

1108201, 2410

Predictability of Weather Regime Transitions

Jeroen Oortwijn

Aan mijn ouders

111 953970

Promotor: dr.ir. J. Grasman,
hoogleraar in de zuivere en toegepaste wiskunde
Copromotor: dr. J. Barkmeijer,
wetenschappelijk onderzoeker aan het European Centre of Medium-Range
Weather Forecasts



This research was supported by the
Netherlands Geosciences Foundation (GOA)
with financial aid from the Netherlands
Organization for Scientific Research (NWO).

Predictability of Weather Regime Transitions

Jeroen Oortwijn

PROEFSCHRIFT

ter verkrijging van de graad van doctor
op gezag van de rector magnificus
van de Landbouwuniversiteit Wageningen,
dr. C.M. Karssen,
in het openbaar te verdedigen
op woensdag 25 maart 1998
des namiddags te vier uur in de Aula

Cover illustration by Corina Bulk

ISBN 90-548-5868-0

BIBLIOTHEEK
LANDBOUWUNIVERSITEIT
WAGENINGEN

Stellingen behorende bij het proefschrift

Predictability of Weather Regime Transitions

1. Atmosferische circulaties die gevoelig zijn voor veranderingen in de sterkte van een blokkade of een zonale stroming, worden vaak gekenmerkt door een krachtige straalstroom ten westen van een diffluente stroming.
2. Voor bijna elke atmosferische circulatie kan een kleine storing gevonden worden, zodanig dat een regime transitie binnen een week zal plaatsvinden.
3. Door niet-lineaire mechanismen is het ontstaan van een blokkade gemiddeld gevoeliger dan het ontstaan van een sterk zonale stroming.
4. Bij het maken van ensemble verwachtingen met behulp van zogenaamde 'singular vectors' dient men in te zien dat lineair langzaam groeiende singular vectors een niet verwaarloosbare bijdrage aan de foutengroei kunnen leveren.
5. Eén vlinder maakt nog geen zomer.
(Een snelle overgang naar een blokkade regime kan, in theorie, het meest effectief op gang gebracht worden door niet op één maar op vele plaatsen in de beginconditie kleine storingen aan te brengen.)
6. Omdat veel mensen geneigd zijn een lineair verband te leggen tussen de oorzaak en het gevolg van een gebeurtenis, dient op middelbare scholen meer aandacht besteed te worden aan niet-lineaire chaotische systemen.
7. Een betweter is pas echt irritant als hij altijd gelijk heeft.
8. Het is vaak een kostbare misvatting dat verantwoordelijkheid overdraagbaar is.

-
9. Juist door relschoppers als groep verantwoordelijk te stellen voor hun daden komt de individuele verantwoordelijkheid beter tot zijn recht.
 10. De 'bestens' order gebruikt voor het aan- of verkopen van effecten zonder koerslimiet geeft ten onrechte de indruk dat het plaatsen van zulke orders een goede keuze is. Een naamswijziging tot 'slechtstens' order zou het gebruik van deze order sterk reduceren en zodoende de particuliere belegger veel geld besparen.
 11. Voor sommige mensen is onrecht een voorrecht, omdat het hun bestaansrecht geeft.
 12. De economiestand op een douchekraan is een goed voorbeeld van een te ver doorge-draaide zuinigheid van de Nederlander.

Jeroen Oortwijn

Wageningen, 25 maart 1998

Contents

1	Introduction	3
1.1	Weather regimes	3
1.2	Error growth	6
1.3	Error growth and regimes in the Lorenz model	10
1.3.1	Lorenz three-variable model	10
1.3.2	Ensemble forecasting	11
1.3.3	Optimal regime transitions	14
1.4	The T21QG model	16
1.5	Overview of this thesis	17
2	Perturbations that optimally trigger weather regimes	19
2.1	Introduction	19
2.2	Characterization of weather regimes	20
2.3	Perturbations triggering a regime onset	22
2.4	Results for the quasi-linear range	24
2.5	Extension to the nonlinear range	27
2.6	Relation to singular vectors	32
2.7	Concluding remarks	37
3	Onset of blocking and strong zonal flow regimes	41
3.1	Introduction	42
3.2	Blocking and strong zonal flow regimes	43
3.2.1	Model performance	43
3.2.2	Characterization of the regimes	44
3.2.3	Inducing transitions	45
3.3	Identification of sensitivity patterns	47
3.3.1	Euro-Atlantic region, short range	47
3.3.2	Pacific region, short range	50
3.3.3	Euro-Atlantic region, medium range	52
3.4	Features of high-sensitivity and low-sensitivity flows	56
3.4.1	Barotropic versus baroclinic mechanisms	58
3.4.2	Precursors of transitions	64
3.5	Summary and discussion	67

4 Growth properties of optimal transition perturbations	71
4.1 Introduction	71
4.2 The evolution of an optimal transition perturbation	73
4.2.1 General remarks	73
4.2.2 A representative example of a sensitive period	74
4.3 A WKB approximation	82
4.4 Numerical simulations	86
4.4.1 Formulation of the WKB model and the barotropic model	86
4.4.2 Zonal mean background flow	87
4.4.3 Non-zonal background flows	90
4.5 Summary and discussion	94
A Vertical modes of the T21QG model	96
B Derivation of the KE-equation	100
References	103
Samenvatting	109
Dankwoord	115
Curriculum vitae	116

Chapter 1

Introduction

This thesis is about the predictability of weather regime transitions. In this introductory chapter the theory of weather regimes is described. Further, the concept of predictability and its limitations due to error growth in numerical weather prediction models are explained. To illustrate all this, a relatively simple example is considered. The quasigeostrophic model is introduced, which is used for numerical experiments in the subsequent chapters, and finally, an overview of the thesis is given.

1.1 Weather regimes

The extratropical atmospheric evolution exhibits variability at various time scales: high-frequency variability (2-6 days), low-frequency variability (10-90 days) and seasonal variability. High-frequency variability is associated with the development of synoptic-scale baroclinic weather systems. Predictability of the atmospheric circulation is directly related to the development of these weather systems. The useful range of deterministic forecasts is limited by the life time of these systems, which is of the order of 5 days. The birth of new high-frequency transients, like depressions, occurs at very small scales that are not fully captured by the numerical weather prediction (NWP) models. Within a couple of days they affect the atmospheric circulation on much larger scales. It is obvious that when this process is not properly resolved in the initial conditions it will be difficult to forecast it correctly. On the other hand, the evolution of systems that are already present in the initial conditions can be forecasted reasonably well. Thus, current deterministic forecasts are useful, on average, up to about 6 days. Low-frequency variability is manifested at time scales in the order of 10 to 90 days. Dole and Gordon (1983) showed that low-frequency variability even explains the largest nonseasonal variance. The predictability is not strongly affected by the presence of low-frequency phenomena.

