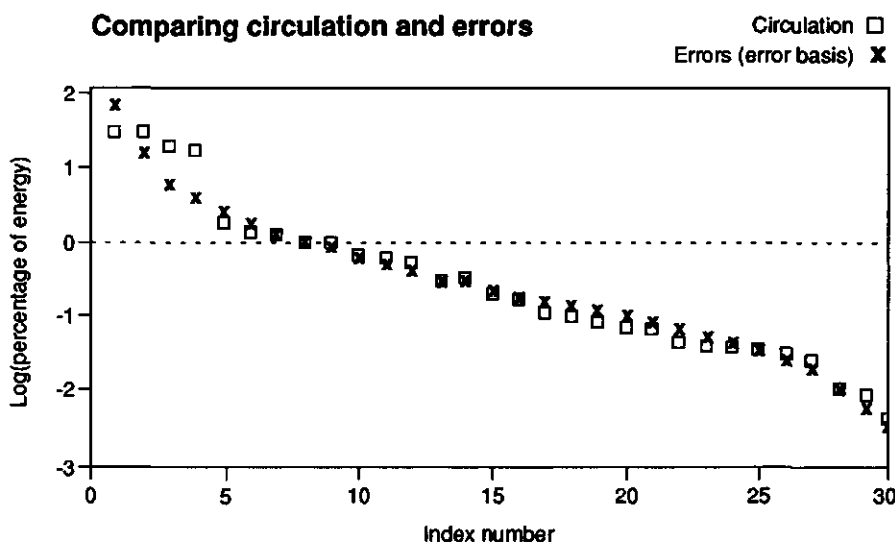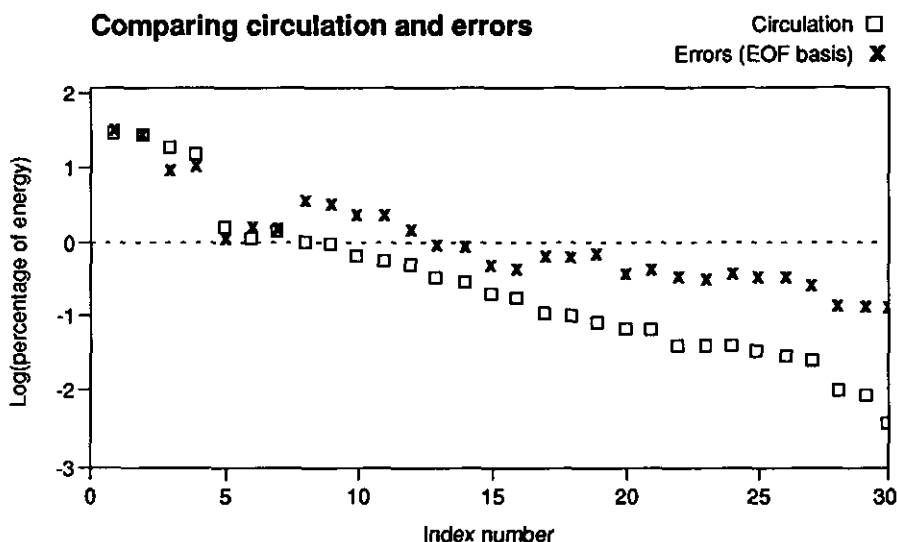
Figure 2.6: Logarithmic eigenvalue graph for the instantaneous circulation and for the errors. The squares give the variance of the instantaneous circulation described with the EOFs. The crosses give the variance of the errors described with the EOFs.

Rinne and Karhila (1979) and Rinne and Järvenoja (1979) give EOFs of the 500 mb instantaneous height field in the northern hemisphere based on a large data sample. They show that 25 EOFs suffice to describe 86 % of the variance. Doubling the number of EOFs to 50 increases the described variance to 94 %. It is assumed that a 2-day error projects as efficiently on the basis of EOFs as an instantaneous field. This means that a 50 dimensional EOF basis suffices to describe 94 % of the mean square error field. In the previous section it appeared that this assumption is optimistic. A projection matrix D is used to project the final error field on the EOF basis. The truncated mean square error R_D^2 becomes:

$$
\begin{aligned}
R_D^2 &= (\mathrm{D}\mathrm{A}u(t=0), \mathrm{D}\mathrm{A}u(t=0)) \\
&= (\mathrm{A}^*\mathrm{D}^*\mathrm{D}\mathrm{A}u(t=0), u(t=0))
\end{aligned}
\tag{2.38}
$$

This equation is almost identical to equation 2.20. The matrix $\mathrm{D}^*\mathrm{D}$ has 50 eigenvalues with value 1.0. Its other eigenvalues are zero. The matrix $\mathrm{A}^*\mathrm{D}^*\mathrm{D}\mathrm{A}$ has 50 non zero eigenvalues. All of them can be found with 51

iterations. So, for 94 % accuracy it is sufficient to do 51 iterations. Having noticed this the projection D can be removed. The iterative procedure is applied. This procedure determines its own optimum basis for the errors. Finally the eigenvalues of the complete system are estimated. Obviously, using a more relevant basis, less than 51 iterations are needed to arrive at the same percentage of described variance. This may compensate for the optimistic assumption at the beginning as it did in the 30-dimensional model. The total computional cost of 51 iterations is roughly the same as the cost of 153 model runs.

The above numbers can be improved if one is interested in geographically localized rather than global predictability. Peagle and Haslam (1982) give an EOF analysis of seven years of winter data over a portion of the Northern Hemisphere centered over the western United States. They find that only 18 EOFs are needed to describe 94 % of the variance. This reduces the dimension and thus also the number of iterations by a factor of 3.

The analyses started with a homogeneous random initial error field. This simplified the exposition of the problem but is not realistic. The errors in the observations can be described by means of a vector w such that:

$$u(0) = Mw \tag{2.39}$$

$$\overline{(w_i, w_j)} = \frac{\delta_{ij}}{N_{obs}} \tag{2.40}$$

$$(w, w) = 1 \tag{2.41}$$

Eq. (2.39) relates the vector w of observation errors with an initial error field $u(0)$. If the measurement errors are small M is a linear operator. In general because of the complexity of assimilation methods the operator M may be cumbersome to evaluate. Eq. (2.40) states that all N_{obs} observations are independent. Eq. (2.41) assigns unit length to the error vector. The adapted equation for the error energy at time t now becomes:

$$R_{DM}^2 = (DAMw, DAMw) = (M^*A^*D^*DAMw, w) \tag{2.42}$$

Because the dimension of the problem is already constrained by the matrix D^*D the basis for the initial error can be made as large as the basis

of u(0) without causing significant computational problems. In practice, though, the number of observations may be much smaller than the number of model variables. Thus, the model matrix A, the projection matrix D, the assimilation matrix M and the maximization procedure in general all have the effect of rapidly eliminating irrelevant initialization directions.

2.6   Discussion and conclusions

In this chapter variations in the deterministic error growth have been studied in a low order spectral model. Only short range error growth has been studied. This allowed the use the tangent linear equations for the evolution of the errors. With these equations and their adjoint the directions with large error growth have been computed. From the growth rates in these directions a probability distribution for the errors could be established. Parameters for the skill of a forecast were obtained from this distribution.

It would be of value to know whether the above method can be applied to a large model. The hypothesis was made that the dimension of the error growth is of the order of the dimension of the instantaneous circulation patterns. The hypothesis has been tested in the low order model, where indeed the number of relevant error growth directions happened to be almost identical to the number of relevant EOFs. This is a plausible result because a run with small initial error will converge to the model attractor. From an analysis of the real atmospheric circulation a number of 50 relevant EOFs can be obtained. According to the hypothesis this is equivalent to 50 relevant error growth directions. Computing these directions is as expensive as about 150 model runs. With these directions and the corresponding growth rates one can get parameters for the skill.

The above methods may still be judged prohibitively expensive. A reduction in the cost may be obtained by choosing a regional instead of a global error norm and by relaxed requirements on the numerical precision of the skill forecast. All of this depends on the yet to be determined spectrum of eigenvalues which might be different from its hypothesized shape. To get a useful estimate of the skill one should use an inhomoge-

neous initial error field which is consistent with errors in the observations. Formally this can be done with the operator M, which transforms a vector with measurement errors into an initial error field. Non-linear error growth is not discussed in this chapter. It limits the applicability of the above method to short-range predictions.

Acknowledgements

REFERENCES

Bourke, W.B., Mc Avaney, B., Puri, K. and R. Thurling, 1977: Global modeling of atmospheric flow by spectral methods. *Methods in Computational Physics*, *17*, 267-334.

Courtier, P. and O. Talagrand, 1987: Variational assimilation of meteorological observations with the adjoint vorticity equation II: Numerical results. *Quart. J. Roy. Meteor. Soc.*, *113*, 1329-1347.

Epstein, E.S., 1969: Stochastic dynamic prediction. *Tellus*, *21*, 739-759.

Holton, J.R., 1979: *An Introduction to Dynamic Meteorology*, [ed. W.L. Donn], New York, Academic Press, 391 pp.

Kaplan, J.L. and J.A. Yorke, 1979: Chaotic behavior of multidimensional difference equations, in *Functional Differential Equations and Approximation of Fixed Points*, [eds. Peitgen, H.O. and Walther H.O.] Heidelberg-New York, Springer, 228-237.

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

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

Leith, C.E., 1974: Theoretical Skill of Monte Carlo Forecasts. *Mon. Wea. Rev.*, *102*, 409-418.

Lorenz, E.N., 1960: Energy and numerical weather prediction. *Tellus*, *12*, 364-373.

Lorenz, E.N., 1963: Deterministic Nonperiodic Flow. *J.Atmos.Sci.*, *20*, 130.

Lorenz, E.N., 1985: The growth of errors in prediction. In *Turbulence and Predictability in Geophysical Fluid Dynamics and Climate Dynamics*, [eds. Ghil, M., Benzi, R. and Parisi, G.] Amsterdam, North-Holland, 243-265.

Orszag, S.A., 1970: Transform method for the calculation of vector-coupled sums: Application to the spectral form of the vorticity equation. *J.Atmos.Sci.*, *27*, 890-895.

Palmer, T.N. and S. Tibaldi, 1988: On the prediction of forecast skill. *Mon. Wea. Rev.*, *116*, 2453-2480.

Paegle, J.N. and R.B. Haslam, 1982: Empirical orthogonal function estimates of local predictability. *J. Appl. Meteor.*, *20*, 117-226.

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

Platzman, G.W., 1960: The spectral form of the vorticity equation. *Journal of Meteorology*, *17*, 635-644.

Preisendorfer, R.W., 1988: *Principal component analysis in meteorology and oceanography*, [ed. C.D. Mobley], Amsterdam, Elsevier, 425 pp.

Reinhold, B.B. and R.T. Pierrehumbert, 1982: Dynamics of weather regimes: Quasi stationary waves and blocking. *Mon. Wea. Rev.*, *110*, 1106-1145.

Rinne, J. and S. Järvenoja, 1979: Truncation of the EOF series representing 500 mb heights. *Quart. J. Roy. Meteor. Soc.*, *105*, 885-897.

Rinne, J. and V. Karhila, 1979: Empirical orthogonal functions of 500 mb height in the northern hemisphere determined from a large data sample. *Quart. J. Roy. Meteor. Soc.*, *105*, 873-884.

Silberman, I., 1954: Planetary waves in the atmosphere. *Journal of Meteorology*, *11*, 27-34.

Schuster, H.G., 1988: *Deterministic Chaos An Introduction*, [ed. W. Greulich], Weinheim, VCH, 270 pp.

Talagrand, O. and P. Courtier, 1987: Variational assimilation of meteorological observations with the adjoint vorticity equation I: theory. *Quart. J. Roy. Meteor. Soc.*, *113*, 1311-1328.

Thiébaux, H.J. and L.L. Morone, 1990: Short-term systematic errors in global forecasts: their estimation and removal. *Tellus*, *42A*, 209-229.

Wolf, A., Swift, J.B., Swinney, H.L. and J.A. Vastano, 1985: Determining Lyapunov exponents from a time series. *Physica D*, *16*, 285-317.

Chapter 3

The quality of skill forecasts

Abstract

A skill forecast gives the probability distribution for the error in the forecast. The purpose of this chapter is to develop a skill-forecasting method. The method is applied to an atmospheric model. It is a spectral two-layer quasi-geostrophic model with a triangular truncation at wavenumber 5. The analysis is restricted to internal error growth. It is investigated how observational errors lead to errors in the analysis. It appears that climatological distributions can be used for the errors in the analysis. In the forecast run the evolution of these distributions is computed. For that purpose the tangent linear equations for the errors are used. Because of this linearization the results are valid for short-range skill forecasts only. The Lanczos algorithm is used to find the structures that dominate the forecast error. This algorithm is intended to be applicable in a realistic model.

3.1   Introduction

Some weather forecasts are more successful than others. The quality of forecasts depends on the accuracy of the initial state and on the sensitivity of the atmosphere to small differences in the initial state. With a skill forecast one predicts the quality of a forecast. Because the forecast error is stochastic a skill forecast must be given in terms of a probability distribution. It allows a forecaster to make a selective use of the predictions

issued by an operational center. In order to compute this distribution it
is studied how stochastic errors in the observations cause stochastic errors
in the forecast. The first problem is how errors in the observations enter
the analysis. The second problem is how these errors propagate during
the forecast run. In this chapter errors due to imperfections of the model
are neglected.

A forecast starts with a data-assimilation process. Data assimilation
updates an old estimate of the state of the atmosphere with new obser-
vations. The resulting estimate is called the analysis. The next data
assimilation uses this analysis, which is then called the previous analysis,
as an old estimate of the state of the atmosphere. In meteorology the
analysis is usually obtained via Optimal Interpolation (OI)(Gandin 1963;
Bergman 1979). Optimal interpolation uses approximate rules for the evo-
lution of the statistics of the error. The resulting analysis is not optimal
because a valid evolution of the statistics of the error would give a more
likely analysis.

An optimal estimate may be obtained with a Kalman filter (Kalman
1960; Ghil et al. 1981; Cohn and Parrish 1991). The computational cost
of the Kalman filter has prohibited its operational use. An alternative
analysis method uses the adjoint of the tangent linear model (Marchuk
1974) to minimize a cost function. An extensive discussion on data assim-
ilation with an adjoint model is given by Courtier and Talagrand (1987)
and by Talagrand and Courtier (1987). Thacker (1989) gives a description
of the adjoint method which incorporates the error statistics of the previ-
ous analysis. Thacker (1989) suggests the name "least squares method"
instead of "adjoint method" because this emphasizes the connection to re-
gressian analysis (see e.g. Draper and Smith 1966). The adjoint method
has a computational cost comparable to the cost of the Kalman-filter if
the evolution of the error statistics is computed. If the errors in the ob-
servations are sufficiently small, and if the model is perfect, then both
methods are optimal. A comparison of a Kalman filter and the adjoint
method is given by Lorenc (1986) and more recently in the review paper
by Ghil and Malanotte-Rizzoli (1991). In the present chapter the adjoint
equation method is used because it is conceptually easier to understand.

A skill forecast may be obtained from the subsequent evolution of the error. Because a rigorous computation is very expensive approximating methods are proposed in this chapter. A brute force method is still needed to test the validity of the approximations. Computational reasons make the use of a simple atmospheric model necessary, but at the same time it is assumed that this model is perfect. This is done because the simplicity of the model precludes a useful study of model error. To study model errors one must compare operational forecasts with observations. This will remain out of the scope of this chapter (see e.g. Dee et al. 1985). Only the sensitivity of the forecast to errors in the observations is studied. For a short range forecast, which has only small errors, "model errors" and "predictability errors" can be treated independently. It is only after some time that the different error terms start to interact non-linearly. Also it is both instructive and convenient to study "predictability errors" in a simple environment before applying the ideas in operational models.

Thus one has to assume that the error growth is linear. This assumption breaks down after about 48 hours (Lacarra and Talagrand 1988). Consequently one is limited to short range skill forecasts. This limitation makes it possible to compute the covariance matrix of the error. To obtain the full matrix the tangent linear equations are integrated once for all coordinates in phase space. The two-layer quasi-geostrophic hemispheric model discussed in Houtekamer (1991)(hereafter referred to as PH) is used. This model, which is truncated at T5, has a 30-dimensional phase space. The tangent linear model, which describes the evolution of small errors, and the adjoint of the tangent linear model are given in PH.

This chapter starts with a description of the observational network and the data-assimilation procedure. The data assimilation is performed by minimizing a quadratic function of the differences between the data and their model counterparts. The Hessian matrix, which contains the second derivatives of this quadratic function, is computed. The inverse of the Hessian is the covariance matrix that establishes to which accuracy the model state is determined by the data (Thacker 1989). A series of brute force computations of the forecast error distributions is performed. The resulting rigorous skill forecasts serve as a standard against which

approximating methods are compared.

As a first test these rigorous skill forecasts that start from the covariance matrix for the error in the previous analysis are compared with skill forecasts that start from a climatological covariance matrix. This first test quantifies the degradation of the skill forecast which must occur if a climatological error-distribution is used for the previous analysis.

From an economical point of view one would like to move the climatology assumption from the previous analysis to the present analysis. Thus information on error growth, that might have been gathered during the data assimilation, is ignored. The second test quantifies the consequences of ignoring this information.

The third simplifying step is a limitation to a small number of error growth directions. As suggested by Lacarra and Talagrand (1988) a Lanczos algorithm (see e.g. Parlett 1980) is used to find these few important error growth directions. Using the climatological covariance matrix for the error in the analysis and using the Lanczos algorithm a method is obtained that can be applied to a realistic model. Again it is tested against the rigorous skill forecast.

In a final test the error in the analysis has a simple equipartition energy spectrum. The results show that the inhomogeneity of the errors in the analysis cannot be ignored. The chapter is concluded with a discussion on the relevance of this work for error growth in more realistic models.

3.2   The data assimilation

The objective of this chapter is to analyse the influence of an inhomogeneous distribution for the analysis error on the accuracy of the resulting forecast. Such inhomogeneous error distributions are simulated with a realistic data-assimilation process. The data assimilation consists of a least squares fit of a model to data. A general description of the least squares method is given by Draper and Smith (1966). An application to an oceanographic model is given by Thacker (1989).

This section first describes the observational network. It proceeds with a discussion on the least squares method. This method gives the analysis
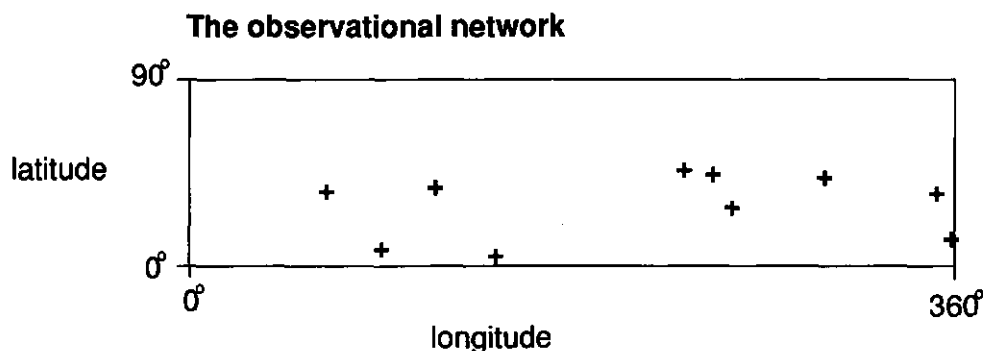
The observational network

Figure 3.1: The observational network. The horizontal coordinate is the geographical longitude of the observing station. The vertical coordinate is for the geographical latitude. The scales are in degrees. The observing stations are indicated with plus signs.

and the covariance matrix of its error using as input the previous analysis, the observations and the covariance matrices for the error in the previous analysis and for the observational error.

The observational network has a total of 10 observing stations. Their locations are shown in fig. 3.1. The locations are chosen with a random number generator. Due to the limited number of observing stations the inhomogeneity of the network comes as a natural property. At the 10 stations independent measurements of the barotropic and the baroclinic values of the streamfunction are simulated at times 12 UT and 24 UT. The errors in the observations have independent normal distributions with variance σ^2. In reality observations may contain gross errors. This causes more excessively large errors than expected from a normal distribution. A quality control is needed to identify and remove such gross errors (Lorenc 1984). The distribution after quality control may be considered to be normal. Note that in the T5 model 30 numbers are enough to define the analysis. The previous analysis and the set of observations consists of a total of 70 numbers so the assimilation problem is overdetermined. It needs to be overdetermined to guarantee a unique and reliable solution.

Before proceeding with a description of the assimilation method some notations for relevant quantities are given. The true state vector for the model atmosphere is $q_m(t)$, with $t = -24$ hours at the beginning of the

assimilation, $t = 0$ at the analysis time and $t = t_f$ at the time of the forecast. The observations at -12 and 0 are used for the assimilation. Unless stated otherwise t_f equals 52 hours. The previous analysis is $\hat{q}(-24)$. The best fit to the data is $q(-24)$. Its forward integration is $q(t)$. The corresponding error vector is $u(t)$:

$$q(t) = q_m(t) + u(t). \tag{3.1}$$

In the following the differences between the previous analysis, the best fit and the true atmospheric state are assumed to be small. Also the measurement errors are taken to be small. These assumptions allow the linearization of operators about the previous analysis rather then about the unknown true atmospheric state. The errors in these approximations appear as higher order terms which are systematically neglected. Because all operators are linear and all errors in the observations have normal distributions the computed error distributions will also be normal (e.g. Example 11.2 in Kendall et al. 1986).

The least squares fit of a model to data is determined by the minimum of a cost function $J(x)$ that is quadratic in the differences between the data d and their modelcounterparts $m(x)$, where x denotes the starting value of the model at $t = -24$:

$$J(x) = \frac{1}{2}(d - m(x))^T S(d - m(x)). \tag{3.2}$$

The data, which consists of both the previous analysis and the new observations, are represented as elements of a column vector d. The matrix S reflects the relative precision of the data. It is the inverse of the covariance matrix for the error in the data:

$$S^{-1} = \overline{(d - m(q_m(-24)))(d - m(q_m(-24)))^T}, \tag{3.3}$$

where the overbar denotes expectation value. This matrix S can be written as:

$$S = \begin{pmatrix} \hat{Q}^{-1} & 0 \\ 0 & \sigma^{-2} I \end{pmatrix}, \tag{3.4}$$

where $\hat{Q}$ is the covariance matrix for the error in the previous analysis, σ^2 is the variance of the measurement errors and I is the identity matrix.

The minimization of Eq. (3.2) is performed using the Hessian matrix H of second derivatives of the cost function:

$$H = (\frac{\partial m}{\partial x})^T S(\frac{\partial m}{\partial x}), \tag{3.5}$$

where $\partial m/\partial x$ represents the Jacobian matrix. To compute the Jacobian, one row at a time, the tangent linear equations are integrated once for unit perturbations in all model coordinates. Finally the best fit is computed as a correction to the previous analysis (Thacker 1989):

$$q(-24) = \hat{q}(-24) + H^{-1}(\frac{\partial m}{\partial x})^T S(d - m(\hat{q}(-24))), \tag{3.6}$$

where one notices that only the information in the new observations is used to update the previous analysis.

After the assimilation of the state vector the error statistics of the best fit can be computed. It will appear that most of the necessary work is done already. The accuracy of the best fit can be described with a covariance matrix $Q(-24)$:

$$Q(-24) \equiv \overline{u(-24)u^T(-24)}. \tag{3.7}$$

The error in the best fit $u(-24)$ is:

$$u(-24) = H^{-1}(\frac{\partial m}{\partial x})^T S(d - m(q_m(-24))). \tag{3.8}$$

This may not appear convenient because the true atmospheric state q_m is not known. However for the present purpose it is sufficient to know the statistics of the error. These statistics are given by Eq. (3.3). Using Eq (3.3), (3.5), (3.7) and (3.8) one obtains for the covariance matrix:

$$Q(-24) = H^{-1}(\frac{\partial m}{\partial x})^T S S^{-1} S^T \frac{\partial m}{\partial x} H^{-1T} = H^{-1}. \tag{3.9}$$

It remains to integrate the covariance matrix $Q(-24)$ to the time of the analysis:

$$Q(0) = R(0, -24)Q(-24)R^T(0, -24), \tag{3.10}$$

where $R(t_1, t_2)$ is the resolvent operator between times t_1 and t_2. It represents the integration of a small perturbation with the tangent linear model from time t_2 until time t_1. The matrix $R(0, -24)$ is a by-product of the computation of the Jacobian $\partial m / \partial x$. The symbol R is used for the mapping of an analysis error to a forecast error:

$$u(t_f) = Ru(0). \tag{3.11}$$

The covariance matrix for the forecast error becomes:

$$Q(t_f) = RQ(0)R^T. \tag{3.12}$$

For the next data assimilation $q(0)$ can be used as the previous analysis and $Q(0)$ as its covariance matrix. Thus the statistics for the analysis error and for the forecast error can be derived from statistics of the error in the previous analysis and of the measurements error. In the next section some modifications are given to make the assimilation feasible with realistic models.

3.3   A feasible assimilation method

The observational network, the data-assimilation method and the method to compute the relevant covariance matrices have been discussed. This gives in principle the necessary tools for a skill forecast. However, because it is intended to analyse the feasibility of actual skill forecasts, two modifications to the above scheme are made. The modifications are necessary because in practice the model is not perfect and the restricted computer power does not allow the computation of the full covariance matrix.

3.3.1   Correction for model error

As discussed in PH the dynamics of the model is chaotic. This has consequences for the evolution of the covariance matrix. An initially spherical distribution of errors will expand in unstable directions and shrink in stable directions. The growing error fields are eventually checked by new observations. In the stable directions the accuracy becomes infinitely high. New unbiased observations can only increase the accuracy in these

directions even further. This is realistic in a perfect model but not in a model that is only an approximation to reality. For this reason errors are not allowed to be much smaller than the errors in the verifying observations. This intuitive notion is incorporated by changing the covariance matrix $\hat{Q}$ of the error in the previous analysis. Any eigenvalue of $\hat{Q}$ less than a certain ad hoc value δ, which reflects in some way the accuracy of the model, is replaced by the value δ.

3.3.2 A climatological covariance matrix

The assimilation procedure requires the covariances $\hat{Q}$ for the error in the previous analysis. To compute these covariances the model is integrated once for all model coordinates. In practice this will never be feasible for a realistic model. So, instead, climatological covariances are used to assimilate the state vector. This leads to a data assimilation technique which is not optimal but which is applicable to a realistic model. Consequently, for the present experiments, a non-perfect but realistic method is used to assimilate the state vector.

The climatological covariance matrix is obtained in the following way. A model run is started somewhere on the attractor. The covariance matrix $Q(0)$ for the first analysis is computed using only the statistics of the observations. For subsequent assimilations a previous analysis with a covariance matrix for its errors can be used. Eigenvalues of $Q(0)$ less than δ are replaced by eigenvalues δ and the adapted matrix $Q_\delta(0)$ is used as covariance matrix $\hat{Q}$ for the error in the previous analysis in the next data assimilation step. This process is continued for 20 days, after which the spin-up of the covariance matrix $\hat{Q}$ has ended. The next 40 assimilations are used to compute successively better approximations $\overline{Q}_i$ of the climatological covariance matrix $\overline{Q}$.

$$\overline{Q}_i = \frac{1}{i} \sum_{day=21}^{20+i} \hat{Q}$$

$$= \frac{1}{i} \sum_{day=20}^{19+i} Q_\delta(0), \qquad i = 1, \ldots, 40 \tag{3.13}$$

$$\overline{Q} = \overline{Q}_{40}. \tag{3.14}$$

During the 40 day experiment the actual estimate $\overline{Q}_i$ of the climatological covariance matrix is used in the data assimilation. This makes the data assimilation non-perfect. A perfect data assimilation uses actual covariances $\hat{Q}$. It will be shown how the actual covariances $\hat{Q}$ for the error in the previous analysis and the non-perfect data-assimilation lead to covariances $Q_\delta(0)$ in the analysis. This derivation is needed for the determination of the climatological covariance matrix with Eq. (3.13) and for the rigorous verfications of skill forecasts performed with more simple methods.

The evolution of covariances $\hat{Q}$ needs to be followed under a data assimilation which uses climatological covariances $\overline{Q}$. Thus the equivalents of Eqs. (3.4), (3.5) and (3.8) become:

$$S_F = \begin{pmatrix} \overline{Q}^{-1} & 0 \\ 0 & \sigma^{-2}I \end{pmatrix}, \tag{3.15}$$

$$H_F = (\frac{\partial m}{\partial x})^T S_F (\frac{\partial m}{\partial x}), \tag{3.16}$$

$$u(-24) = H_F^{-1}(\frac{\partial m}{\partial x})^T S_F(d - m(q_m(-24))), \tag{3.17}$$

where the subscript F indicates the use of a climatological covariance matrix by the data assimilation. To obtain the actual covariances Eq. (3.3), (3.7) and (3.17) are combined:

$$Q(-24) = H_F^{-1}(\frac{\partial m}{\partial x})^T S_F S^{-1} S_F^T \frac{\partial m}{\partial x} H_F^{-1T}. \tag{3.18}$$

These covariances are integrated to the time of the analysis using Eq. (3.10). The matrix $Q(0)$ is then modified to give $Q_\delta(0)$ which completes the algorithm to update the covariance matrix in the case of a non-perfect data-assimilation. In the following experiments the climatological covariances $\overline{Q}$ are used to assimilate the state vector. If the minimization of the cost function is done using adjoint equations the assimilation is also feasible for more realistic models. In this case one does not obtain actual covariances for the analysis error, but it will be shown that a skill forecast can be based on climatological covariances for the analysis error.

3.4 The experiments

In this section actual distributions for the forecast error are compared with some approximate distributions. The actual distribution of the forecast error is obtained from the actual covariances for the error in the forecast. The covariances will be computed rigorously by starting from actual covariances of the error in the previous analysis. The non-perfect data assimilation, which is based on climatological covariances for the error in the previous analysis, is used to obtain covariances for the analysis error. Next the tangent linear model is used for the evolution of the covariances to the forecast time. Then finally the distribution for the kinetic energy of the forecast error can be computed. This provides the reference distribution against which more feasible skill-forecasting methods are tested.

As a first approximation climatological covariances are used at the time of the previous analysis. So the climatological covariances of the error in the previous analysis are mapped with the non-perfect data assimilation and the tangent linear model to the covariances of the analysis error and of the forecast error. This makes the data assimilation and the skill forecast a consistent pair based on the same information.

In a second approximation the skill forecast is started from climatological covariances for the analysis error. These covariances are identical to the climatological covariances of the error in the previous analysis. In doing so the error information which might have been collected during the data assimilation is neglected.

As a third approximation a Lanczos algorithm is used to find the few dominating error growth structures. This approximation makes the skill forecast computationally feasible when using a realistic forecast model.

A final approximation uses a homogeneous distribution for the analysis errors. In this way information on the distribution of initial errors is ignored. The approximations are tested against the rigorous skill forecast.

3.4.1 The reference distribution

During a 1600-day run the covariance matrix $Q(t_f)$ is computed for each successive day. It is shown how the probability distributions for the kinetic

energy K of the forecast error can be computed from the covariance matrix $Q(t_f)$. For the error energy one can write:

$$K = u^T(t_f)u(t_f) = \langle u(t_f), u(t_f) \rangle, \tag{3.19}$$

where the brackets $\langle , \rangle$ denote an inner product proportional to kinetic energy. Equation (3.19) requires a proper scaling of the coordinates for u. The probability distribution for the error energy K can be obtained from an eigenvector analysis of $Q(t_f)$. The vectors $\{q_i, i = 1, \cdots, 30\}$ are the eigenvectors of $Q(t_f)$ at eigenvalues $\{\lambda_i^2, i = 1, \cdots, 30\}$. The forecast error $u(t_f)$ can be written in a basis spanned by the eigenvectors:

$$u(t_f) = \sum_{i=1}^{30} d_i \lambda_i q_i, \tag{3.20}$$

where the projection coefficients $\{d_i, i = 1, \cdots, 30\}$ have independent normal distributions with zero mean and unit variance. For the energy K of the forecast error one has:

$$K = \langle \sum_{i=1}^{30} d_i \lambda_i q_i, \sum_{i=1}^{30} d_i \lambda_i q_i \rangle = \sum_{i=1}^{30} d_i^2 \lambda_i^2. \tag{3.21}$$

table 3.1: quality of the approximations t_f=52								
	correlation with K			average values					
	16%	50%	84%	16%	σ	50%	σ	84%	σ
R^2	1.00	1.00	1.00	5.7	0.5	11.6	1.5	24.6	4.4
R_F^2	0.77	0.78	0.80	6.0	0.4	11.7	1.2	23.7	3.3
R_A^2	0.63	0.60	0.60	6.7	0.4	12.4	1.1	23.4	2.6
R_L^2	0.58	0.60	0.60	4.4	0.4	9.9	1.1	20.9	2.7
R_I^2	0.56	0.48	0.42	12.4	0.5	22.4	1.1	42.8	2.8

    With a random number generator 4000 random sets of projection coefficients are generated. The resulting 4000 values of K give a Monte Carlo estimate of the probability distribution. The values of K below which one has 16 %, 50 % and 84 % of all realizations are computed. The value below which one has 50 % of the points is the median. Fig. 3.2 displays the 16 % point, the median and the 84 % point during the first 100 days

Figure 3.2: The actual error distribution. The horizontal coordinate gives the day from which the forecast was made. The vertical coordinate gives the forecast error energy below which the forecast is expected with 16 % probability (dotted line), 50 % probability (dashed line) and 84 % probability (solid line).

of the 1600-day run. It is noted that the predictability is quite variable. Observe for example that the prediction starting at day 38 has an error variance which is almost twice as large as the error variance for the prediction starting at day 40. Fig. 3.2 can be considered to display the true probability distribution of the errors. The curve connecting the 84 % points, provides the reference against which approximations can be tested. To diagnose systematic errors table 3.1 displays the average energy of the error at the 16 % point, the median and the 84 % point. As a measure for the variability of the predictability the root mean square of the day-to-day variations in these values is given. Random errors are identified by the computed correlations.