The geographical distributions of both high-frequency and low-frequency variability are inhomogeneous. The atmospheric circulation of the Northern Hemisphere is characterized by a localized strong westerly flow, the so-called jet stream, which exists due to the rotation of the earth and the temperature difference between the equator and north pole. Due to orography and land-sea contrasts the jet stream has mainly two maxima to the east of the Himalayas and the Rocky Mountains. Large velocity gradients, differences in the heat capacity between land and sea and differences in the moisture distribution in the air over

land and sea give rise to the existence of storm tracks at the north-western parts of the Atlantic and Pacific oceans. These regions mark the maxima of high-frequency variability. The maxima of low-frequency variability are located at the end of the storm tracks at the north-eastern parts of the oceans.

The patterns of low-frequency variability have also a preferred structure. They are characterized by recurrent large-scale anomaly patterns. One of the best known weather phenomena that may persist for weeks is the so-called blocking event. Blocked flows are characterized by a high-amplitude large-scale anomaly pattern consisting of a high pressure cell to the north of a low pressure cell. This pattern blocks the normally westerly flow and splits the jet into a northern and southern branch. Because the high-frequency transients are forced to bend northward or southward this may result in a long period of quiet weather. The local weather varies only little during a blocking episode, sometimes causing extremely dry and warm periods in summer or dry and cold periods in winter. To the north and south of the blocking the weather may be (extremely) wet and stormy. Blocking episodes do not show a preferred time scale. Onset and breakdown of blockings are usually very rapid. The predictability of blocking maintenance is on average good, whereas the predictability of the onset is still a challenge for numerical models (see e.g. Tibaldi et al. 1994).

To explain the existence of large-scale low-frequency phenomena Charney and DeVore (1979) employed the concept of multiple equilibria. In their study they assumed that the atmospheric circulation possesses several equilibrium solutions. A stationary equilibrium is maintained by the nonlinear interaction of the external (nonzonal) forcing and the zonal flow. Driven by small-scale instabilities, the atmosphere moves from one quasi-stable equilibrium to another. A blocking event thus corresponds to a situation in which the flow approaches a quasi-stable equilibrium. However, all their solutions (which were obtained in a low-order model) were unstable to synoptic-scale disturbances and will not be maintained in realistic flows.

This concept has therefore been modified by Reinhold and Pierrehumbert (1982) who introduced the term *weather regime*. They hypothesized that the low-frequency variability of the atmospheric circulation is a consequence of transitions between several weather regimes. A weather regime is the result of the quasi-equilibrium between the large-scale flow and the synoptic-scale weather systems. The large scales not only organize the evolution of the synoptic scales into storm tracks, vice versa, the organized storm tracks determine the structure of the large-scale flow and balance the tendency of the large scales. Weather regimes are quasi-stationary in the large scales only, due to the systematic feedback of the high-frequency transients. The reason why low-frequency patterns are not predictable for periods much longer than the synoptic-time scale is probably because they are coupled to the high-frequency transitions to maintain their structure.

Two possible configurations of the balance between the large-scale flow and the synoptic-scale flow may occur. In the first configuration the large-scale pattern is not related to a steady state of the large-scale flow, and can only be quasi-stationary by the forcing of high-frequency transients. In other words, the large-scale pattern must be successful in organizing the transients in a self-sustaining pattern otherwise it will not be persistent. The second possibility is that the large-scale pattern approaches a steady state of the large-scale equations of motion. In this hypothesis the small tendency of this large-scale flow pattern in

the full system must therefore be balanced by the high-frequency transients. Both described configurations are supported by several studies. Vautard et al. (1988) and Vautard and Legras (1988) showed that blocking flows, which are not stationary states of the large-scale equations, can be maintained against dissipation by the high-frequency transients. Haines and Marshall (1987) found a similar result but for modonlike blocking flows, which are steady states of the simplified large-scale equations.

In the weather regime theory, onset and breakdown of weather regimes can be explained by the evolution of a single disturbance that triggers or breaks down the quasi-equilibrium. Thus, transitions may occur very rapidly and almost at random. This is in agreement with the observational results that onset and breakdown of persistent large-scale anomalies (e.g. blockings) are rather abrupt (on the synoptic-time scale) compared to the duration of the event (Dole and Gordon 1983; Hansen and Sutra 1986; Yang and Reinhold 1991). Furthermore, no preferred time scale is found (Dole and Gordon 1983; Legras and Ghil 1985; Liu 1994). Transitions between weather regimes may thus be purely due to the internal dynamics of the atmospheric flow. This view is in contrast with the ideas of, for instance, Hoskins and Karoly (1981) who described the onset of certain persistent anomalies in terms of propagation of wave trains into the midlatitudes produced by tropical sea surface temperature anomalies. Such an external slowly-varying forcing can not account for the abrupt changes in the flow structure. However, external forcing could favour the occurrence of certain regimes within a season (see e.g. Palmer 1993). In addition to quasi-stationary weather regimes, low-frequency oscillations are observed (e.g. Anderson and Rosen 1983; Branstator 1987; Kushnir 1987). These low-frequency oscillations may favour the occurrence of particular weather regimes and influence transitions between regimes (Plaut and Vautard 1994; Michelangeli and Vautard 1998).

The regime-like behaviour of the atmospheric circulation may be described using the theory of dynamical systems. For forced-dissipative dynamical systems the set of possible states, called the attractor of the system, is believed to be much smaller (even of probability zero in phase space) than the phase space of the whole system (which is infinitely dimensional). An atmospheric state is given by a point in phase space and its evolution by a trajectory. The states are expected to propagate aperiodically (chaotically) but not randomly. Regimes may be recognized by regions of phase space that are densely populated. Atmospheric states will spend more time in these regions and will also visit them more often. Transitions between regimes are given by trajectories that connect the regimes. Some regimes may be connected by many trajectories, others by few trajectories. The first represent preferred transitions between two weather regimes. In section 1.3 we will demonstrate that this behaviour even occurs in low-dimensional systems.

In this thesis, we will focus on the predictability of transitions between a blocking regime and a strong zonal flow regime. This study will be performed within the context of a three-level quasigeostrophic model triangularly truncated at wavenumber 21 (T21QG). Its tangent linear and adjoint models are used to perform sensitivity experiments. In the next section we describe error growth in numerical models and explain how tangent linear and adjoint models may be used to compute maximum error growth. In section 1.3 we apply these methods for a relatively simple example. In section 1.4 the T21QG model is briefly described. Finally, an overview of this thesis is given in section 1.5.