3.4.2 Climatology at the previous analysis

As a first approximation it is assumed that the covariances of the error in the previous analysis take their climatological values. The estimated

covariances at the time of the analysis and at the time of the forecast are respectively:

$$Q_F(0) = R(0, -24)H_F^{-1}R^T(0, -24) \tag{3.22}$$

$$Q_F(t_f) = RQ_F(0)R^T. \tag{3.23}$$

The above method is used to derive a skill forecast based on $Q_F(t_f)$. The location of the 84 % point is shown in fig. 3.3b. Although the position of most of the peaks is the same, substantial differences do exist between the figures 3.3a and 3.3b. For instance, at day 32, a favourable probability distribution is predicted. This conflicts with the model's reality. A priori the variability of the predictability is expected to decrease, because taking a climatological covariance matrix is bound to smooth the errors. Indeed, in table 3.1, the root mean square of the energy for the 84 % point has decreased from 4.4 to 3.3. The correlation coefficients of the 16 % point, the median and the 84 % point between K and K_F are 0.77, 0.78 and 0.80. These correlations are computed from the full 1600-day run. The correlation coefficients are collected in table 3.1. Table 3.1 contains the same information on the next three approximations. To get a feeling for the accuracy of the estimated correlations it may be assumed that the percentage points and the median have normal distributions. If two normal distributions have correlation ρ, an estimate $\hat{\rho}$ of ρ based on n pairs has variance (Kendall et al. 1986):

$$\overline{(\rho - \hat{\rho})^2} = \frac{1}{n}(1 - \rho^2)^2 \approx \frac{1}{n}(1 - \hat{\rho}^2)^2. \tag{3.24}$$

Substitution of $n = 1600$ and the values of table 3.1 into Eq. (3.24) shows that the root mean-square error in the correlation is 0.02 in the worst case ($\hat{\rho} = 0.42$) and 0.009 in the best case ($\hat{\rho} = 0.80$). For the present purposes the estimated correlation coefficients are sufficiently reliable.

If a climatological approximation $\overline{Q}$ is taken for the covariance matrix $\hat{Q}$ of the error in the previous analysis the skill forecast is far from being perfect.
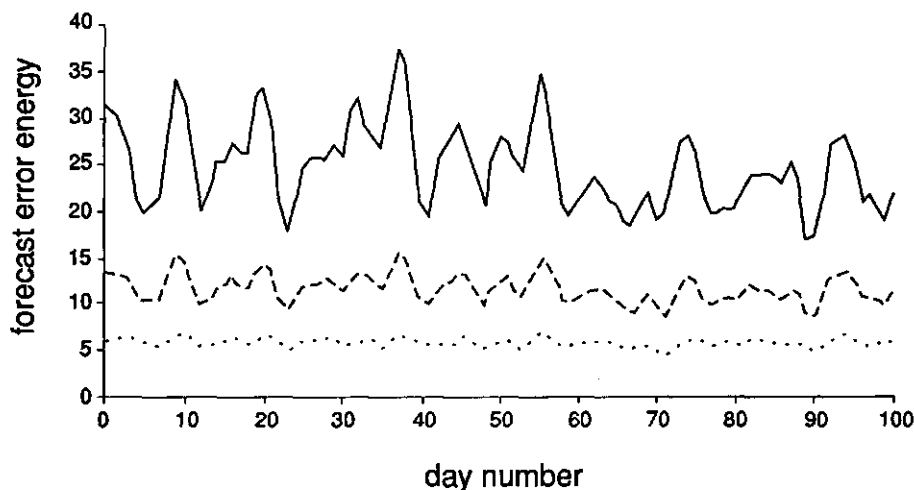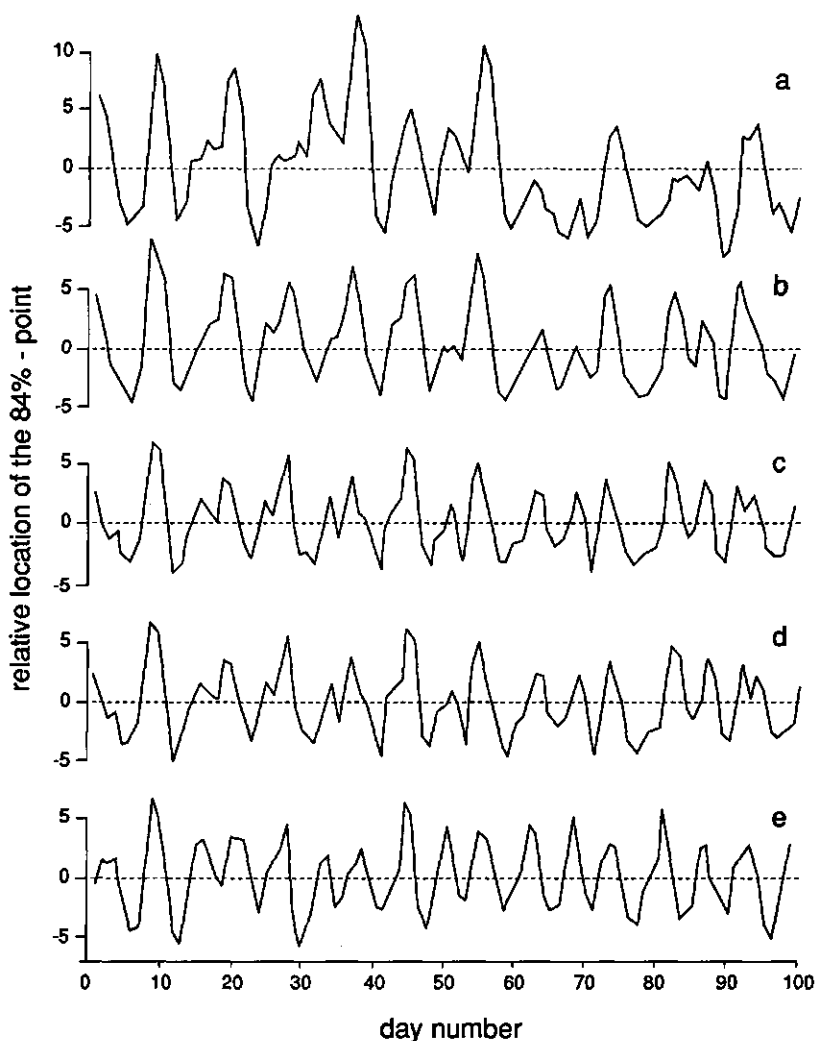
Figure 3.3: Actual and predicted variation of the 84 % point. The horizontal coordinate gives the day from which the forecast was made. The vertical coordinate gives the forecast error energy below which the forecast is expected with 84 % probability. Different predictions are separated with arbitrary vertical displacements. Shown are the 84 % points for: (a) reality, (b) a skill forecast starting from a climatological covariance for the previous analysis, (c) a skill forecast starting from a climatological covariance for the analysis, (d) a skill forecast as (c), but using a Lanczos algorithm. (e) a skill forecast starting from an equipartition distribution for the error in the analysis.

3.4.3   Analysis climatology

In a second approximation the climatological covariance matrix $\overline{Q}$ is used for the error in the analysis. An average assimilation operator B is computed from the eigenvectors of $\overline{Q}$:

$$\overline{Q} = BB^T. \tag{3.25}$$

Thus the columns of B contain the eigenvectors times the root of the eigenvalues. Because B is a climatological operator it has to be determined only once. The subscript A denotes quantities based on this assumption. The new estimate for the covariance matrix of the forecast error becomes:

$$Q_A(t_f) = RBB^TR^T = RBB^*R^*, \tag{3.26}$$

where the asterisk denotes adjoint. With euclidian innerproducts, which are used in this chapter, adjoint operators are identical to transpose operators. The application of R^* to a vector requires either the backward integration of the homogeneous adjoint of the tangent linear model (PH) or the computation of the full matrix R and its transpose. The spectrum of eigenvalues has the property:

$$\sigma(Q_A(t_f)) = \sigma(B^*R^*RB). \tag{3.27}$$

This general property is used in the next section by the Lanczos algorithm.

For a skill forecast based on $Q_A(t_f)$ the position of the 84 % point varies as shown in fig. 3.3c. A comparison with fig. 3.3b suggests that the variability is less chaotic and more periodic. As would be expected the variability of the predictability has decreased even further. The correlation coefficients of the 16 % point, the median and the 84 % point for K and K_A have dropped to 0.63, 0.60 and 0.60.

3.4.4   Reduction of error space

The third approximation, and the last step towards a "feasible" skill forecast, is the use of a Lanczos algorithm. It is feasible because it no longer requires perturbations in all model coordinates. The Lanczos algorithm gives estimates for the j largest eigenvalues of the matrix B^*R^*RB. The

algorithm is analogous to the one used in PH. The operator RB takes the role of his operator A. For completeness the new algorithm is given.

In a series of j iterative steps a basis $\{\Delta_1, \Delta_2, \cdots, \Delta_j\}$ is constructed which is an approximation to the basis for the j most important eigenvectors of B*R*RB. The first vector Δ_1, which is of length one, is chosen at random. New directions are obtained with the operator B*R*RB. They are made orthonormal to the previous directions with a Gram-Schmidt procedure:

$$\Delta'_{i+1} = B^*R^*RB\Delta_i - \sum_{k=1}^{i}(B^*R^*RB\Delta_i, \Delta_k)\Delta_k, \tag{3.28}$$

$$\Delta_{i+1} = \frac{\Delta'_{i+1}}{(\Delta'_{i+1}, \Delta'_{i+1})^{1/2}}. \tag{3.29}$$

After every iteration step the j largest eigenvalues of the operator B*R*RB are estimated with the eigenvalues of the tridiagonal j x j-matrix $E_j^T E_j$.

$$E_j = [RB\Delta_1, RB\Delta_2, \cdots, RB\Delta_j]. \tag{3.30}$$

Figure 3.4 shows the average percentage of the sum of the eigenvalues of B*R*RB obtained from the sum of the eigenvalues of $E_j^T E_j$ as a function of j. The Lanczos algorithm starts from a random initial vector Δ_1. This makes the matrix $E_j^T E_j$ and the estimate of its eigenvalues a stochastic process. The performance will also depend on the spectrum of eigenvalues of B*R*RB. These two effects make it necessary to average over a number of cases. The averaging is over the 100 sets of percentages obtained from the first 100 cases provided by the experiment. With the Lanczos algorithm the error growth is underestimated, because only j eigenvalues are estimated on a total of 30. There is no computational reason to do more than 10 iteration steps. After 10 iteration steps a brute-force determination of R becomes cheaper. From fig. 3.4 a suitable value of j is selected. With $j = 8$ about 84 % of the variance is explained on average. Also indicated in this figure is the error energy described with the j most rapidly growing eigenvectors. This shows the steepness of the spectrum of eigenvalues of the matrix $Q_A(t_f)$. This spectrum is only approximated by the eigenvalues of $E_j^T E_j$. The 8 most rapidly growing eigenvectors of $Q_A(t_f)$ explain 91 %
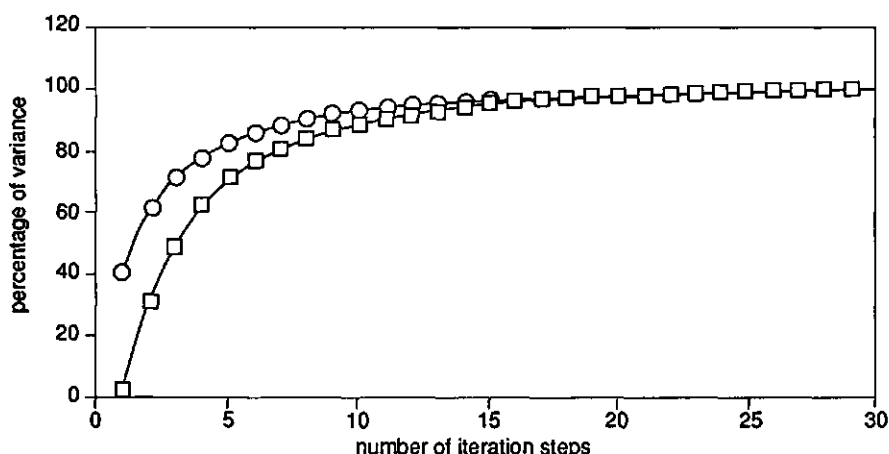
Figure 3.4: The efficiency of the Lanczos algorithm. The horizontal coordinate is the number of iteration steps. The squares give the average percentage of the error variance captured with a Lanczos procedure. Due to numerical errors this percentage may become larger than 100 %. The circles give the percentage of error energy described by the most rapidly growing vectors.

of the variance. The 84 % level is reached with only 6 eigenvectors. Thus, due to the approximate character of the Lanczos algorithm, 8 rather than 6 iteration steps are needed to capture 84 % of the error energy.

The subscript L denotes quantities based on the Lanczos algorithm and the operator B*R*RB. The location of the 84 % point for K_L is shown in fig. 3.3d. As a comparison of fig. 3.3c and fig. 3.3d shows, the Lanczos algorithm leads to almost perfect predictions of the variations of the probability distribution. The median and the 84 % point, which were computed with the Lanczos algorithm, give the same correlations with reality as the values from the algorithm which produces 30 eigenvalues. However the correlation for the 16 % point has dropped significantly. As explained by PH this is because 22 eigenvalues are taken to be zero. The Lanczos algorithm is only practical for estimating the dominant error structures. In addition, as can be deduced from table 3.1, the Lanczos algorithm causes a systematic underestimation of the error. The correlation coefficients are insensitive to this systematic error. Comparing reality, as plotted in fig.

3.3a, with the feasible forecast, as plotted in fig. 3.3d it is observed that the skill forecast carries information on the quality of the forecast.

3.4.5 Initial unit error sphere

As a final approximation the inhomogeneity of the error in the analysis is ignored. The distribution of the forecast errors is determined completely by the stability of the flow during the forecast run. The analysis error covariances are taken proportional to the Identity matrix:

$$Q_I(0) = \frac{\alpha}{30} I, \tag{3.31}$$

$$\alpha = \sum_{i=1}^{30} \lambda_i^2(\overline{Q}), \tag{3.32}$$

where the subscript I denotes this approximation. The $\lambda_i^2(\overline{Q})$ are the eigenvalues of $\overline{Q}$. The estimated covariance matrix the forecast error is:

$$Q_I(t_f) = \frac{\alpha}{30} RR^*. \tag{3.33}$$

The forecast of the 84 % point based on $Q_I(t_f)$ is shown in fig. 3.3e.

Comparing fig. 3.3a and fig. 3.3e and using table 3.1 the following shortcomings of the estimate K_I are noted. The range of values of K is larger than the range of K_I (a factor 4.4/24.6 against a factor 2.8/42.8). Thus, with K_I, the variability of the predictability is underestimated. On the basis of the limited variability in the distribution for K_I one might conclude that skill forecasts have little practical value. With each subsequent approximation in the skill prediction method the variations in the predicted probability distribution have been reduced. Thus skill forecasts appear less interesting than they really are. The last skill forecast shows the remarkably poor correlation coefficients: 0.56, 0.48 and 0.42. Finally it is remarked that K_I gives a significant overestimate of the error. Thus ignoring the data assimilation leads to almost useless skill forecasts.

table 3.2: quality of the approximations $t_f{=}95$								
	correlation with K			average values					
	16%	50%	84%	16%	σ	50%	σ	84%	σ
R^2	1.00	1.00	1.00	8.2	0.9	18.8	2.8	45.0	8.8
R_F^2	0.84	0.82	0.78	8.5	0.9	19.0	2.9	43.6	8.5
R_A^2	0.76	0.73	0.68	9.6	0.9	21.0	3.0	47.5	9.0
R_L^2	0.71	0.72	0.68	7.4	0.9	18.7	2.9	45.2	8.9
R_I^2	0.68	0.66	0.55	19.3	1.4	44.0	3.6	107.6	11.3

In validating the skill forecast one needs to realize that at the time of the analysis the climatological covariance matrix $\overline{Q}$ has, by definition, no predictive value for $Q(0)$. During the forecast integrations are done with the tangent linear equations. This adds relevant temporal information to the initially climatological error distribution. Consequently the predictive value of the error distribution increases during the course of the integration until after some time the linearity assumption breaks down. Using a linearity assumption both for the errors and for their validation , as done in this chapter, a constantly improving skill forecast is to be expected. For $t_f = 95$, which is the integration length used by PH, the experiment has been repeated. The quality information is given in table 3.2. As anticipated the quality of the skill forecasts has increased compared to the case with $t_f = 52$. The correlations for K_I are 0.68, 0.66 and 0.55. This characterizes the quality of the skill forecasts obtained by PH. Comparing K_I and K_L in table 3.2 the figures show that the Lanczos algorithm and the climatology assumption lead to an improved prediction, especially in the tail of the distribution. Moreover, the systematic errors in the prediction are considerably reduced.

3.5   Discussion

Skill forecasts for a low order T5 model are presented. The skill forecasts show obvious skill. However, one has to be cautious in interpreting the results to skill forecasts for realistic models.

In the T5 model environment the use of climatological covariances for the analysis error has a big negative impact on the quality of the skill forecast. No problem of similar magnitude occurs during the forecast run. So, in the environment of the T5 model, these results suggest the development of an updating strategy for these covariances. Because of the assumed near perfection of the model (only the confidence in decreasing error directions was limited) it takes a long time before old observations lose their value, and thus it pays to look far back in time. In this the T5 model may differ from a realistic model.