1.2 Error growth

Errors in the estimation of the initial conditions and errors in the formulation of numerical models give rise to error growth. Given the initial state of the atmosphere, future states are computed by integrating complex numerical general circulation models (GCMs). Although the atmosphere is a deterministic system and its laws of motion are known, some fundamental difficulties arise in the predictability of its future states. The atmospheric circulation is nonlinear, mainly due to the presence of non-linear advection, and appears to be fundamentally chaotic. An important property of a chaotic system is that it shows sensitivity to initial conditions. This means that a small error in the initial conditions will eventually lead to a large forecast error. In an infinitely-dimensional system like the atmosphere, infinitesimal errors may evolve into large errors even within a finite time. This has resulted in the estimation of a predictability horizon of about one or two weeks. Behind this horizon a single deterministic forecast is useless. Due to errors in the initial state and imperfections in the model formulation current GCMs are useful until forecast day 6, i.e., they have a higher skill than forecasts based on climate or persistence. *

The atmospheric state at a certain time may be represented by a vector ψ in phase space. The components of ψ contain all essential information, namely, vorticity, divergence, temperature, and relative humidity at all points in the atmosphere and pressure at the surface. In principle the phase space is infinitely dimensional. In practice, the vector ψ is computed at a grid so that the phase space of current GCMs is about 5×10^6 .

Suppose that the atmospheric evolution is given by the following equation

$$\frac{\partial \psi}{\partial t} = \mathbf{F}(\psi), \quad (1.1)$$

where \mathbf{F} is a nonlinear evolution operator. A forecast can be made by integrating equation (1.1). The atmospheric circulation is represented in phase space by an orbit $\psi(t)$. Orbits starting at slightly different initial states will diverge sooner or later. Since the exact initial state of the atmosphere is unknown we will have to take into account error growth. During the first 2-3 days error growth appears to be dominated by linear terms. It is therefore useful to linearise equation (1.1). This results in the evolution equation for a small error ε ,

$$\frac{\partial \varepsilon}{\partial t} = \mathbf{F}'(t)\varepsilon. \quad (1.2)$$

This equation is called the tangent linear equation. The operator $\mathbf{F}'(t)$ is obtained by differentiation of \mathbf{F} along the orbit of $\psi(t)$ in phase space. Since equation (1.2) is linear, we may relate the error at forecast time T to the initial error

$$\varepsilon(T) = \mathbf{R}(0, T)\varepsilon(0), \quad (1.3)$$

where $\mathbf{R}(0, T)$ is the linear evolution operator.

*Forecast skill is measured in terms of verification scores like the root mean square error (RMSE) and anomaly correlation coefficient (ACC).

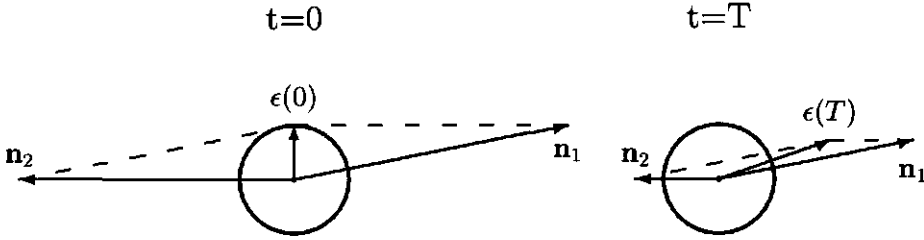


Figure 1.1: Initial error $\varepsilon(0)$ is sum of eigenvectors \mathbf{n}_1 and \mathbf{n}_2 . Large initial contributions of the normal modes are 'hidden' in $\varepsilon(0)$ due to their opposite working. At forecast time T the error has grown because of the unequal decay of \mathbf{n}_1 and \mathbf{n}_2 .

In order to demonstrate that error growth may be fast or even explosive we shall assume that $\mathbf{F}'(t)$ is independent of time. Now, the error at forecast time T is found by integration of equation (1.2),

$$\varepsilon(T) = e^{\mathbf{F}'T} \varepsilon(0). \quad (1.4)$$

In a classical stability analysis the eigenvalues and the eigenvectors of \mathbf{F}' are calculated. The eigenvalues of $e^{\mathbf{F}'T}$ are equal to the exponent of the eigenvalues of \mathbf{F}' multiplied by the time T . The eigenvectors, which are equal to these of $e^{\mathbf{F}'T}$, are usually called normal modes. Their modal structure is constant in time. It will only amplify or decay exponentially and oscillate periodically in time (when the eigenvalue is complex). An initial error can be written as a sum of these normal modes. A solution is then said to be asymptotically stable when all normal modes decay exponentially. However, even in this case (temporary) explosive error growth may occur.

We illustrate this with a simple example, where $\mathbf{R} = e^{\mathbf{F}'T}$ has the following matrix structure

$$\mathbf{R} = \begin{pmatrix} 0.3 & 2 \\ 0 & 0.7 \end{pmatrix}. \quad (1.5)$$

The eigenvectors of \mathbf{R} (equal to the normal modes of \mathbf{F}') are equal to $\mathbf{n}_1 = (0.981, 0.196)$ with eigenvalue $\lambda_1 = 0.7$ and $\mathbf{n}_2 = (1, 0)$ with eigenvalue $\lambda_2 = 0.3$. Hence, at forecast time T the normal modes have shrunk to values of 0.7 and 0.3 of their initial lengths. Both normal modes decay and the solution is asymptotically stable. However, when we take an initial error vector $\varepsilon(0) = (\varepsilon_1(0), \varepsilon_2(0)) = (0, 1)$, its growth at time T is equal to a factor 2.12, $\varepsilon(T) = (2, 0.7)$. The situations at initial and final time are shown in Fig. 1.1.

This error growth is possible because the normal modes are not orthogonal. As a result, a small error can have relatively large components with respect to the basis of normal modes. The unequal decay of the normal modes results in the temporary growth of the error. [†] That this linear mechanism is important for error growth was already realized by Orr (1907). The

[†]In addition to growth due to the unequal decay of the normal modes another process may contribute to temporary growth when the dimension of the system is larger. A real matrix operator with at least rank 4 can have different complex eigenvector pairs. Each eigenvector pair will span a two-dimensional subspace in

amplification of errors may be much larger, certainly for systems with more dimensions.[†] When errors are amplified sufficiently, nonlinear terms become important. In this case, the linear (asymptotic) stability analysis has become irrelevant.