In a realistic experiment a few additional problems have to be faced. A discussion of some of them is given by Farrell (1990). The starting point, the determination of a climatological covariance matrix, requires a major effort. Both local and global information on the covariances of analysis errors is needed. In data-rich areas, such as the area above the United States, existing diagnostics of the analysis system performance could be used to get climatological covariances (Hollingsworth and Lönnberg 1989). For the determination of global errorfields and errorfields above data-void areas one probably has to perform an Observation System Simulation Experiment. Such an experiment involves the generation of a number of independent analyses. From the differences between analyses one gets statistical information on the analysis error. An example is the work by Daley and Mayer (1986).

Once a reliable climatological covariance matrix has been obtained one would like to have un updating strategy for the most energetic components of the analysis error. Courtier and Talagrand (1990) state that information gathered during the data assimilation may be useful to determine the covariance matrix for the analysis.

It is not clear at what rate information from old observations becomes irrelevant in a realistic model. Due to model errors this will be much faster than suggested by the present study. In fact Courtier and Talagrand (1987) do a data assimilation, where the previous analysis is not used in the cost function. It still appears to provide a reasonable analysis. Over the eastern Pacific Ocean they doubt the quality of the analysis because of the lack of observations in that region or upstream of it. In this area

the use of a previous analysis could have been helpful. If information on the covariances of the errors in the analysis would be available, one would surely be tempted to use it, both for the next data assimilation and for the actual skill forecast.

During the forecast run the imperfections of the model increase the error. The present method ignores these imperfections. An independent estimate of the contribution of the model errors needs to be added to the estimated forecast error. Due to the limited knowledge of the structure of model errors this will make the predicted probability distributions less variable in space and time.

For the computational cost the spectrum of eigenvalues, i.e. the number of dominating growth directions, is the key factor. This spectrum translates into a number of required iterations with the Lanczos algorithm. Spectra for eigenvalues are given by Farrell (1990). In a case with undamped baroclinic shear he finds a good separation of the eigenvalues. In a case with a barotropic jet the separation is less convincing. Based on an EOF analysis, a linearity assumption and results from the T5 model, PH gives an estimate for the dimension of the errorspace (i.e. for the separation of the eigenvalues). It is not clear to what extent this estimate relies on the linearity assumption or on the T5 model. One may have to wait for a spectrum of eigenvalues determined for a realistic model to get an impression of the feasability of global skill forecasts.

A practical method to limit the number of eigenvalues is the projection of the forecast error on a limited number of localized forecast parameters. This is discussed in a recent paper by Barkmeijer (1992).

Summarizing the present analysis of error growth in a low-order model, it is concluded that short-range skill-forecasts with large operational models are useless, if no attention is paid to the initial distribution of errors. The obvious way to make this feasible is the use of climatological covariances for the analysis error.

Acknowledgements

REFERENCES

Barkmeijer, J., 1992: Local error growth in a barotropic model. Preprint to *Tellus*

Bergman, K.H., 1979: Multivariate analysis of temperatures and winds using optimum interpolation. *Mon. Wea. Rev.*, 107, 1423-1444.

Cohn, S.E., and D.F. Parrish, 1991: The behavior of forecast error covariances for a Kalman filter in two dimensions. *Mon. Wea. Rev.*, 119, 1757-1785.

Courtier, P., and O. Talagrand, 1987: Variational assimilation of meteorological observations with the adjoint vorticity equation II: Numerical results. *Quart. J. Roy. Meteor. Soc.*, 113, 1329-1347.

Courtier, P., and O. Talagrand, 1990: Variational assimilation of meteorological observations with the direct and adjoint shallow-water equations. *Tellus*, 42A, 531-549.

Daley, R., and T. Mayer, 1986: Estimates of the global analysis error from the global weather experiment observational network. *Mon. Wea. Rev.*, 114, 1642-1653.

Dee, D.P., S.E. Cohn and M. Ghil, 1985: Systematic estimation of forecast and observation error covariances in four-dimensional data assimilation. Preprints, *Seventh Conference on Numerical Weather Prediction*, Boston, Amer. Meteor. Soc., pp. 9-16.

Draper, N.R., and H. Smith, 1966: *Applied Regression Analysis*, 407pp., John Wiley, New York.

Farrell, B.F., 1990: Small error dynamics and the predictability of atmospheric flows. *J. Atmos. Sci.*, 47, 2409-2416.

Gandin, L.S., 1963: Objective analysis of meteorological fields. *Gidrometeorologicheskoe Izdatelstvo* , Leningrad. English translation by: Israel Program for Scientific Translations, Jerusalem, 1965, 242pp. [NTIS N6618047, Library of Congress QC 996.G3313].

Ghil, M., S. Cohn, J. Tavantzis, K. Bube and E. Isaacson, 1981: Applications of estimation theory to numerical weather prediction. In *Dynamic Meteorology: Data Assimilation methods.* [eds: Bengtsson, L., M. Ghil and E. Källén] New York, Springer-Verlag, 139-224.

Ghil, M., and P. Malanotte-Rizzoli, 1991: Data assimilation in meteorology and oceanography. *Adv. Geophys.*, 33, 141-266.

Hollingsworth, A., and P. Lönnberg, 1989: The verification of objective analyses: Diagnostics of analysis system performance. *Meteorol. Atmos. Phys.*,40,3-27.

Houtekamer, P.L., 1991: Variation of the predictability in a low order spectral model of the atmospheric circulation. *Tellus*, 43A, 177-190.

Kalman, R.E., 1960: A new approach to linear filtering and prediction problems. Trans. ASME, Ser. D., J. Basic Eng., 82, 35-45.

Kendall, Sir, M., A. Stuart and J.K. Ord, 1986: *Kendall's advanced theory of statistics- 5 th ed, Vol 1: distribution theory,* Charles Griffin and Company Limited, London, 604 pp.

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

Lorenc, A.C., 1984: Analysis methods for the quality control of observations, in *Proceedings of the ECMWF workshop on the use and quality control of meteorological observations for numerical weather prediction,* 6-9 Nov. 1984, pp. 397-428.

Lorenc, A.C., 1986: Analysis methods for numerical weather prediction. *Quart. J. Roy. Meteor. Soc.*, 112,1177-1194.

Marchuk, G.I., 1974: The Numerical Solution of Problems of Atmospheric

and Oceanic Dynamics (in Russian), 387pp., Gidrometeoizdat, Leningrad, USSR. (English translation, Rainbow Systems, Alexandria, Va.).

Parlett, P., 1980: *The symmetric eigenvalue problem.* Series in computational mathematics, Prentice Hall, Engle wood Cliffs, New Yersey, 348pp.

Talagrand, O., and P. Courtier, 1987: Variational assimilation of meteorological observations with the adjoint vorticity equation I: theory. *Quart. J. Roy. Meteor. Soc.*, 113, 1311-1328.

Thacker, W.C., 1989: The role of the Hessian Matrix in Fitting Models to Measurements. *J. Geophys. Res.*, 94, 6177-6196.

Chapter 4

Global and local skill forecasts

Abstract

A skill forecast gives the probability distribution for the error in a forecast. Statistically well founded skill forecasting methods have so far only been applied within the context of simple models. In this chapter, the growth of internal errors is studied. This means that errors, which are already present in the estimate of the initial state, can grow only in accordance with the dynamics of a model. Errors in the description of the model itself are neglected. A three level quasi-geostrophic spectral model of the atmospheric circulation is used. It is trunctated at T21. It is shown, that a linear theory for the evolution of errors can be used for the first four days of a forecast. For the description of the global error, Monte Carlo methods are more efficient than methods based on the use of the adjoint of the tangent linear equations. The limitation to spatially local errors dramatically reduces the dimension of the error vectors. In that case, adjoint methods are the most efficient ones. Local skill forecasts for three days ahead are computed for a period of 24 consecutive days, using the T21 model and the adjoint of its tangent linear equations. The variability in the predicted distributions for the local errors is fitted with a 2 parameter stochastic model. Within the context of a perfect model assumption, providing perfect skill forecasts, the variability in the distribution of the error at day three is such that for equal quality forecasts the maximum extension of the forecast length is two days.

4.1   Introduction

Recently, skill forecasting techniques that are based on the use of the tangent linear model have been developed (Lacarra and Talagrand 1988; Barkmeijer 1992; Houtekamer 1991, 1992). Because the techniques assume linear error growth, and because they do not consider the effect of model errors, they may be applied only to short-range weather forecasts. The idea is to search, with a Lanczos algorithm, for the few directions in which the forecast is sensitive to errors (Lacarra and Talagrand 1988; Houtekamer 1991). The number of relevant error directions can be small as a consequence of the specific formulation of the problem (Barkmeijer 1992). In this case, which occurs naturally for local skill forecasts, it is not necessary to use a Lanczos algorithm.

The author has used a simple quasi-geostrophic 2-level model with a triangular truncation at T5, to propagate the use of a Lanczos algorithm (Houtekamer 1991), and to stress the importance of accurate probability distributions for the initial error (Houtekamer 1992). The conclusions may depend on the order of the model, which has only 30 components, and consequently some work is needed before the methods and conclusions can be applied to state-of-the-art models.

In this chapter, the relevance of methods based on the use of an adjoint is studied for a model with a resolution that is sufficient to describe the large scale circulation in the atmosphere. It is a quasi-geostrophic T21 model with three levels. The model has been developed at the European Centre for Medium-Range Weather Forecasts (ECMWF) by F. Molteni (Marshall and Molteni 1992). It is still simple enough to allow thousands of short term integrations.

In section 4.2, the data assimilation technique (Thacker 1989; Houtekamer 1992) is discussed briefly. The covariance matrix for the initial errors is obtained from this method. This matrix forms the starting point of the actual error growth analysis. The method is not feasible with state-of-the-art primitive equation models. A possibly feasible method based on a simplification of the Kalman filter was recently discussed by Dee (1991).

In section 4.3, a Monte Carlo technique is used to validate the assump-

tion of linear error growth. An ensemble of initial states is integrated both with the tangent linear model and with the full non-linear model. The differences are a measure for the degree of nonlinearity in the evolution of the ensemble.

It has been argued (Phillips 1986) that the dimension of the error vectors decreases as the length of the forecast period increases. In section 4.4, a Lanczos algorithm is used to estimate the dimension of the global error as a function of forecast time. This information is the basis of a comparison of the convergence of a Lanczos algorithm and a Monte Carlo algorithm. It will appear that, due to the high dimension of the global error, Monte Carlo methods are more efficient. In section 4.5, the same comparison is made for local skill forecasts, where probability distributions for only a few specific local forecast parameters at one specific forecast time are needed. In this case, a method using the adjoint equations proves to be more efficient than the Monte Carlo technique.

Skill forecasts are useful if the day to day variability of the width of the probability distributions is comparable to the width of the daily distributions. If this is not so, a climatology skill forecast is sufficient. In section 4.6, the day to day variability is quantified for a local skill forecast. Finally in section 4.7, the use of skill forecasts, under such variable conditions, is discussed.

In the discussion, an account of the current state-of-the-art with respect to skill forecasts is presented. Although the effect of model errors has largely been ignored one is now in the position to validate the proposed methods.

4.2   The data assimilation

Skill forecasts start from a covariance matrix for the error in the analysis. This matrix can be obtained as a by-product of the data assimilation. In this chapter, it is computed during a data assimilation based on a minimization of a quadratic cost function (Thacker 1989; Houtekamer 1992). For the analysis itself the estimate of the ECMWF is used. The error statistics for the observations and the covariance matrix of the previous

analysis field are used to compute the covariances for the analysis error. In this section, the error statistics for the observations are discussed. Next the data assimilation and the assumptions on model error are described. Finally some statistics on the analysis error are given.

4.2.1   The observations

The available radiosonde measurements of February 24, 1990 are taken to be representative for the observing network. All other measurements such as surface reports, aircraft reports and wind measurements are neglected. Because, for the present purpose, only error statistics need to be computed, only the error statistics of the geopotential height observations of the radiosondes are required. The observed values themselves are not necessary.

The assimilation period is 24 hours and starts at 12 GMT. The total number of radiosonde-observations, of which the statistics are used for one assimilation, is 3969. The spatially inhomogeneous radiosonde observing network will lead to a spatially inhomogeneous distribution for the analysis error.

Error statistics for a set of 3 observations from one radiosonde can be obtained from Lönnberg and Hollingsworth (1986). The root mean square (rms) errors σ_{200}, σ_{500} and σ_{800} of the geopotential heights Φ_{200}, Φ_{500} and Φ_{800} in meters at the model levels 200 mb, 500 mb and 800 mb are:

$$\sigma_{200} = 15m, \quad \sigma_{500} = 10m, \quad \sigma_{800} = 5m. \tag{4.1}$$

The correlations ρ of the errors ϵ_{200}, ϵ_{500} and ϵ_{800} are:

$$\rho(\epsilon_{200}, \epsilon_{800}) = 0.35, \quad \rho(\epsilon_{200}, \epsilon_{500}) = 0.50, \quad \rho(\epsilon_{500}, \epsilon_{800}) = 0.50. \tag{4.2}$$

These statistics are for the North-American radiosonde network. For simplicity, it is assumed that all observations have the same quality. Different radiosondes give independent errors. Thus the covariance matrix for the observational errors can be written as a matrix with at its diagonal three by three matrices for the covariances of the errors in the three measurements by one individual radiosonde. It is assumed that the observations

have Gaussian error statistics. This implies that stochastic variables which depend linearly on measurements will also have Gaussian distributions.

With the above error statistics the covariance matrix W expressed in m^2 for one radiosonde becomes:

$$W = \overline{(\epsilon_{200}, \epsilon_{500}, \epsilon_{800})(\epsilon_{200}, \epsilon_{500}, \epsilon_{800})^T} = \begin{pmatrix} 225 & 75 & 26.25 \\ 75 & 100 & 25 \\ 26.25 & 25 & 25 \end{pmatrix} \quad (4.3)$$

To get an impression of the vertical structure of the error introduced by the set of three measurements, an eigenvector analysis of the matrix W is performed. The eigenvectors of W are multiplied with the root of the eigenvalues to get three independent vectors in meters:

$$w_1 = \begin{pmatrix} -0.1 \\ -1.1 \\ 4.0 \end{pmatrix} \ w_2 = \begin{pmatrix} -3.7 \\ 7.1 \\ 1.9 \end{pmatrix} \ w_3 = \begin{pmatrix} 14.5 \\ 7.0 \\ 2.3 \end{pmatrix} \quad (4.4)$$

The projections of an ensemble of measurement errors on these vectors have standard normal probability distributions. In the following the notation $N(0, 1)$ is used for a standard normal distribution with expectation zero and unit variance. The vertical structure of the vector w_3, which has the highest amplitude, is equivalent barotropic. The term "equivalent barotropic" refers to the observation that it has the highest amplitude at the highest level, that its amplitude increases with height and that it has the same sign at the 3 levels of the model. The vectors w_1 and w_2 peak at respectively the 800 mb and 500 mb level. Thus the structure of errors in radiosonde measurements is mainly equivalent barotropic.

4.2.2 Least squares fitting

The model is fitted to the data from the previous analysis and the observations with a least squares method (Thacker 1989; Houtekamer 1992). The least squares fit is determined by the minimum of a costfunction $J(x)$ that is quadratic in the differences between the data d and their model counterpart m(x), where x is the value of the state vector at the beginning of the assimilation:

$$J(\mathbf{x}) = 0.5(\mathbf{d} - \mathbf{m}(\mathbf{x}))^T \mathsf{S}(\mathbf{d} - \mathbf{m}(\mathbf{x})). \tag{4.5}$$

Here the vector $\mathbf{d}$ consists of both the radiosonde observations and the previous analysis. The previous analysis is the analysis which was determined during the previous data assimilation. Thus, it is at the beginning of the present assimilation period. The vector $\mathbf{d} - \mathbf{m}(\mathbf{x})$ gives the difference between the model run and the data. The matrix S is the inverse of the covariance matrix for the errors in the data. It is constructed from a submatrix for the measurement errors, which has been discussed already, and a submatrix for the error in the previous analysis, which is output from the previous data assimilation as is shown below.

The accuracy of the least squares fit is given by the inverse of the Hessian matrix H which consists of the second derivatives of the cost function:

$$\mathsf{H} = \left(\frac{\partial \mathbf{m}}{\partial \mathbf{x}}\right)^T \mathsf{S} \left(\frac{\partial \mathbf{m}}{\partial \mathbf{x}}\right), \tag{4.6}$$

where $\partial \mathbf{m} / \partial \mathbf{x}$ is the Jacobian matrix. To compute the Jacobian, one row at a time, the tangent linear equations are integrated once for perturbations in all model coordinates. To compute model counterparts of the data, streamfunction fields have to be transformed to geopotential height fields. For this the linear balance equation is used. This approach is not allowed near the equator (Daley 1983). It appears that no easy solution exists for this problem (Daley 1978; Tribbia 1981), because any mass-wind law is inaccurate near the equator. Consequently, the skill forecast, which is produced, will have a low quality near the equator.

The inverse of the Hessian matrix gives the covariances at the beginning, at time t_0, of the assimilation period. They are integrated to the time t_1 of the analysis:

$$\mathsf{Q} = \mathsf{R}(t_1, t_0)\mathsf{H}^{-1}\mathsf{R}^T(t_1, t_0), \tag{4.7}$$

where $\mathsf{R}(t_1, t_0)$ is the resolvent operator of the tangent linear equation between times t_1 and t_0. It represents the integration of a small perturbation with the tangent linear equations from time t_0 until t_1. The matrix $\mathsf{R}(t_1, t_0)$ is a by-product of the computation of the Jacobian.

It has been shown that, the covariance matrix for the error in the analysis can be computed from measurements and a previous analysis, both

with complete error statistics. This covariance matrix and the ECMWF-analysis provide the required data for the skill forecast. Before the covariance matrix is used in the next assimilation, a correction for the model errors is added. This is discussed in the next subsection.

4.2.3 Model Errors

In the data assimilation, the model is assumed to be perfect. However, due to model errors, the previous analysis will not be as accurate as anticipated. During the previous data assimilation, a covariance matrix Q was computed for the then actual analysis errors. This matrix now serves as a covariance matrix for the error in the previous analysis, where the previous analysis is now defined to be at the initial time of the present assimilation period. A matrix M for the model error is added to the covariance matrix for the error in the previous analysis:

$$Q_\delta = Q + M \tag{4.8}$$

This matrix Q_δ is subsequently inverted to get the part of the matrix S for the accuracy of the previous analysis. Little is known about the model error. Dee et al (1985a, 1985b) propose a method, which is based on an extension of the Kalman filter, to estimate the structure of the model error. Due to the computational cost of their filter, it has not yet been put to practice. One can assume that the model error is randomly distributed over all modes. Thus the model error is of white noise type (Phillips 1986; Cohn and Parrish 1991). In a large model, one has to limit the model error which is assumed to be present at modes with a large two dimensional wavenumber n (Bennet and Budgell 1987). This problem does not yet occur at $n = 21$. In fact, due to the strong damping at the smallest scales of the model, an extra source of model errors is added for waves that are close to the truncation limit ($n \geq 19$). The following expression is introduced for the model error:

$$M_{ij} = 0.5\delta_{ij}(1 - h(n_i))10^{-9} + 0.5h(n_i)h(n_j)V_{ij} \tag{4.9}$$

$$h(n) = \begin{cases} 0 & n < 19 \\ 1 & n \geq 19 \end{cases}, \tag{4.10}$$

where n_i and n_j are the two-dimensional wavenumbers corresponding to mode number i and j. The number of modes in the model is 1449. The stepfunction h is used to select the highest three wavenumbers. The matrix V for the variability is obtained from observational data:

$$V = \frac{1}{248} \sum_{i=1}^{248} (x_i - \overline{x})(x_i - \overline{x})^T, \qquad (4.11)$$

where the vectors $\{x_i, i = 1, \cdots, 248\}$ are the ECMWF analyses of the periods December 1, 1989 - January 31, 1990 and December 1, 1990 - January 31 ,1991. The ECMWF-analysis at 200 mb, 500 mb, 700 mb and 850 mb is used. The values at the 800 mb model level are obtained from a linear interpolation. Successive ECMWF analyses are 12 hours apart.

The low-wavenumber part of Eq. (4.9) corresponds to white noise. To each individual mode the same amount of error kinetic energy is added. The constant $0.5 \cdot 10^{-9}$ implies a rms geopotential height error from 1 meter at the equator to 7 meters at the poles.

The high-wavenumber part of Eq. (4.9) acts as a source of errors at the smallest scales. The added uncertainty is 50 % of the variability V at these scales. So this added source term becomes smaller at smaller scales (Farrell 1990).

Figure 4.1 gives the distribution of the kinetic energy at 200 mb as a function of two dimensional wavenumber n. At each value of n the energy, in the corresponding $2n+1$ modes with zonal wavenumber between $-n$ and n, is added. The figure displays the energy distribution for the atmosphere, for the variability in the atmosphere, for the model error and for the analysis error. The distributions at 500 mb and 800 mb (not shown) are similar though they have less kinetic energy. The variability (transient energy) is maximal at $n = 8$. Beyond $n = 10$ almost all energy is in the variability (e.g. Daley and Mayer 1986). The white noise part of the model error increases linearly, with the number of modes, below $n = 19$. At $n = 19$ the energy of the model error jumps to 50 % of the energy of the variability. The curve for the analysis error is discussed in the next section.

Energy distribution at 200 mb
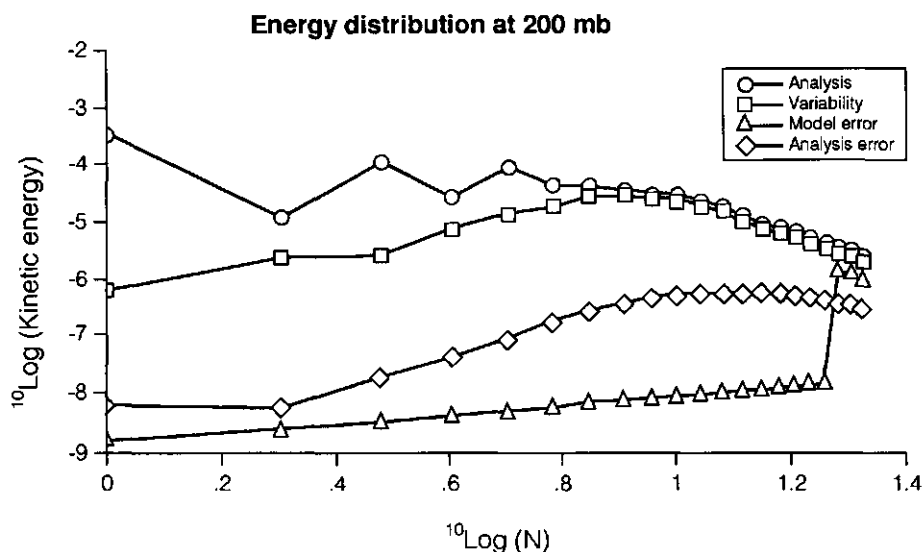
Figure 4.1: The kinetic energy distribution at 200 mb as a function of two dimensional wavenumber N. Shown are the energy distributions for the analysis (circles), the variability(squares), the model error (triangles) and the analysis error at Dec 8, 12 GMT (diamonds).

4.2.4 Structure of the analysis error

The data-assimilation starts at December 1, 12 GMT 1989. The ECMWF analysis for this date is the starting point of a 24 hour reference run. These 24 hours form the first assimilation period. To start the assimilation cycle one can initially take the diagonal of the matrix V for the variability as the covariance matrix for the error in the "previous analysis":

$$Q_{ij} = \delta_{ij}V_{ij} \tag{4.12}$$

The assimilation is then performed with as result the covariance matrix Q for the error in the analysis of December 2, 12 GMT. The matrix Q is updated for model errors with Eq. (4.8). For the next assimilation period the ECMWF analysis for December 2, 12 GMT and the matrix Q_δ can be used. This assimilation-cycle is repeated for a total of 7 days after which

the covariance matrix Q_δ describes reasonable values and structures. Starting from December 8, the covariance matrix can be used for predictability experiments. Before reporting on these experiments, a description of the spectrum, the vertical structure and the horizontal structure of the analysis error at December 8 is given. These properties are studied with an ensemble of random analysis errors, which is generated as follows.

As a preliminary step an eigenvector analysis of the matrix Q has to be done. This gives a set of eigenvectors $\{q_i, \ i = 1, \cdots, N\}$ and eigenvalues $\{\lambda_{i,0}^2, \ i = 1, \cdots, N\}$. The number N of model coordinates is 1449 for the T21 model. The subscript 0 refers to the initial time. Members of an ensemble of initial perturbations have independent random projections $\{d_i, \ i = 1, \cdots, N\}$, with $N(0,1)$ distribution, on the vectors $\{\lambda_{i,0}q_i, \ i = 1, \cdots, N\}$. Thus a random initial perturbation $\epsilon(t_1)$ is obtained using:

$$\epsilon(t_1) = \sum_{i=1}^{N} d_i \lambda_{i,0} q_i. \tag{4.13}$$

Writing d for the vector with random coeffients and writing B for the matrix which has in its columns the products of the roots of the eigenvalues with the corresponding eigenvectors one obtains the equivalent expression:

$$\epsilon(t_1) = \mathsf{B}d. \tag{4.14}$$

An ensemble is generated from successive sets of N random numbers. In the limit of an infinite number of ensemble members the covariance matrix Q can be reconstructed from the ensemble statistics:

$$Q = \mathsf{B}\overline{dd^T}\mathsf{B}^T = \mathsf{B}\mathsf{B}^T. \tag{4.15}$$

An ensemble, generated with Eq. (4.14), is used to study properties of the analysis error. First, in figure 4.1, the analysis error is given as a function of two-dimensional wavenumber n. For n is 1 to 8, the error energy is about 1 % of the transient energy. The fractional magnitude of the errors then starts to increase, until it is about 10 % of the transient energy at $n = 21$.

To get an impression of the vertical structure of the analysis error, a (3×3) covariance matrix is computed. The element at column i and row j is proportional to the covariance between the error at level i and the error

at level j. Here an inner product for kinetic energy is used to compute covariances. The eigenvectors z_1, z_2 and z_3 of the covariance matrix are:

$$
z_1 = \begin{pmatrix} 0.2 \\ -2.3 \\ 5.2 \end{pmatrix}, z_2 = \begin{pmatrix} -4.9 \\ 7.6 \\ 3.6 \end{pmatrix}, z_3 = \begin{pmatrix} 14.5 \\ 7.8 \\ 3.1 \end{pmatrix}, \tag{4.16}
$$

where the eigenvectors are scaled such that the projections on them have equal probability distributions and such that the maximum number in Eq. (4.16) is equal to the maximum number in Eq. (4.4). Comparing the two equations, it can be seen that the eigenvectors are very similar. As expected, most of the error is in the equivalent barotropic mode z_3. The two baroclinic modes carry less energy. It may be concluded that, the vertical structure of the observational errors is still present in the vertical structure of the analysis error.

The rms error of the 500 mb analysis, for the Northern Atlantic ocean and Europe is shown in figure 4.2. The maximum amplitude of 8.8 meters is above the Atlantic. The minimum of 2.2 meters is over Europe, where one has many observing stations. The exact structure of the rms-field above the oceans changes slightly from one analysis to the next (not shown), because the advection of information depends on the flow. The constant radiosonde network inserts information at fixed times and places and this restricts variability that can occur in the distribution for the analysis error. This is a first suggestion, that a climatology approximation for the analysis error may indeed be reasonable (Houtekamer 1992). The consequences of such an approximation are quantified, in the section on the variability of the local spread. A comparison with Daley and Mayer (1986) shows that, their estimate of error amplitudes is roughly twice as large as the present one. The cause of the underestimation is probably to be found in the hypothesis on model errors and in the insufficient resolution of the T21-model. Working with a T21-model, the error of representativeness should be taken larger than computed by Lönnberg and Hollingsworth (1986). To take this type of error properly into account, one has to study the effect of the unresolved scales. The error of representativeness at nearby observing stations will be correlated and thus the covariance matrix for the observational error will loose its very simple structure. In the

rms value of the analysis error at 500 mb
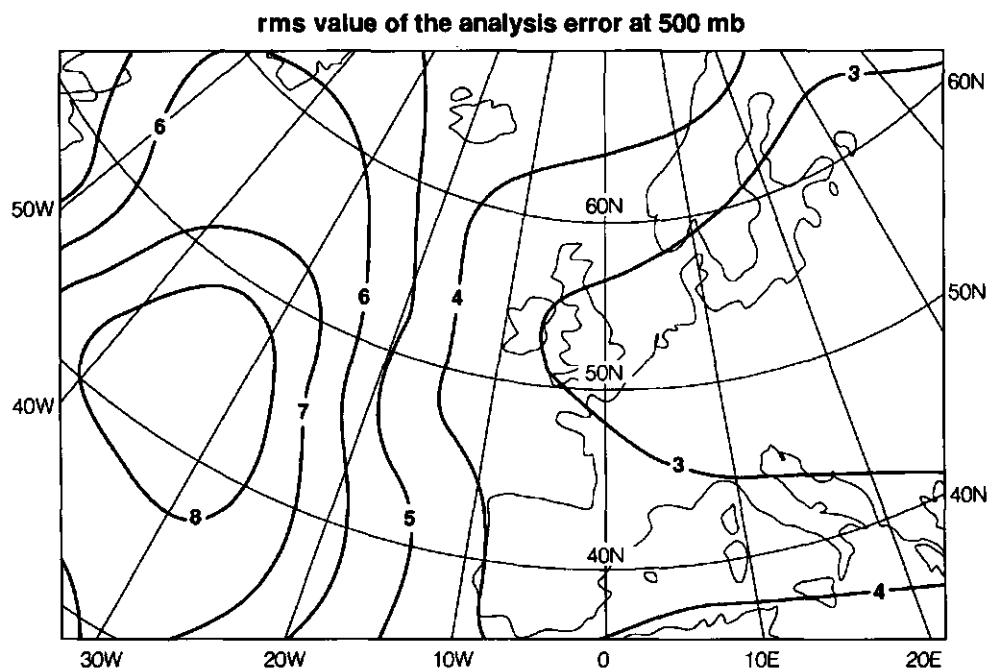
Figure 4.2: The rms geopotential height error at the 500 mb level for December 8, 12 GMT. Only the Northern Atlantic and Europe are shown

following experiments the analysis error is simply multiplied with a factor of 2 before a skill forecast is performed. Thus the general spatial and spectral characteristics of the error field are maintained. This ensures that no problem with transient dynamics (spin-up) will occur. It is not expected that the conclusions are influenced by these choices because, as will be demonstrated, the results are not very sensitive to the exact specification of the covariances for the analysis error.

4.3   Linear error growth

After some forecast period the linearity assumption for the evolution of errors does not hold anymore. According to Lacarra and Talagrand (1988)

the linearity assumption is valid for at least two days. Vukićević (1991) performed three case studies with a state-of-the-art primitive equation limited-area model. Her conclusion is that the linear theory gives a good representation of errors for a 36 hour forecast. In this section, the length of the time domain over which the linear approximations are valid is studied more precisely, because skill forecasts will be most useful near the ultimate range of their validity.

Predictability is a scale and frequency dependent property (e.g: Boer 1984; Lacarra and Talagrand 1988; Van den Dool and Saha 1990). Different aspects of the forecast error may become non-linear at a different time. In this section, the forecast time is estimated at which the effect of non-linearity becomes evident, in the kinetic energy of the global error.

A Monte Carlo experiment is performed with an ensemble of 200 random initial perturbations to the analysis of December 8, 12 GMT, 1989. The ensemble of analysis errors is generated using Eq. (4.14) and a set of random vectors. From all perturbations two model runs are started. One uses the full non-linear 3 level T21 model. The other uses the tangent linear version of the T21 model to integrate the perturbation to the time of the forecast. The reference orbit itself is computed from an integration of the ECMWF analysis with the T21 model. The forward integration of the ensemble provides information on the nonlinearity of error growth. To study the nonlinearity the following vector is defined:

$$\delta(t) = \epsilon(t) - \epsilon_{NL}(t) \tag{4.17}$$

where $\epsilon(t)$ is found from the forward integration with the tangent linear model and $\epsilon_{NL}(t)$ from the forward integration with the full non-linear model. Initially the two errorvectors are identical:

$$\epsilon_{NL}(t_1) \equiv \epsilon(t_1) \tag{4.18}$$

To measure distance a norm $\| \cdot \|$ is used which is proportional to the root of the kinetic energy. Thus $\| \delta(t) \|^2$ is a measure for the difference in kinetic energy between the linear and the non-linear run.

Table 4.1 summarizes, in arbitrary units, the results of a 5-day Monte Carlo run starting at December 8, 12 GMT. Ensemble averages are given

Table 4.1: Linearity of global errors

	Non-linear		Linear		
day	$\|\epsilon_{NL}\|^2$	σ	$\|\epsilon\|^2$	σ	$\|\delta\|^2$
0	57	6	57	6	0
1	73	9	73	9	4
2	100	15	104	16	20
3	140	24	156	33	58
4	207	40	260	74	151
5	304	61	447	137	362

as well as the rms differences with the average values. The energy of $\epsilon(t)$ is too high on average. This overestimation occurs especially for large errors. At day five it is already 47 %. Apparently an important property of nonlinearity is to damp errors. This occurs because non-linear error growth is bounded by the size and shape of the attractor. With small errors nonlinearities may lead both to larger and to smaller errors. In this respect, the observation by Vukićević (1991) is mentioned that the pattern similarity between $\epsilon(t)$ and $\epsilon_{NL}(t)$ is greater than the similarity in amplitudes. This suggests that a skill forecast, based on linear theory, will show better performance in terms of anomaly correlations than in terms of rms errors.

As one can see, from the values of $\|\delta\|^2$ in table 4.1, the error growth is essentially linear upto day 2. Between day 3 and day 5, the error growth becomes non-linear. At day 5 the distance, between the linear and the non-linear errorvector, is comparable to the length of the individual error vectors. Linear errorgrowth methods should not be used beyond day four. Non-linear estimates of errorgrowth may be valid beyond day four. Their validity will eventually be limited by the non-linear interaction of model errors with internal errors.