It is evident that the calculation of normal modes is not the best choice for estimating the maximum possible error growth. Maximum growth may be found by maximizing the length $\|\varepsilon(T)\|$ of $\varepsilon(T)$. In order to measure lengths we have to equip the vector space with a norm. Various norms may be defined depending on their application. In fact, we already used a norm induced by the Euclidean inner product $\langle \cdot, \cdot \rangle$:

$$\|\varepsilon\|^2 = \langle \varepsilon, \varepsilon \rangle = \sum_{i=1}^2 \varepsilon_i^2. \quad (1.6)$$

Now we can write

$$\|\varepsilon(T)\|^2 = \langle \varepsilon(T), \varepsilon(T) \rangle = \langle \mathbf{R}\varepsilon(0), \mathbf{R}\varepsilon(0) \rangle = \langle \mathbf{R}^* \mathbf{R} \varepsilon(0), \varepsilon(0) \rangle. \quad (1.7)$$

Here, we introduced the operator \mathbf{R}^* , which is the adjoint of \mathbf{R} with respect to the Euclidean inner product. For realistic atmospheric flows the linear evolution operator \mathbf{R} is always non-normal, i.e., the operator does not commute with its adjoint. This is due to the always present vertical and horizontal shear (Farrell and Ioannou 1996). Hence, the set of eigenvectors is not orthogonal. The eigenvector of \mathbf{R} with the largest real part of the eigenvalue will therefore not bound the maximum error growth, as we have seen above. However, the product operator $\mathbf{R}^* \mathbf{R}$ is normal and its eigenvectors, so-called singular vectors, do form an orthogonal basis with respect to the Euclidean inner product. Hence, the dominant eigenvalue, or singular value, of $\mathbf{R}^* \mathbf{R}$ gives an upperbound for the amplification of errors at forecast time T .

In our example the adjoint of \mathbf{R} is just its transpose. Hence,

$$\mathbf{R}^* \mathbf{R} = \mathbf{R}^T \mathbf{R} = \begin{pmatrix} 0.09 & 0.6 \\ 0.6 & 4.49 \end{pmatrix}. \quad (1.8)$$

Its singular vectors are $\mathbf{v}_1 = (0.133, 0.991)$ with singular value $\mu_1 = 4.57$ and $\mathbf{v}_2 = (0.991, -0.133)$ with singular value $\mu_2 = 0.0096$. So, even though all normal modes decay the maximum amplification is a factor of $\sqrt{4.57} = 2.14$. The largest decay is determined by the smallest singular value, a factor of $\sqrt{0.0096} = 0.098$. The initial and evolved singular vectors are shown in Fig. 1.2.

The singular vectors form an orthogonal basis. Hence, each perturbation can be written as a sum of singular vectors. When the amplification properties of all singular vectors are

the real phase space. When the imaginary parts of the eigenvalues are different, corresponding to different oscillation periods, the eigenvector pairs will rotate with different velocities in the real phase space. This may lead to error growth. For example, the eigenvectors may be first orthogonal and at a later time be partially aligned in the same direction, such that the sum of the components of an error along these eigenvectors will increase.

[†]Trefethen et al. (1993) show that temporary growth may amplify errors by factors of the order of 10^5 . They use this linear mechanism as an explanation for the onset of turbulence for Couette and Poiseuille flow at subcritical Reynolds numbers (for which all normal modes decay).

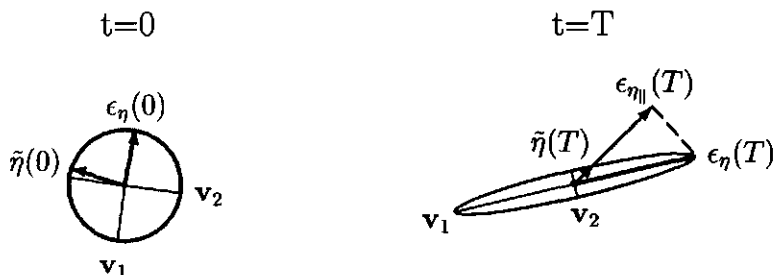


Figure 1.2: The error circle evolves into an ellipsoid with principal axes given by the singular vectors \mathbf{v}_1 and \mathbf{v}_2 . The error vector $\epsilon_\eta(0)$ is almost equal to \mathbf{v}_1 . At forecast time it has evolved into the vector $\epsilon_\eta(T)$, which has the largest possible projection onto the vector $\eta = (1, 1)$, given by $\epsilon_{\eta\parallel}(T)$. The vector $\tilde{\eta}(0)$ evolves into $\tilde{\eta}(T)$, which is exactly aligned along η . This vector is not optimal because it projects onto decaying directions of phase space.

known maximum error growth in any specific direction can thus be computed. However, the computation of singular vectors is rather expensive. When only maximum error growth in a specific direction is required, a more efficient method may be found. Such an approach is useful to determine large error growth into the direction of blocking and strong zonal flow regimes. Suppose that we want to maximize error growth at forecast time T into the direction of phase space given by the vector η (which represents, for instance, a blocking anomaly pattern), or, equivalently to maximize the inner product between $\epsilon(T)$ and η . Using the tangent linear and adjoint operators, we can write

$$\langle \epsilon(T), \eta \rangle = \langle \mathbf{R}\epsilon(0), \eta \rangle = \langle \epsilon(0), \mathbf{R}^*\eta \rangle. \quad (1.9)$$

We conclude that the error vector $\epsilon_\eta(0)$ of unit norm that maximizes the inner product is equal to

$$\epsilon_\eta(0) = \frac{\mathbf{R}^*\eta}{\|\mathbf{R}^*\eta\|}. \quad (1.10)$$

In the following chapters this method is used to trigger regime transitions. To illustrate it in a simple context we return to the two-dimensional example given by Eq. (1.5). Suppose that we are interested in error growth in the direction of $\eta = (1, 1)$. We then find that

$$\epsilon_\eta(0) = (0.110, 0.994) \quad \text{and} \quad \epsilon_\eta(T) = (2.02, 0.696). \quad (1.11)$$

These vectors are almost equal to the fastest growing singular vector \mathbf{v}_1 at initial and optimization time, respectively. The maximum amplification in the direction of η is given by

$$\frac{\langle \epsilon_\eta(T), \eta \rangle}{\|\eta\|} = 1.92. \quad (1.12)$$

Notice that the optimal error vector $\epsilon_\eta(T)$ points into another direction of phase space than η . However, its projection onto η is largest at forecast time T . In Fig. 1.2 the projection onto η is denoted by $\epsilon_{\eta\parallel}(T) = (1.36, 1.36)$.

The initial perturbation $\tilde{\eta}(0)$ of unit length which at optimization time is exactly aligned along η is found by the backward integration of the linear model starting with η ; that is, $\tilde{\eta}(0) = \frac{\mathbf{R}^{-1}\eta}{\|\mathbf{R}^{-1}\eta\|} = (-0.974, 0.225)$. The evolved vector $\tilde{\eta}(T) = (0.157, 0.157)$ is clearly smaller than the projection $\varepsilon_{\eta}(T)$ of the optimal vector $\varepsilon_{\eta}(T)$ onto η at time T .

1.3 Error growth and regimes in the Lorenz model

In this section we will describe a simple dynamical system used by Lorenz (1963) to study Rayleigh-Bénard convection. The Lorenz model has only three variables and is one of the most simple models that exhibits chaos and sensitive dependence to initial conditions. We will use the Lorenz model as a metaphor for the atmospheric circulation and illustrate that error growth depends strongly on the value of the initial conditions. Furthermore, it is shown that regimes can be triggered with the method described in the previous section.

1.3.1 Lorenz three-variable model

The Lorenz model is given by the following set of nonlinear differential equations

$$\begin{aligned}\dot{x} &= -\sigma x + \sigma y, \\ \dot{y} &= -xz + rx - y, \\ \dot{z} &= xy - bz.\end{aligned}\tag{1.13}$$

Here x, y, z are the three variables and σ, r and b are parameters. The parameters are set to the original values $\sigma = 10$, $r = 28$ and $b = 8/3$. These equations are derived by Lorenz in order to represent Rayleigh-Bénard convection to a rough approximation. It describes the motion of a fluid between two horizontal plates with a temperature difference due to the heating of the lower one. An unstable situation arises: the warm and lighter fluid at the bottom rises and sets into motion the colder fluid at the top. The equations describe the three dominant modes of the system. There are three fixed points: the first represents a state of rest (no convection at all), the other two represent circulation cells, which rotate clockwise or anticlockwise. For the above given values of the parameters all three solutions are unstable.

A fourth-order Runge Kutta integration scheme is used to integrate the Lorenz model with a nondimensional time step of 0.01. The solutions of equations (1.13) appear to be aperiodic and chaotic. Due to dissipation, trajectories in the three-dimensional phase space are attracted to the structure shown in Fig. 1.3. Since the system shows extreme sensitivity to initial conditions this attractor is called a strange attractor (see, e.g., Schuster 1989). The attractor resembles a butterfly with two wings on which trajectories circle around.[§] The centers of the two wings represent the two unstable steady convections. The motion of the fluid switches in a chaotic manner between the two convection regimes. The number

[§]The Hausdorff dimension is a noninteger (fractal) and equal to 2.06. Locally the attractor appears to be approximately two-dimensional. All points on the attractor lie very close to a two-dimensional plane tangent to the attractor. On a micro-scale it appears that the attractor consists of an infinite set of sheets that are densely packed together.

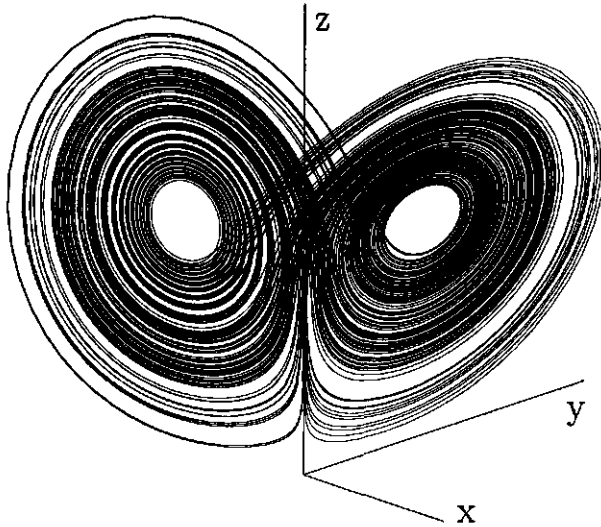


Figure 1.3: The Lorenz attractor in the three-dimensional phase space.

of circulations on a wing between two regime transitions change in a random manner. Two characteristic time scales can be recognized: a slow one representing the residence time within a regime and a fast one representing the transition time scale.

The Lorenz model may be considered as a simple model for the atmospheric circulation that oscillates between two regimes, e.g., a blocking and a strong zonal flow regime. The fast transition time scale in the model can be associated with the synoptic-time scale in the atmosphere (about 5 days). As discussed in section 1.1, regime transitions are often rapid and on a time scale of the order of a few days. The slower time scale represents the residence time within a regime (of the order of a few weeks).

1.3.2 Ensemble forecasting

Because we do not know the exact state of the atmosphere, we always have an uncertainty in the initial conditions. This uncertainty can be represented in terms of a probability density function. Given the uncertainty in the initial conditions represented by the probability density function, an optimal forecast is one for which the complete evolution of the probability density function can be computed. Forecasting has therefore become an inherently probabilistic problem (Epstein 1969).

Let us return to the Lorenz model and look at some ensemble forecasts for various initial conditions (cf. Palmer 1993). We assume that each component of the three-dimensional state vector \mathbf{x} has a normal distribution with a uniform standard deviation σ around the initial condition. The probability density function thus forms initially a three-dimensional sphere in phase space with radius σ . The ensembles may be interpreted as the three-dimensional