4.4   Dimension of the global forecast errors

The dimension of the analysis error is quite high. Independent observational errors are made everywhere. As time evolves the errors organize

in patterns. When the forecast provides no longer information, the error vector is like a random state vector. Thus the dimension of the error vectors must be identical to the dimension of the state vectors. Houtekamer (1991) proposed to use a Lanczos algorithm for the computation of the preferential error directions in short-range forecasts. For this method to be efficient, the number of important error growth directions must be small. To get relevant results from the algorithm, the error growth must be predominantly linear. It is not a priori clear that, the required reduction in dimension is obtained within the first four days of linear error growth. A Lanczos algorithm is used, and its rate of convergence is determined, for short-range skill forecasts starting from the ECMWF analysis for Dec 8, 12 GMT, 1989. More specifically, the largest eigenvalues of the matrix RQR^T are estimated. The matrix R is the resolvent of the integration, with the tangent linear model, to the time of the forecast. The matrix Q consists of the covariances for the analysis error. The coordinates are chosen such that adjoint operators are identical to transpose operators. This means that an innerproduct of two vectors is simply the sum of their multiplied coefficients. In principle one may also use a standard coordinate system. The transformation of adjoint to transpose operators or vice versa is then given by an additional linear operator.

From the Monte Carlo result (table 4.1), one has an unbiased estimate of the kinetic energy in the forecast error. From 200 iterations with the Lanczos algorithm, an alternative estimate of the kinetic energy in the forecast error is obtained. After every iteration, the explained variance is computed and scaled with the Monte Carlo estimate. The result is shown in figure 4.3. Looking at the day 1 forecast, a rather slow convergence is observed. One needs 72 iterations to explain 50 % of the global error. For a two day forecast, 44 iterations are needed to obtain the same percentage. A three day forecast requires 28 iterations and a four day forecast requires 18 iterations. It is concluded that, the dimension of the forecast error shrinks rapidly when the forecast time increases.

It is not clear how many iterations are acceptable in the operational practice. At this time it is simply demanded that the proposed method is more efficient than a Monte Carlo experiment. To estimate the random

Dimension reduction
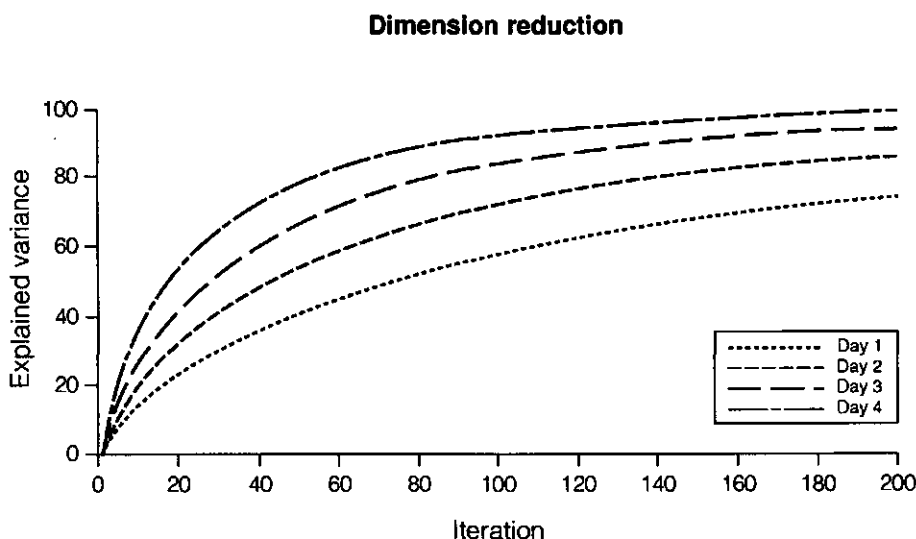
Figure 4.3: Dimension reduction. The horizontal coordinate is the number of iterations used by the Lanczos algorithm. The vertical coordinate displays the cumulative variance. Results are show for a 1-day, 2-day, 3-day and a 4-day forecast.

error in the four day forecast with a Monte Carlo technique the spectrum of 200 eigenvalues that has been obtained from the Lanczos algorithm can be used. The average forecast error energy $\overline{K}$ is:

$$\overline{K} = \sum_{i=1}^{200} \lambda_{i,4}^2, \tag{4.19}$$

where the $\{\lambda_{i,4}^2, \ i = 1, \cdots, 200\}$ are the estimated eigenvalues of the co-variance matrix at day 4. A Monte Carlo run, which starts from a random perturbation as obtained from Eq. (4.14), has random projections, with $N(0,1)$ distributions, on the eigenvectors of the covariance matrix for the forecast error. The forecast error energy K associated with one run is simulated with:

$$K = \sum_{i=1}^{200} \lambda_{i,4}^2 d_i^2, \tag{4.20}$$

where the $\{d_i, \ i = 1, \cdots, 200\}$ are a new set of random numbers with independent $N(0,1)$ distributions. An ensemble with N_{MC} sets of random

numbers is constructed. The average forecast error energy $\overline{K}$ is estimated, with the average value $\hat{K}$ of K in the ensemble. To obtain the rms error $\sigma(\hat{\overline{K}})$ in the Monte Carlo estimate $\hat{\overline{K}}$ of $\overline{K}$, several such experiments are performed. One obtains:

$$\sigma(\hat{\overline{K}}) \approx 0.25 N_{MC}^{-1/2}\overline{K}. \tag{4.21}$$

The dependence on the root of N_{MC} is accurate for large values of N_{MC} as one expects from the Central Limit theorem (e.g. Kendall et al. 1986). The constant 0.25 was obtained from the experiment. Equation (4.21) implies that, after one run, the fractional uncertainty, in the estimate of the average error energy, is only about 25 %. This implies that, the probability distribution for the global forecast error energy is narrow. After 25 Monte Carlo computations, the average forecast error energy is estimated with 5 % rms error and no bias. If a simplified tangent linear model is used for the Monte Carlo experiment then 2 Monte Carlo integrations can be done at the cost of 1 iteration with the Lanczos algorithm. So 12 iterations are about as expensive as 25 Monte Carlo runs. After 12 iterations, see figure 4.3, the Lanczos algorithm still has a systematic error of 59 %. In addition, it is remarked that a Monte Carlo integration up to four days supplies as a by-product information on the error at day 1,2 and 3, whereas the Lanczos algorithm gives only results for the forecast error at day 4. As has been seen, the Lanczos algorithm converges most rapidly at day 4. This is due to the steepness of the spectrum of eigenvalues. A Monte Carlo estimate performs better with a flat spectrum. It must be concluded that the Lanczos algorithm, even when it is applied for its most favourable forecast period, is inferior to the Monte Carlo technique if one wishes to estimate global forecast errors.

4.5 Local forecast errors

In this section local error growth is considered. In this case one is interested only in the probability distribution of some local properties. This reduces the dimension of the problem. Attention is restricted to the geopotential height errors η_i at 500 mb at nine locations $\{L_i, \quad i = 1, 2, \cdots, 9\}$

Table 4.2: Linearity of local errors

day	Non-linear		Linear		
	$\parallel \epsilon_{NL} \parallel_L^2$	σ	$\parallel \epsilon \parallel_L^2$	σ	$\parallel \delta \parallel_L^2$
0	625	305	625	305	0
1	1659	893	1651	881	40
2	4835	2912	4900	3087	365
3	12310	9208	13389	10350	2092
4	23031	17789	27142	20397	9643
5	37882	28087	41921	32529	30677

above the Northern Atlantic and above Europe. The coordinates are:

$L_1 : (-11^o.3E, 36^o.0N)$ $L_2 : (-16^o.9E, 52^o.6N)$ $L_3 : (-28^o.1E, 69^o.2N)$
$L_4 : (+5^o.6E, 36^o.0N)$ $L_5 : (+5^o.6E, 52^o.6N)$ $L_6 : (+5^o.6E, 69^o.2N)$
$L_7 : (+22^o.5E, 36^o.0N)$ $L_8 : (+28^o.1E, 52^o.6N)$ $L_9 : (+39^o.4E, 69^o.2N)$

The locations are chosen at points on the Gaussian grid. The nine points are chosen far apart, with the objective to get nine rather independent local errors. The errors $\{\eta_i, \ i = 1, \cdots, 9\}$ in the geopotential height are computed from the global error $\epsilon(t_f)$ in the forecast at time t_f with a projection operator P:

$$\eta = \mathsf{P}\epsilon(t_f), \tag{4.22}$$

where η is expressed in meters. The cost-function H_{FC} for the local quality of the forecast is:

$$H_{FC} = (\mathsf{P}\epsilon(T), \mathsf{P}\epsilon(T)) = \sum_{i=1}^{9} \eta_i^2 = \parallel \epsilon(T) \parallel_L^2 \tag{4.23}$$

The subscript L refers to a local norm based on the error in geopotential at 9 points. Table 4.2 gives an analysis of the nonlinearity of the local error. Again this result is for the forecast starting on December 8, 12 GMT. The units in this table are m^2. The results, definitions and conclusions are similar to the ones for the global error growth. One should not use linear methods for the prediction of the local error beyond day 4. In the following only 3-day forecasts are discussed. Writing the cost function in terms of the analysis errors gives:

$$H_{FC} = (\mathsf{PRBd}, \mathsf{PRBd}) = (\mathsf{B}^*\mathsf{R}^*\mathsf{P}^*\mathsf{PRBd}, \mathsf{d}) \qquad (4.24)$$

Again, all coordinates are chosen such that adjoint operators correspond to transpose operators thus one has for instance:

$$\mathsf{PRB} = (\mathsf{B}^T\mathsf{R}^T\mathsf{P}^T)^T = (\mathsf{B}^*\mathsf{R}^*\mathsf{P}^*)^T \qquad (4.25)$$

The projection, to a nine dimensional space, reduces the number of non-zero eigenvalues of the (1449×1449) matrix $\mathsf{B}^*\mathsf{R}^*\mathsf{P}^*\mathsf{PRB}$ to nine. These nine eigenvalues are identical to the eigenvalues of the (9×9) covariance matrix $\mathsf{PRQR}^*\mathsf{P}^*$ for the local forecast errors. It follows that, the eigenvalue problem can be solved using only the application of the operator $\mathsf{B}^*\mathsf{R}^*\mathsf{P}^*$ to the nine unit vectors of the low dimensional phase space. The operator PRB follows from a transposition using Eq. (4.25). If B is known already, the computational cost is dominated by the nine backward integrations with the adjoint model. For the 3-day forecast starting December 8, 12 GMT, the eigenvalues $\{\mu_i^2, \ i = 1, \cdots, 9\}$ of the local covariance matrix $\mathsf{PRQR}^*\mathsf{P}^*$ are: $6600m^2$, $3328m^2$, $1444m^2$, $1168m^2$, $720m^2$, $628m^2$, $340m^2$, $184m^2$ and $148m^2$. The first eigenvalue explains 45 % of the variance. This can happen because the errors in the 9 points are correlated. Error patterns have structures which, in spite of the significant distance between grid points, extend over several points on the Gaussian grid.

The eigenvalues $\{\mu_i^2, \ i = 1, \cdots, 9\}$ correspond to the eigenvectors $\{v_i, \ i = 1, \cdots, 9\}$ of the matrix $\mathsf{B}^T\mathsf{R}^T\mathsf{P}^T\mathsf{PRB}$. The vectors $\{\mathsf{B}v_i, \ i = 1, \cdots, 9\}$ give the initial perturbation fields which cause the forecast errors $\{\mathsf{PRB}v_i, \ i = 1, \cdots, 9\}$. For a random initial perturbation $\epsilon(0)$, the projections on the modes $\{\mathsf{B}v_i, \ i = 1, \cdots, 9\}$ have expectation zero and unit-variance. Inspection of these fields would show which initial features are critical for the forecast. Such studies have been advocated by Errico and Vukićević (1992). Of use to the forecaster are the fields $\{\mathsf{PRB}v_i, \ i = 1, \cdots, 9\}$, which give the structure of the forecast error.

The fact that 45% of the local forecast error may come from one single mode has consequences for a Monte Carlo experiment. If, like in the previous case of global errors, one has many relevant modes then an accidentally small projection on one mode may be "compensated" by an accidentaly large projection on another. In the present case, much more

runs are needed to get the same fractional accuracy. Using the eigenvalues for December 8, 12 GMT, the rms error $\sigma(\hat{\overline{H_{FC}}})$ in the estimate $\overline{H_{FC}}$ of the ensemble average $\overline{H_{FC}}$ of the local cost function H_{FC} becomes:

$$\sigma(\hat{\overline{H_{FC}}}) \approx 0.78 N_{MC}^{-1/2} \overline{H_{FC}} \tag{4.26}$$

Thus after 9 Monte Carlo runs, which corresponds in cost to the Adjoint Equation algorithm, the fractional error is 26 %. Thus, the Adjoint Equation algorithm is more efficient for the determination of local error distributions. The high number of required Monte Carlo runs indicates that H_{FC} has a broad probability distribution. This implies that random effects, as opposed to dynamical effects, have a strong influence on the success of the local forecast. The dynamical effects, as quantified by the variations in the distributions, must be large for a skill forecast to provide useful information.

4.6   Variability of the local spread

This section deals with the temporal variability in the rms value of the probability distribution for the local error in geopotential height. This rms value is called the spread. It can be used by forecasters to assess the reliability of their product. For a useful application, the skill forecast should add information to the experience of the forecaster. It will do so only if it shows significant variability in time. To validate a skill forecast, one should perform a large number of computations of the spread and compare these values with the forecast errors that have actually occurred. From this comparison, one obtains information on the quality of the skill forecast, because the spread must show a correlation with the forecast error. This correlation will be less than unity because the spread of a distribution does not uniquely determine an individual realization. Therefore, before such an experiment is done, one would like to quantify the amount of success that can be expected. For this purpose, local skill forecasts for the 3-day forecast are made during a period of 24 days. This information is used to quantify the temporal variability in the local spread with a two-parameter stochastic model. The two parameters are fitted to the 24 local values of

the spread. In the next section, this two-parameter model will actually be used to generate a hypothetical large experiment with 4×10^6 forecasts of the spread and corresponding forecast errors. This will then provide a preliminary view on the use of actual skill forecasts which are based on an analysis of the stability of the flow against perturbations.

For each of the 24 days the 9×9 local covariance matrix $\overline{\eta \eta^T}$ is computed:

$$\overline{\eta \eta^T} = \mathsf{PRQR}^*\mathsf{P}^*. \qquad (4.27)$$

The mean square value $S(i)^2$ of the geopotential height error η_i at location L_i is at diagonal element i of this matrix $\overline{\eta \eta^T}$. For the 9 points L_i the rms value $S(i)$ is given in table 4.3. The overall magnitude of the spread agrees with observed errors in operational forecasts (e.g: ECMWF 1990). Apparently, after the multiplication of the analysis error with a factor of 2, a reasonable estimate of the forecast errors can be obtained.

The variability of the spread is quantified with a model which is similar to the one introduced by Kruizinga and Kok (1988). In fact results from the present work, which are based on considerations of internal dynamics, will be compared with their purely statistical analysis of data. The value $S(i)$ for the spread of the error at grid point i is modelled by:

$$S(i) = S_m(i) \exp(\beta_i d_{1i}), \qquad (4.28)$$

where β_i is a measure for the day to day variability of the spread at point i. The number d_{1i} represents a daily draw from a $N(0,1)$ distribution. The amplitude of the error is scaled with $S_m(i)$. Thus the spread $S(i)$ of the distribution at point i is modelled with a stochastic variable. A negative random number d_{1i} corresponds to a small spread. For $d_{1i} = 0$ the spread assumes its median value $S_m(i)$. With a skill forecast, one does actually compute the variation in the spread $S(i)$ based on the stability properties of the flow. For the hypothetical large experiment, the stochastic model is used to simulate such a computation.

The parameters $S_m(i)$ and β_i are estimated, with the maximum likelihood method, from the 24 daily rms values at the 9 points in table 4.3. That is, for the 24 values of the spread the natural logarithm of Eq. (4.28) is taken. The maximum likelihood estimate of the constants $\log S_m(i)$ and β_i is:

$$\log S_m(i) \;=\; \overline{\log S(i)} \tag{4.29}$$
$$\beta_i^2 \;=\; \overline{(\log S(i) - \log S_m(i))^2}. \tag{4.30}$$

The bottom lines of table 4.3 show the 9 fitted pairs of parameters. The values of β are around 0.35. This is higher than the observational estimate $\beta = 0.2$ of Kruizinga and Kok (1988). In part this may be due to minor differences in the design of the experiment. However, the present study cannot support the assumption, made by Kruizinga and Kok, that values of the spread at grid points as far apart as the points $L_1, \cdots, L_9$ have a strong positive correlation. For the 12 pairs of neighbouring points, one finds 7 negative and 5 positive correlations between the sets of 24 values of the spread. From this it is concluded that the correlations are significantly less than plus one. This may be the explanation for the apparently lower variability in their study.