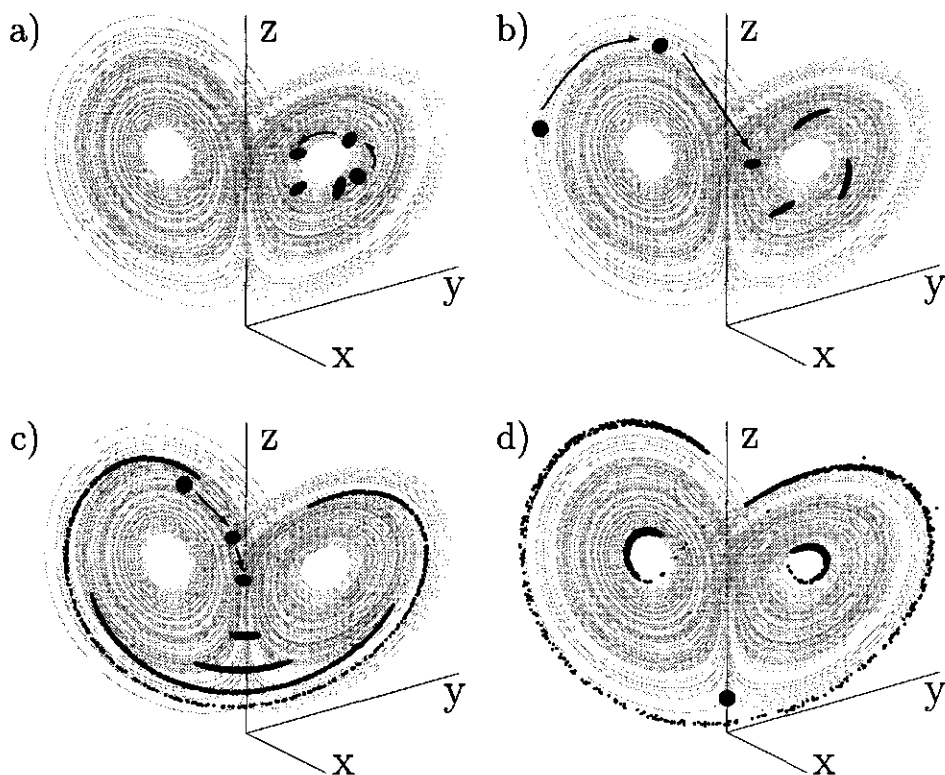


Figure 1.4: The evolution of ensemble forecasts starting at four different initial conditions. Each ensemble consists of 1000 members initially distributed within a sphere centered around a reference point. In the first three panels the evolution stages are plotted at intervals of 16 time steps. In panel (d) the time interval is 64 time steps.

probability density functions. Each ensemble consist of 1000 points, initially distributed within a three-dimensional sphere centered around a reference point. The evolution of the ensemble is performed by integrating each point with the nonlinear Lorenz equations (1.13). The evolutions are shown in Fig. 1.4. In each panel several stages are plotted. In Fig. 1.4a the complete ensemble remains in one regime. The sensitivity to the initial conditions is small in this part of phase space and so the evolution is very predictable. In Fig. 1.4b the complete ensemble moves from one regime to the other. The regime transition is very predictable. Figure 1.4c shows a sensitive evolution. The probability that a regime transition will occur is about 50 %. During the first four stages the sphere evolves into an ellipsoid. This is the linear stage of the evolution. The principal axes of the ellipsoid can be obtained by computing singular vectors, as described in the previous section. In the next stage the ellipsoid is slightly deformed by nonlinear processes, the weakly nonlinear stage. The singular vectors still give

a reasonable approximation of the fastest growing directions. After this stage the evolution of errors is strongly nonlinear. Linear calculations of maximum error growth fail. Total loss of predictability occurs when the ensemble is dispersed over the whole attractor (not shown). By that time, the probability density function can not be distinguished anymore from the climatological distribution. In Fig. 1.4d the ensemble is shown at three stages of the evolution. The linear stage is much shorter than in the former examples. At the final stage the probability density function has a sharp bimodal distribution. The regime type is unpredictable. However, it is still predictable that the system will remain in one of the regimes for a very long time. One speaks of a return of skill. This convergence of the ensemble is due to phase-space regions in which all (infinitesimal) errors shrink with time.

The examples illustrate that the predictability depends considerably on the initial conditions. This leads to a variability in the skill of various forecasts. By making use of ensemble forecasts the skill may be estimated a priori. Knowledge of the skill could enlarge the useful range of forecasts. How to compute forecast skill of operational forecast models is one of the leading research themes of numerical weather prediction. Current operational forecast models have about $N = 10^6$ components. The N -dimensional probability density function will evolve in the same way as illustrated in the Lorenz model. The duration of the linear phase is estimated at about 2-3 days, the duration of the weakly nonlinear phase at about 5-7 days. Because of the many degrees of freedom and the high costs of integrating the operational forecast models, it is impossible to compute the complete evolution of the probability density function. Therefore, several techniques are applied to estimate the forecast skill.

Here, we will describe briefly three approaches, which differ mainly in how initial perturbations are selected to create an ensemble. The first technique is Monte-Carlo forecasting: ensemble members are randomly selected from the initial probability density function (cf. Leith 1974). Due to the high numerical costs only a small ensemble can be integrated, which will sample only a small subspace of the N -dimensional phase space. This has been the reason to develop other techniques that dynamically constrain the selection of perturbations to create an ensemble. One of them is applied in the ensemble prediction system (EPS) at the European Centre for Medium-Range Weather Forecasts (ECMWF; Molteni et al. 1996). In this approach an ensemble is created by computing the fastest growing perturbations that will stretch the probability density function along the directions of these dominant linear instabilities. In the current ensemble prediction system singular vectors are computed (see section 1.2) for a simplified version of the operational model so that 50 alternative forecasts can be made. Based on this ensemble probabilities of various scenarios are derived. The third technique is the so-called 'breeding' method used at the National Centers for Environmental Prediction (NCEP; Toth and Kalnay 1993, 1997). A small ensemble of perturbations is created (currently about 17 members are used). The perturbations represent directions in which growth has been dominant over the past. The idea is that these 'bred vectors' represent directions in which the analysis uncertainty is large. The bred vectors are obtained by performing nonlinear perturbed integrations. After each integration period the bred vector is scaled to its initial amplitude. After many integration periods the direction in which error growth has been large will remain.

It is not clear yet which approach is more successful. Studies of Houtekamer and Derome (1995) with a T21QG model and Anderson (1997) with the Lorenz model indicate that

the use of constrained ensembles may have no impact on the ensemble mean. The forecast of higher moment quantities such as forecast spread may even be worse (Anderson 1997). Although the importance of ensemble forecasting is evident, much theoretical and practical work has to be done.

1.3.3 Optimal regime transitions

Ensemble forecast may be especially useful in situations when the atmosphere is about to make a regime transition (Palmer 1993; Buizza and Palmer 1995). As we have illustrated for the Lorenz model, some regime transitions are predictable, others are unpredictable. In the next chapters, we will show that the predictability of the circulation in the T21QG model also varies considerably. We will extensively consider initial conditions sensitive and insensitive with respect to regime transitions.

In section 1.2 we have already explained how we can study optimal regime transitions. We first illustrate the procedure for the Lorenz model. In order to maximize a regime transition we will use the following property of the Lorenz system. The divergence of trajectories is largest along the z -axis near the unstable fixed point $(0,0,0)$, see Fig. 1.4d. Trajectories that pass this point closely will move to the center of one of the regimes. We will therefore try to minimize the z -value of the perturbed orbit at forecast time T in order to induce a transition later.

Let us first rewrite the Lorenz model in the form $\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x})$. The linearized equations are then given by $\dot{\varepsilon} = \mathbf{F}'(t)\varepsilon$. The operator $\mathbf{F}'(t)$ is found by differentiation of \mathbf{F} along the reference orbit,

$$\mathbf{F}'(t) = \begin{pmatrix} -\sigma & \sigma & 0 \\ r - z(t) & -1 & -x(t) \\ y(t) & x(t) & -b \end{pmatrix}. \quad (1.14)$$

In order to find the optimal transition vector we maximize the inner product

$$\langle \varepsilon(T), -\mathbf{e}_z \rangle, \quad (1.15)$$

where $\mathbf{e}_z = (0, 0, 1)$ is the unit vector in the z -direction. As in equations (1.9) and (1.10) the optimal initial transition vector is, apart from a scaling, given by

$$\varepsilon_{tr}(0) = -\mathbf{R}^* \mathbf{e}_z. \quad (1.16)$$

The adjoint operator \mathbf{R}^* is equal to the transpose \mathbf{R}^T of \mathbf{R} , with \mathbf{R} formally given by

$$\mathbf{R} = e^{\int_0^T \mathbf{F}'(t) dt}. \quad (1.17)$$

We will demonstrate how regimes are triggered with two examples. In both examples the sensitivity is very large. The optimization time T is equal to 100 time steps. In Fig. 1.5a a reference orbit and a perturbed orbit are shown. The reference flow makes first one loop in the left regime. Next, it moves to the right regime where it will remain a long period. At optimization time T the reference flow is located in the upper left corner of the left wing. The

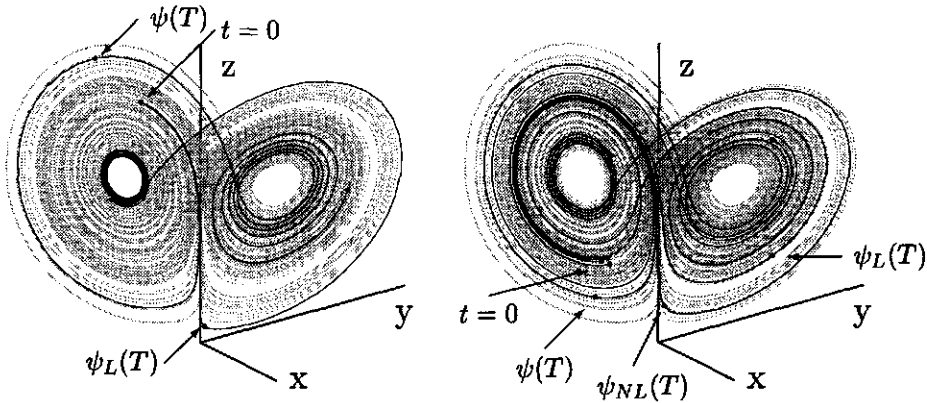


Figure 1.5: Two examples of evolutions that are extremely sensitive to the initial conditions. The reference flows are given by $\psi(t)$. In (a) the trajectory ψ_L starting at a point found by adding the linearly optimal perturbation to the reference initial condition triggers a regime for a very long period, while the reference flow remains in the other regime. In (b) the linear optimal evolution ψ_L does not trigger a regime for a long period. The evolution ψ_{NL} , obtained by iteratively modifying the optimal perturbation, is more successful. The locations at initial, optimization and final time are marked by dots.

linearly optimal perturbation is computed according to equation (1.16). The perturbation is scaled such that it lies on a sphere centered around the reference initial condition with a radius equal to 0.1. At optimization time the perturbed orbit is located near the z -axis close to the origin. After that, the perturbed flow makes a transition to the left regime, where it will remain for a very long time. A regime is successfully triggered.

After the optimization time T the perturbed flow moves in a large loop around the right wing in a region of phase space that is rarely visited. This indicates that the perturbation is rather special. We will return to this characteristic of optimal perturbations in the T21QG model (see section 5 of chapter 2). The advantage of the Lorenz model is that it has only three variables, so that we are able to study the optimal perturbation in more detail (cf. Anderson 1994, 1997). A first remark is that the perturbation is constructed to induce extreme events and must be special in this sense. Secondly, it appears that the initial optimal perturbation is not located exactly on the attractor. [¶] The initial route of the

[¶]Very long integrations with the Lorenz model can be performed. It is even feasible to compute analogs that are very close to the initial condition. These analogs are on the attractor and are all located in a nearly two-dimensional plane. The initial optimal perturbation is located just outside this plane. By selecting the analog that at optimization time is closest to the optimally perturbed forecast, a comparison can be made between these two trajectories. Each time step, the local tangent plane to the attractor has been computed at the location of this analog. It appears that the optimal perturbed trajectory is attracted into the tangent plane within about 40 time steps; that is, it is indistinguishable from it. So, although the initial route of the optimal perturbation is located outside the attractor, it seems that the further evolution lies on the attractor.

perturbed trajectory is located just outside the attractor plane. This indicates that also in more general dynamical systems optimal perturbations or singular vectors may be not on the attractor. They may therefore realize perturbation growth that can not occur for evolutions of states on the attractor.

In Fig. 1.5b a sensitive period is selected for which the linear optimal perturbation fails to predict the maximum possible deviation from the reference flow. The perturbation is scaled such that it lies on a sphere centered around the reference initial condition with a radius equal to 1. Using an iterative extension of the maximization procedure, a perturbation with the same initial amplitude is computed, which results in a much larger deviation. In chapter 2 the iterative procedure will be described in the context of the T21QG model. The trajectory of the iteratively modified initial state remains after the optimization time in the left regime for a very long period. The linearly computed optimal perturbation is not successful in inducing a regime for such a long time. Through the iterative procedure, we can account for nonlinear error growth and extend the maximization procedure into the (weakly) nonlinear stage.

1.4 The T21QG model

The extratropical circulation can be described approximately by the quasigeostrophic equations. The advantage of the quasigeostrophic equations above the full primitive equations is that they are filtered (absence of gravity waves) and can be written in terms of only one variable, the quasigeostrophic potential vorticity. In the absence of forcing and dissipation the quasigeostrophic potential vorticity is conserved following the geostrophic motion. All other variables can be found by the inversion of the quasigeostrophic potential vorticity.

In this section we will briefly describe the quasigeostrophic model developed by Molteni (Marshall and Molteni 1993). It is a three-level spectral model triangularly truncated at wave number 21. The T21QG model integrates prognostic equations for quasigeostrophic potential vorticity q at 200, 500, and 800 hPa (levels 1,2,3 respectively),

$$\frac{\partial q_i}{\partial t} = -\mathbf{v}_{\mathbf{g}i} \cdot \nabla q_i - D_i + S_i, \quad i = 1, 2, 3. \quad (1.18)$$

Here $\mathbf{v}_{\mathbf{g}}$ is the geostrophic velocity, D represents dissipation and S an artificial forcing. A leap-frog scheme is used to integrate the model with a time step of 40 minutes. At each level the quasigeostrophic potential vorticity has 483 independent spectral components, so that the model has 1449 degrees of freedom.