To quantify the influence of the inhomogeneous statistics of the analysis error, additional skill forecasts are made, based on the assumption of white noise statistics of the analysis error. In this case, PRR*P* is the local covariance matrix. A table similar to table 4.3 is computed (not shown). The 9 correlations between the corresponding columns of the two experiments are: 0.93, 0.95, 0.96, 0.97, 0.98, 0.93, 0.91, 0.97 and 0.94. The estimation error in the individual correlations is about 0.02. Assuming that the spread values predicted with the inhomegeneous statistics for the analysis error are correct, it is concluded that still about 90 % of the variance in the spread can be predicted with homogeneous statistics for the analysis error. Next skill predictions starting from a constant inhomogeneous covariance matrix are checked against those starting from variable covariances. The covariance matrix for the analysis error at December 8, 12 GMT is the constant matrix. The 9 correlations between the two sets of numbers are: 0.99, 0.96, 0.99, 0.99 0.98, 0.97, 0.96, 0.99 and 0.98.

Table 4.3: Variability of the local spread of the 3-day forecast error. Units are meters of geopotential height at 500 mb.

initial	Local spread values at 9 locations								
day	$S(1)$	$S(2)$	$S(3)$	$S(4)$	$S(5)$	$S(6)$	$S(7)$	$S(8)$	$S(9)$
Dec 8	59	60	46	37	34	29	28	28	18
Dec 9	47	69	47	52	33	33	27	33	15
Dec 10	43	52	44	27	75	37	21	26	20
Dec 11	65	36	47	53	46	51	21	34	29
Dec 12	49	52	43	50	52	44	25	39	29
Dec 13	59	59	37	48	44	31	28	50	28
Dec 14	59	44	30	67	29	41	36	110	35
Dec 15	53	39	27	42	27	38	33	91	44
Dec 16	54	42	25	36	42	28	43	55	32
Dec 17	61	34	20	38	57	23	23	32	30
Dec 18	47	29	27	33	72	30	20	48	26
Dec 19	27	65	33	36	31	29	37	78	28
Dec 20	28	94	37	28	36	33	29	48	25
Dec 21	64	44	38	22	71	45	25	34	28
Dec 22	42	32	30	26	74	59	25	36	60
Dec 23	84	73	46	23	72	41	34	27	60
Dec 24	63	38	66	44	45	84	30	38	43
Dec 25	23	74	67	32	29	53	32	27	37
Dec 26	32	55	78	32	26	39	42	21	44
Dec 27	29	43	46	25	32	74	50	25	58
Dec 28	34	47	55	26	26	46	35	20	29
Dec 29	56	44	78	32	27	37	21	22	25
Dec 30	38	27	51	28	20	58	23	26	36
Dec 31	27	58	36	40	33	40	35	33	27
$S_m(i)$	45	48	41	35	39	41	29	36	32
β_i	0.35	0.32	0.36	0.29	0.40	0.31	0.25	0.46	0.35

An impressive 96 % of the variability of the local spread values can be explained with time independent covariances for the analysis error. This observation suggests an enormous simplication of the computations for a skill forecast. It does not support the earlier conclusion of Houtekamer (1992), based on the results from a low order model, that proper attention needs to be paid to the initial distribution of errors.

4.7   Using the local skill prediction

The prediction of the spread can be used as a measure for the expected quality of the forecast. As was mentioned already, a skill forecast is useful if the temporal variability of the spread is large. This section discusses a few hypothetical large experiments which quantify the use of skill forecasts, done for the geopotential error at one grid point, under different conditions of variability. A stochastic model, with 2 random variables, is used to model the forecast error and the forecast of the spread:

$$S_M = \exp(\beta_M d_1) \tag{4.31}$$
$$\eta_M = d_2 S_M, \tag{4.32}$$

where d_1 and d_2 are independent random variables. The temporal variability of the spread is determined by the parameter β_M. The subscript M denotes the use of a stochastic model. The modelled value of the local spread is S_M. The realization of the local forecast error is η_M. The median of the spread is taken to be unity. This does not influence the correlation between spread and skill, because both are proportional to the choice for the median.

The hypothesis is made that the skill forecast is perfect. This means that the probability distribution of the forecast error, and thus the spread of the ensemble, are assumed to be known without error. In this section this "knowledge" is simulated with the draw d_1 from a $N(0,1)$ distribution. The forecast error η_M is determined by S_M and the random variable d_2. In the first experiments d_2 is drawn from a $N(0,1)$ distribution. This is consistent with Gaussian errors in the observations. The sensitivity to the choice for the distribution will be investigated later.

A large ensemble of S_M and η_M is obtained from Eq. (4.31) and (4.32) with $4 \cdot 10^6$ random numbers d_1 and d_2. The experiment is repeated for fixed values of β_M between 0.001 and 4.0. The correlation $\rho(S_M, |\eta_M|)$ between S_M and $|\eta_M|$ is a usual measure for the performance of a skill forecast. It may be obtained from the above data or from the expression (A. Otten, private communication):

$$\rho^2(S_M, |\eta_M|) = \frac{2}{\pi} \frac{1 - \exp(-\beta_M^2)}{1 - 2\exp(-\beta_M^2)/\pi} \tag{4.33}$$

In addition the probablity P_u, that the forecast is unsuccessful under the condition that the spread of the ensemble is large, is computed:

$$P_u(\beta_M) = P(|\eta_M| > |\eta_M|_m \,|S_M > 1), \tag{4.34}$$

where the subscript m denotes the median value. The median of the spread is one in this experiment. In practice, a skill forecaster has a computed value of spread. If this value is above the median, he states that the spread is larger than usual. After such a statement, the probability of a bad forecast is given by P_u. The deviation of P_u from 0.5 shows the value of the skill forecast in warning against bad forecasts. Similarly P_s is the probability that a forecast is successful under the condition that the spread is small:

$$P_s(\beta_M) = P(|\eta_M| < |\eta_M|_m \,|S_M < 1). \tag{4.35}$$

The deviation of P_s from 0.5 gives a measure of the confidence in the forecast which one may have after a positive statement by the skill forecaster. Because both $|\eta_M|$ and S_M are compared with their median value, it can be shown that the probabilities P_u and P_s are equal.

Table 4.4 gives the result of the experiment. In the case of vanishing variability of the spread, its prediction is not needed. When the variability increases the skill forecast becomes useful. As stated above, realistic values for the parameter β_M are between 0.3 and 0.4. Having a look at the numbers, one may not be impressed by the success of the perfect skill forecast. Even in the case of extreme variability of the spread ($\beta_M = 4$) the correlation is only 0.79.

Table 4.4: Performance of a perfect local skill forecast. β_M measures the temporal variability. ρ is the correlation between spread and skill. P_u measures the success in warning against a bad forecast.

β_M	$\rho(S_M, \mid \eta_M \mid)$	P_u
0.001	0.02	0.50
0.1	0.13	0.53
0.2	0.26	0.57
0.3	0.36	0.60
0.4	0.46	0.63
0.5	0.53	0.66
0.6	0.59	0.68
0.8	0.68	0.72
1.0	0.73	0.76
2.0	0.78	0.85
4.0	0.79	0.92

Table 4.5: Local contingency table for a 3-day forecast. The skill has a Gaussian distribution.

skill	spread of the ensemble			
	0- 25 %	25 - 50 %	50 - 75 %	75 - 100 %
0 - 25 %	0.367	0.267	0.214	0.156
25 - 50 %	0.322	0.272	0.228	0.172
50 - 75 %	0.234	0.276	0.266	0.227
75 -100 %	0.074	0.187	0.290	0.448

For a typical value of $\beta_M = 0.35$ a more detailed skill prediction is performed. Both the spread and the skill are divided into four classes with 25 % of the points, rather than into two classes each with 50 % of the points as in the previous experiment. The spread increases with the number of the class. The skill (or the success) of a forecast decreases with the number of the class. The probability that the skill is in a particular class under the condition that the spread is in a particular class is computed. In the case of vanishing variability of the spread or vanishing quality of the skill forecast all probabilities are 0.25. For the present case ($\beta_M = 0.35$) results

Table 4.6: Local contingency table. The skill has a Cauchy distribution.

skill	location of the first quartile.			
	0- 25 %	25 - 50 %	50 - 75 %	75 - 100 %
0 - 25 %	0.355	0.267	0.216	0.162
25 - 50 %	0.282	0.270	0.245	0.203
50 - 75 %	0.205	0.245	0.271	0.280
75 -100 %	0.159	0.218	0.268	0.355

are given in table 4.5. The rows give the same class of skill. The columns give the same class of spread. Especially in the corners of the table the deviation from 0.25 is large. Therefore it is in the prediction of extreme cases that a skill forecast is most useful and easiest to validate. If for instance the spread is very small, the probability of a very low skill is only 0.074. If on the other hand the spread is very large this probability is about six times higher. It must be emphasized that a large number of data points is needed to get accurate statistics for the contingency table.

There is some evidence that observations have, after quality control, Gaussian error statistics (Lorenc, 1984). The Gaussian distribution is stable. This means that the sum of 2 Gaussian variables has a Gaussian distribution. In particular, the linearity of all operators implies Gaussian statistics of the forecast error. In the previous experiments, this is reflected in the choice of a $N(0,1)$ distribution for the random variable d_2. One might instead assume that the errors in observations have a Cauchy distribution (e.g. Holt and Crow 1973). A Cauchy distribution is also stable. Consequently, in this case the forecast error will also have a Cauchy distribution. A Cauchy distribution has much more points far away in the tail of the distribution than a Gaussian distribution. These points would correspond to gross errors in the observations. In fact the higher moments of this distribution do not exist (e.g. Kendall et al. 1986). The results for this experiment, which is again done with $\beta_M = 0.35$, are given in table 4.6. Because the second moment of the distribution does not exist, S_M is now identified with the location of the first quartile. At the first quartile 25% of the cases have occurred. It is a measure of the width of a distribution which is defined for a Cauchy distribution. Comparing with

Table 4.7: Local contingency table. The skill has a homogeneous distribution.

skill	spread of the ensemble			
	0- 25 %	25 - 50 %	50 - 75 %	75 - 100 %
0 - 25 %	0.373	0.264	0.209	0.154
25 - 50 %	0.368	0.267	0.211	0.153
50 - 75 %	0.254	0.310	0.253	0.182
75 -100 %	0.004	0.158	0.327	0.511

table 4.5, one observes that the main difference is in the lower left-hand corner. This is to be expected. Due to the large number of points in the tail of the distribution for d_2 a stable flow does not imply a small error in the forecast.

Alternatively, one might reason that nonlinear effects tend to reduce the number of cases with a large forecast error. In the most extreme case, with no tail at all, the forecast error is modelled with a homogeneous distribution for d_2. The homogeneous distribution is not stable. It is selected on the basis of heuristic arguments on nonlinear dynamics and not on statistical theory. Comparing the result with 4.5 one observes the opposite effect as before. The number of points in the lower left corner has become much smaller. This happens because a low value of the spread and a homogeneous distribution exclude the possibility of a large forecast error.

As has been shown, the results of the experiment depend on the assumptions for the stochastic model. In an actual large experiment, one may use the computed set of values for the spread as a definition of a probability distribution function. For each value of the spread, one may simulate a number of realizations with a Gaussian probability distribution function. Unexpected results may possibly have to be explained in terms of deviations from a Gaussian.

The skill forecast could be used to extend or limit the range of a forecast. At day 3, the rms value of the local forecast error increases with 40 % a day. With values of β_M of about 0.35, this implies that with perfect knowledge of the spread, a forecaster has the possibility to extend or limit the forecast time with about one day. Or alternatively if a forecaster gave

only high quality forecasts, by a limitation of the forecast time, he may extend the forecast time with one day in an average case and two days in a reliable case.

In reality a skill forecast will not be perfect and it will show less performance. In view of the statistical problems with the validation of the quality of skill forecasts, it may prove to be difficult to convince forecasters to use skill forecasts to their full potential. For the time being, forecasters are likely to make a more cautious use of skill predictions. At the same time, skill forecasters should be careful in rejecting their methods on the basis of a low correlation between a limited number of values of spread and skill (e.g. Branković et al. 1990).

4.8   Discussion

An environment is created which contains the most essential properties of a forecasting system. The response of a data assimilation system to errors in the measurements of the geopotential heights is studied. With a 3 level quasi-geostrophic atmospheric model, the analysis errors are integrated to the time of the forecast. The predictability of the forecast error and more in particular methods to estimate temporal and spatial variations in the predictability have been studied.

It appears that linear methods, for instance methods using adjoint equations, are useful for the first four days of a skill forecast. After four days non-linear errorgrowth and the interaction between model errors and internal errors become important.

The global forecast error has a complex structure during the regime of linear error growth. As a consequence, a Lanczos algorithm as proposed by Houtekamer (1991) does not efficiently lead to a skill forecast. For global errors a Monte Carlo method is superior.

More useful results are obtained in the case of local error growth. Here the dimension of the errorspace is low by definition. The adjoint method gives results of infinite precision, at least within the assumption of linear errorgrowth, and is thus to be preferred over a Monte Carlo method.

The variability of the spread is quantified using 24 3-day local skill

forecasts. It appears that due to this variability the length of the useful forecast period varies with about two days. Based on the 24 skill forecasts, on some assumptions on the statistical properties of temporal variability of the spread, and on an assumed distribution to couple spread and skill, a 2-parameter stochastic model is fitted to the data. This model is used to evaluate the use of a hypothetical perfect skill forecast. It appears that skill forecasts are difficult to validate. It is recommended, for future validations, that the temporal variability of the computed spread values is combined with a statistical model of the forecast error to estimate the performance of a hypothetical perfect skill forecast, before conclusions are drawn on the quality of the skill forecasting procedure.

Part of the experiment was repeated with very simple assumptions on the analysis error. The use of white noise statistics for the analysis error led to almost the same skill predictions. Thus error statistics, which are obviously different at the time of the analysis, may be very similar at the time of the forecast. This may happen because unrealistic initial states rapidly converge to atmospheric states. This insensitivity towards the initial error statistics may explain the success of the skill forecasts by Barkmeijer and Opsteegh (1992), who use a white noise formulation of the analysis error statistics.

Following the suggestion of Houtekamer (1992), a constant inhomogeneous covariance matrix was used for the statistics of the analysis error. This showed an even closer agreement with the reference result. It is concluded that, the variability of the flow during the assimilation period is hardly relevant for the statistics of the forecast error at a forecast period of three days. In principle, variations due to changes in the observing network may still be important. One should expect that the sensitivity of the forecast error to the structure of the analysis error increases with decreasing forecast period.

This study neglects the effect of model errors. This causes an underestimation of the forecast error (e.g. Branković et al. 1990). This assumption probably causes a decrease in the correlation between the observed skill and the predicted spread. A consistent method to account for model error is beyond the scope of this chapter. The only purpose is to demonstrate

the feasibility of skill forecasting methods. The study indicates that the problem is easier than expected on the basis of studies with a simple model (Houtekamer 1991, 1992). The computational costs are small, because a simple 3 level quasi-geostrophic T21 model is used for evolution of the errorfields. In fact the cost of the proposed methods is negligible compared to the cost of obtaining a reference forecast. It would seem that time is ready for a validation with real data.

Acknowledgements

REFERENCES

Barkmeijer, J., 1992: Local error growth in a barotropic model, *Tellus A.*

Barkmeijer, J., and J.D. Opsteegh, 1992: Local skill prediction with a simple model, in *ECMWF workshop on predictability.*

Bennett, A.F., and W.P. Budgell, 1987: Ocean data assimilation and the Kalman filter: Spatial regularity. *J. Phys. Oceanogr., 17*, 1583-1601.

Boer, G.J., 1984: A spectral analysis of predictability and error in and operational forecast system. *Mon. Wea. Rev., 112*, 1183-1197.

Branković, č., T.N. Palmer, F. Molteni, S. Tibaldi and U. Cubasch, 1990: Extended-range predictions with ECMWF models: Time-lagged ensemble forecasting, *Quart. J. Roy. Meteor. Soc., 166*, 867-912.

Cohn, S.E., and D.F. Parrish, 1991: The behavior of forecast error covariances for a Kalman filter in two dimensions. *Mon. Wea. Rev.*, *119*, 1757-1785.

Daley, R., 1978: Variational non-linear normal mode initialization. *Tellus*, *30*, 201-218.

Daley, R., 1983: Linear non-divergent mass-wind laws on the sphere. *Tellus*, *35A*, 17-27.

Daley, R., and T. Mayer, 1986: Estimates of Global Analysis Error from the Global Weather Observational Network. *Mon. Wea. Rev*, *114*, 1642-1653.

Dee, D.P., 1991: Simplification of the Kalman filter for meteorological data assimilation. *Quart. J. Roy. Meteor. Soc.*,*117*, 365-384.

Dee, D.P., S.E. Cohn and M. Ghil, 1985a: Systematic estimation of forecast and observation error covariances in four dimensional data assimilation. Preprints, *Seventh Conference on Numerical Weather Prediction*, Boston, Amer. Meteor. Soc., pp 9-16.

Dee, D.P., S.E. Cohn, Dalcher, A. and M. Ghil, 1985b: An efficient algorithm for estimating noise covariances in distributed systems, *IEEE Trans. Automat. Contr.*, *AC-30*, 1057-1065.

ECMWF, 1990: ECMWF Forecast report 49 December 1989 - February 1990, *ECMWF*, Reading, UK.

Errico, R. and T. Vukićević, 1992: Sensitivity analysis using an adjoint of the PSU/NCAR mesoscale model. *Mon. Wea. Rev.*, *120*, 1644-1660.

Farrell, B.F., 1990: Small error dynamics and the predictability of atmospheric flows. *J. Atmos. Sci.*, *47*, 2409-2416.

Holt, D.R. and E.L. Crow, 1973: Tables and Graphs of the stable probability density functions.
J. Res. Nat. Bur. Standards-B. Math. Sc.,*77B*, 143-198.

Houtekamer, P.L., 1991: Variation of the predictability in a low order spectral model of the atmospheric circulation. *Tellus*,*43A*,177-190.

Houtekamer, P.L., 1992: The quality of skill forecasts for a low order spectral model. Accepted by *Mon. Wea. Rev.*

Kendall, Sir, M., Stuart, A. and J.K. Ord, 1986: *Kendall's advanced theory of statistics, 5 th ed, Vol 1: distribution theory*, Charles Griffin and Company Limited, London, 604 pp.

Kruizinga, S. and C.J. Kok, 1988: Evaluation of the ECMWF experimental skill prediction scheme and a statistical analysis of forecast errors. In *Proceedings of the ECMWF workshop on predictability in the medium and extended range*, ECMWF, Reading, UK, 403 - 415.

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

Lönnberg, P. and A. Hollingsworth, 1986: The statistical stucture of short-range forecast errors as determined from radiosonde data. Part II: The covariance of height and wind errors. *Tellus, 38A*, 137-161.

Lorenc, A.C., 1984: Analysis methods for the quality control of observations, in *Proceedings of the ECMWF workshop on the use and quality control meteorological observations for numerical weather prediction*, 6-9 Nov. 1984, pp. 397-428.

Marshall, J. and F. Molteni, 1992: Towards a dynamical understanding of planetary-scale flow regimes. Submitted to *J. Atmos. Sci.*

Phillips, N.A., 1986: The spatial statistics of random geostrophic modes and first-guess error. *Tellus, 38A*, 314-332.

Thacker, W.C., 1989: The role of the Hessian matrix in fitting models to measurements. *J. Geophys. Res., 94*, 6177-6196.

Tribbia, J., 1981: Normal mode balancing and the ellipticity condition. *Mon. Wea. Rev., 109*, 1751-1761.

Van den Dool, H.M. and S. Saha, 1990, Frequency dependence in forecast skill, *Mon. Wea. Rev., 118*, 128-137.

Vukićević, T., 1991: Nonlinear and linear evolution of initial forecast errors. *Mon. Wea. Rev., 119*, 1602-1611.

114

Samenvatting

Voorspelbaarheid in modellen van de atmosferische circulatie

In de weersvoorspellingen is de laatste dertig jaar veel vooruitgang geboekt. De gebruikte modellen hebben een hoge graad van perfektie bereikt. Het wordt dan ook steeds moeilijker om zinvolle verbeteringen in de modellen aan te brengen. Desondanks komt het regelmatig voor dat een weersvoorspelling niet goed uitkomt. Voor een deel is dit mislopen van een voorspelling een voorspelbaar verschijnsel. Wanneer het weer slechts langzaam van karakter verandert kan men denken dat het weerbericht relatief betrouwbaar zal zijn. In de winter kan bijvoorbeeld een lang aanhoudende vorstperiode optreden. Omdat er weinig verandert lijkt het makkelijk om een weerbericht te maken. In andere gevallen, bijvoorbeeld als depressies in hoog tempo over de oceaan in de richting van Nederland trekken, lijken de voorspellingen veel moeilijker. Het liefst heeft men natuurlijk in alle gevallen een duidelijke uitspraak over de betrouwbaarheid van het weerbericht en niet alleen als bepaalde vuistregels toepasbaar zijn. Een dergelijke uitspraak is nuttig omdat in het geval van een goed te vertrouwen voorspelling besloten kan worden het weerbericht met een aantal dagen te verlengen. In een wat onduidelijker situatie kan de voorspeltermijn juist verkort worden.

In dit proefschrift worden methoden ontwikkeld om iets te zeggen over de betrouwbaarheid van het weerbericht. Hiertoe wordt het ontstaan en het groeien van fouten onderzocht. Fouten ontstaan direkt op het moment van het meten. De meetfout wordt niet met opzet gemaakt en is onbekend. Het gevolg van de meetfout is dat niet precies bekend is vanuit welke toestand men de berekeningen moet starten. De onbekende beginfout zal gaan groeien totdat zij, hoewel nog steeds onbekend, uiteindelijk ontoelaatbaar groot is. Het weerbericht heeft zijn waarde dan verloren. Van de onbekende meetfout is wel iets te zeggen. Van meetinstrumenten is meestal bekend hoe nauwkeurig ze zijn. Men kan bijvoorbeeld een thermometer hebben waarmee op 0.5 graden nauwkeurig kan worden afgelezen. Dit vertaalt zich in het drie-daagse weerbericht dan bijvoorbeeld in een onzekerheid van 2 graden in het geval van een stabiele situatie en een

onzekerheid van 4 graden in het geval van een onstabiele situatie. In beide gevallen kan men geluk hebben en zowel een meetfout als een voorspelfout van 0.0 graden hebben. Van te voren kan men zeggen dat de voorspelling in het stabiele geval het betrouwbaarste is. Men moet er echter rekening mee houden dat een weersvoorspelling voor een deel een kansproces is. Het zal altijd een verrassing blijven hoe goed een voorspelling uitkomt. Vanwege dit altijd aanwezige toevalselement worden uitspraken over de betrouwbaarheid van het weerbericht kansverwachtingen genoemd.

De basisgedachte achter het maken van kansverwachtingen is eenvoudig. Men kan net doen alsof ergens een meetfout gemaakt is. Deze fout wordt nu met opzet aangebracht. Vervolgens doet men opnieuw alle dingen die nodig zijn om een weerbericht te maken. De verandering die hierdoor optreedt in het weerbericht geeft aan hoe groot het effekt van de aangebrachte meetfout is. Voor de genoemde verandering maakt het geen verschil of de meetfout met opzet geïntroduceerd is, zoals in dit experiment, of toevallig gemaakt is, wat in werkelijkheid gebeurt. Als meetfouten aangebracht worden in alle metingen tegelijkertijd, ontstaat een indruk van de gevoeligheid van het weerbericht voor deze set meetfouten. Om een goede kansverwachting te maken moet het effekt van meerdere sets van meetfouten doorgerekend worden. Deze methode van werken wordt de Monte-Carlo methode genoemd. Theoretisch gezien heeft de Monte-Carlo methode geen beperkingen. In de praktijk is de methode niet haalbaar omdat ze teveel rekenwerk vereist. Het weerbericht is dan al achterhaald voordat de kansverwachting beschikbaar komt.

Om toch een kansverwachting te maken moet men zich twee beperkingen opleggen. Een eerste beperking is de veronderstelling dat de fout in het weerbericht klein is. In de praktijk betekent dit dat alleen uitspraken gedaan kunnen worden over de eerste drie of vier dagen van een voorspelling. Deze veronderstelling houdt in dat de fouten beschouwd kunnen worden als kleine afwijkingen op een redelijk goed uitkomende voorspelling. Voor de beschrijving van de ontwikkeling van deze afwijkingen kan een eenvoudig model gebruikt worden. Voor de redelijk goed uitkomende voorspelling kan de al beschikbare voorspelling van het meest succesvolle en vaak ook meest ingewikkelde weermodel overgenomen wor-

den. Het gebruik van een eenvoudig model verhoogt de haalbaarheid van kansverwachtingen doordat de benodigde rekentijd belangrijk gereduceerd wordt.

Ten tweede wordt de kansverwachting geografisch beperkt. Aangenomen wordt dat alleen de betrouwbaarheid van het lokale weerbericht interessant gevonden wordt. Voor het Nederlandse publiek is het in het algemeen niet interessant hoe betrouwbaar de verwachting voor bijvoorbeeld IJsland is. Deze tweede veronderstelling kan gebruikt worden om de aanpak van het probleem om te draaien. Bij de Monte-Carlo methode wordt steeds gekeken hoe een set meetfouten doorwerkt in het totale weerbericht. Er kan voor een soort voorspelfout echter ook selektief gekeken worden naar de meetfouten die deze fout kunnen veroorzaken. Door de beperking tot een lokale kansverwachting is het aantal soorten voorspelfouten waar men in geïnteresseerd is klein. Men kan zich bijvoorbeeld afvragen hoe gevoelig de driedaagse temperatuur- en windvoorspelling voor De Bilt is voor fouten in alle afzonderlijke meetstations. Voor de beantwoording van deze vraag is een techniek ontwikkeld door Russische en Franse meteorologen. Na een aantal wiskundige manipulaties wordt van het simpele model de "toegevoegde" versie verkregen. Dit toegevoegde model wordt gebruikt om de genoemde voorspelfouten terug te ontwikkelen in de tijd van het tijdstip waarvoor de voorspelling gedaan wordt naar het tijdstip waarop de voorspelling uitgegeven wordt. Als men slechts geïnteresseerd is in de betrouwbaarheid van een beperkt aantal onderdelen van de voorspelling kan volstaan worden met een even beperkt aantal terugontwikkelingen met het toegevoegde model. Omdat het toegevoegde model is afgeleid van een simpel model kost de hele procedure erg weinig rekentijd.

In dit proefschrift wordt uitgelegd waarom en hoe technieken op basis van toegevoegde modellen gebruikt kunnen worden voor het maken van kansverwachtingen. In eerste instantie worden methoden ontwikkeld binnen de context van een simpel model. Het model beschrijft het weer op het noordelijk halfrond met slechts dertig getallen. Net als in meer complexe en meer realistische modellen van de atmosfeer kunnen kleine beginfouten in dit simpele model snel groeien. De eenvoud van het model maakt het echter mogelijk om de voorgestelde methoden goed uit te testen. Dit is be-

langrijk omdat experimenten met realistische modellen, zoals in het laatste hoofdstuk van dit proefschrift, vaak ingewikkeld en duur zijn. Men moet dan van te voren goed weten wat men gaat doen. In hoofdstuk 2 wordt een rekenwijze gepresenteerd waarmee gezocht kan worden naar die elementen van een globale of lokale voorspelling die het gevoeligst zijn voor fouten in de begintoestand. De methode maakt gebruik van toegevoegde modellen en zal verder "Lanczos-algoritme" genoemd worden naar de ontwerper van het zoekalgoritme.

Voor de berekeningen van een weerbericht kunnen starten, moeten eerst de metingen verzameld worden. De computer voert berekeningen uit met behulp van een rooster met punten. Op iedere punt van het rooster is kennis van, onder andere, de temperatuur, de luchtdruk en de windsnelheid noodzakelijk. De waarneemstations liggen als regel niet op de roosterpunten. Sommige roosterpunten, zoals op de oceanen, zijn ver verwijderd van waarneemstations. De manier waarop de her en der gedane waarnemingen omgezet worden naar een vorm waar de computer mee kan gaan rekenen wordt in hoofdstuk 3 besproken. Er wordt een verband gelegd tussen fouten in metingen of zelfs het geheel ontbreken van metingen en fouten in de begintoestand van een voorspelling. Ruwweg kan gezegd worden dat de begintoestand altijd het minst nauwkeurig is boven de oceanen waar weinig metingen gedaan worden. Aangetoond wordt dat een goede statistische beschrijving van de fouten in de begintoestand een noodzakelijke voorwaarde is voor het maken van een kansverwachting. Het verkrijgen van zo'n beschrijving is erg duur. Men wil daarom zoveel mogelijk vereenvoudigingen aanbrengen. Men kan er bijvoorbeeld van uitgaan dat de fouten in de begintoestand iedere dag dezelfde statistische eigenschappen hebben. Deze eigenschappen hoeven maar éénmaal bepaald te worden en het blijkt dat er nog steeds een redelijk goede kansverwachting gegeven kan worden.

In hoofdstuk 4 worden de mogelijkheden voor een toepassing in de praktijk besproken. Hiertoe dienen zowel de normale weersvoorspelling als een redelijk simpel model gebruikt te worden. Het blijkt dat in een globaal weerbericht veel verschillende foutenpatronen kunnen optreden. Een Lanczos-algoritme heeft veel rekentijd nodig om al deze patronen

vinden. Het algorithme heeft zelfs meer rekentijd nodig als een, ook al onhaalbare, Monte-Carlo methode. Voor lokale kansverwachtingen zijn methoden die toegevoegde modellen gebruiken het meest efficiënt. Aangetoond wordt dat gedurende ongeveer vier dagen de fouten zo klein zijn dat het gebruik van de methode te rechtvaardigen is. Met betrekking tot de resultaten van hoofdstuk 3 kan een zeer hoopvolle conclusie getrokken worden. De gesuggereerde vereenvoudiging: "de begintoestand van de voorspelling is iedere dag even betrouwbaar" geeft veel betere resultaten als op grond van hoofdstuk 3 verwacht mocht worden. Voor de kansverwachtingen echt gebruikt kunnen gaan worden moet hun kwaliteit eerst getest worden. Door het al genoemde kanselement is het moeilijk te bewijzen dat een kansverwachting te vertrouwen is. Er zijn heel veel kansverwachtingen en mislopende voorspellingen nodig voor men weet of de kansverwachtingen goed werken. Gelukkig zijn de voorgestelde methoden zo goedkoop dat snel voor een groot aantal oude weersvoorspellingen alsnog een kansverwachting gemaakt kan worden. Als de methode werkt voor historische weersvoorspellingen kan men vertrouwen hebben dat ze ook voor nieuwe voorspellingen zal werken. In de huidige voorspelpraktijk is men meestal voorzichtig. Er wordt altijd rekening gehouden met een slecht uitkomende voorspelling. Als men echter over een goede kansverwachting beschikt, dan kan men besluiten de weersverwachting in stabiele situaties met één of twee dagen te verlengen.

Curriculum Vitae

Peter Houtekamer is in 1962 geboren in de gemeente Noordoostpolder. Het diploma VWO behaalde hij in 1981 aan het Christelijk Lyceum voor Zeeland te Goes. Daarna studeerde hij sterrenkunde aan de Rijksuniversiteit te Leiden. In 1984 behaalde hij het kandidaatsexamen met als hoofdvakken sterrenkunde en natuurkunde en als bijvak wiskunde. Het doctoraalexamen haalde hij in 1987 met als hoofdvak sterrenkunde en als bijvak wiskunde.

In 1988 trad hij in dienst als Onderzoeker In Opleiding bij de werkgemeenschap Meteorologie en Fysische Oceanografie van de Nederlandse Organisatie voor Wetenschappelijk Onderzoek. Het eerste jaar werkte hij vier dagen per week op het Mathematisch Instituut van de Rijksuniversiteit te Utrecht en één dag per week op het Instituut voor Meteorologie en Oceanografie dat eveneens tot de Rijksuniversiteit te Utrecht behoort. Het onderzoek is voorgezet op het Koninklijk Nederlands Meteorologisch Instituut. Gedurende het promotie onderzoek heeft hij drie conferenties, twee workshops, één zomerschool en acht buitenlandse instituten bezocht.

Dankwoord

Als eerste bedank ik Rien van de Weygaert die me wees op de vacature voor het promotie onderzoek. Mijn studievrienden Jean-Paul Koninx, Rien van de Weygaert, Henk Latour, Ignaz Wanders en Erwin van Soest dank ik voor de gezellige maaltijden en de enerverende risklessen.

Mijn moeder, Nelie, Onno, Gilbert en Monique dank ik voor hun vertrouwen en de rustige weekenden.

Met mijn (ex)-huisgenoten Maaike en Hans heb ik me 's ochtends vaak wakker gekletst.

Met Theo Opsteegh en Johan Grasman heb ik vrijwel wekelijks een voortgangsgesprek gehad. Zij hebben blunders in de kiem gesmoord en vele ideeën aangedragen. Zij hebben mij geleerd hoe een artikel geschreven moet worden maar weigerden consequent om als medeauteur op te treden.

Roger Daley heeft op een moment waarop ik volkomen vast zat een bruikbare oplossing aangedragen. Zijn milde en stimulerende kritiek hebben mij belangrijk geholpen.

Ik dank mijn collega's uit de voorspelbaarheidsgroep voor de hulp bij het onderzoek. Jan Barkmeijer heeft mij geholpen bij het ontwikkelen van toegevoegde modellen. Frank Selten heeft velerlei problemen voor mij opgelost. Ruben Pasmanter heeft mij aan stabiele kansverdelingen geholpen.

Ik dank de medewerkers van de studio van het KNMI voor het verzorgen van alle tekeningen in dit proefschrift. De medewerkers van de bibliotheek hebben een groot aantal referenties aangedragen.

Van de werkgemeenschap Meteorologie en Fysische Oceanografie dank ik met name A. Bijlsma en Mevr. I. van Esch die mij van een afstand op geruststellende wijze gesteund hebben.

Mijn collega's op het KNMI en op het Mathematisch Instituut hebben de promotietijd tot een aangename periode gemaakt. Van hen wil ik hier noemen: Marc Allaart, Sylvia Barlag, Barteld Braaksma, Cisco de Bruijn, Hans Cuijpers, Arjen Doelman, Paul de Groot, Rein Haarsma, Kees Kok, Maarten Krol, Wim de Rooy, Peter Siegmund, Liu Qing en Xueli Wang.

Mijn verste dank gaat naar Elena Kartashova. Toen de druk allang van de ketel was heeft zij voor afleiding gezorgd.