The quasigeostrophic potential vorticity is related to the streamfunction as

$$\begin{aligned} q_1 &= \nabla^2 \psi_1 + f - R_1^{-2}(\psi_1 - \psi_2), \\ q_2 &= \nabla^2 \psi_2 + f + R_1^{-2}(\psi_1 - \psi_2) - R_2^{-2}(\psi_2 - \psi_3), \\ q_3 &= \nabla^2 \psi_3 + f \left(1 + \frac{h}{H_0}\right) + R_2^{-2}(\psi_2 - \psi_3). \end{aligned} \quad (1.19)$$

where f is the planetary vorticity, R_1 ($= 700$ km) and R_2 ($= 450$ km) are Rossby radii of deformation, h is the orographic height and H_0 ($= 9$ km) is a scale height. The relation

between the geostrophic velocity and the streamfunction is given by

$$\mathbf{v}_{\mathbf{g}_i} = \mathbf{k} \times \nabla \psi_i, \quad i = 1, 2, 3. \quad (1.20)$$

Here \mathbf{k} is the local vertical unit vector.

The dissipation terms include Ekman friction, temperature relaxation (Newtonian) and horizontal diffusion (hyperviscosity). The forcing is constant but spatially varying. It is computed by requiring that the average potential vorticity tendencies (eq. (1.18)) of a large number of observed atmospheric states will vanish (cf. Roads 1987). The observed states that are used, are ECMWF analyses of streamfunction at the three pressure levels for each day in January and February 1984 to 1989. The procedure to compute the artificial forcing appears to be very successful. The mean state of the model turns out to be very realistic. Furthermore, the model simulates the wintertime extratropical variability remarkably well. Even statistics and characteristics of blocking and strong zonal flow regimes are well simulated. For more details about the model formulation we refer to Marshall and Molteni (1993). The model performance is also described by Liu and Opsteegh (1995) and in section 2.1 of chapter 3.

In addition to the nonlinear T21QG model, a tangent model and an adjoint model are used. The tangent model integrates linear equations (forward in time) which are obtained by linearising the prognostic equations (1.18) around a nonlinear orbit of the T21QG model. It describes the evolution of small perturbations along this nonlinear orbit. The adjoint model integrates the adjoint of the tangent linear equations backward in time. Using the adjoint model the sensitivity of a forecast aspect to all uncertainties in the initial conditions can be computed. We will use the adjoint model to compute the sensitivity of regime transitions to the initial conditions. An overview of the literature on the use of adjoint models in meteorology is given in Courtier et al. (1993).

1.5 Overview of this thesis

In this thesis we study the sensitivity to the initial conditions of weather regime transitions. We focus on transitions between two weather regimes, a Euro-Atlantic blocking regime and a Euro-Atlantic strong zonal regime within the context of a quasigeostrophic model. The goal of this thesis is to investigate which characteristics of a flow determine the sensitivity with respect to these regime transitions and which mechanisms contribute to the growth of the optimal transition perturbations. The thesis consists of four chapters.

In chapter 2 we explain the method to construct perturbations that optimally trigger the onset of a blocking or strong zonal flow regime. These regimes are characterized by a blocking index. The perturbations maximize the difference in the blocking index between reference and perturbed flow at a prescribed forecast time (3 days). A linear sensitivity is defined on the basis of the amplification of the perturbations. It is demonstrated that periods of high linear sensitivity are very short. The method is also extended to the (weakly) nonlinear range (6 days) by applying an iterative procedure to the linear method. This procedure modifies the perturbations considerably. The optimal transition perturbations are compared to regional singular vectors, which maximize error growth within the Atlantic-European area. An optimal perturbation can be written as a sum of regional singular

vectors. In the linear range, optimal transition perturbations and fastest-growing regional singular vectors are strongly correlated during high-sensitivity periods. In the medium range, linearly slower-growing regional singular vectors also contribute to the growth of the optimal transition perturbations. This chapter has been published as an article in the *Journal of the Atmospheric Sciences* (Oortwijn and Barkmeijer 1995).

In chapter 3 we study characteristics of high- and low-sensitivity periods. We consider blocking and strong zonal flow regimes located over the Euro-Atlantic area and over the North-Pacific. High-sensitivity periods are characterized by an intensified jet stream to the west of a diffuent flow. Low-sensitivity flows have a jet stream that is weaker and more zonal. It is shown that sensitivity toward blocking is slightly larger than sensitivity to strong zonal flow in the medium range, due to nonlinear feedback mechanisms. Transitions to blocking appear to have, on average, a higher sensitivity than the mean sensitivity, whereas transitions to strong zonal flow have, on average, a mean sensitivity. This is probably an important reason why especially blocking onset is difficult to predict. This chapter will be published as an article in the *Journal of the Atmospheric Sciences* (Oortwijn 1998a).

In chapter 4 the growth properties of optimal transition perturbations are studied. The evolution of an optimal transition perturbation during a sensitive 72-hour period is extensively studied. The evolution can be divided in two phases. The first phase is short (12 to 24 hours), the growth is strongly non-modal and baroclinic. After that, the growth is still non-modal but not as strong and almost barotropic. The initial baroclinic process leads to a larger amplification than with barotropic growth only. The further evolution and the structure of the perturbation are determined by barotropic interactions between the reference flow and the perturbation. The barotropic evolution is also studied using a leading order WKB approximation. The WKB approach is used as a diagnostic tool to interpret the evolution of optimal perturbations qualitatively. Thus, we are able to determine characteristics of reference flows and perturbations that are important for the development of the perturbations. This chapter has been submitted for publication as an article in the *Journal of the Atmospheric Sciences* (Oortwijn 1998b).

Chapter 2

Perturbations that optimally trigger weather regimes

The sensitivity of the onset of two weather regimes with respect to initial conditions is studied. The weather regimes are a Euro-Atlantic blocking regime and a Euro-Atlantic strong zonal flow regime. Both regimes are characterized by the same anomaly pattern but with opposite sign. Using a three-level quasigeostrophic T21 model and its tangent linear and adjoint versions, initial perturbations are computed that have the largest projection on this anomaly pattern at a prescribed forecast time. The tangent linear and adjoint techniques can only be used to describe linear error growth. However, with an iterative procedure nonlinear error growth can be taken into account. In this way, perturbations can be computed that trigger the onset of a weather regime in the linear range (even optimally) as well as in the nonlinear range. It is shown that moderate initial perturbations occasionally trigger a transition from a blocking regime to a zonal flow regime, or vice versa, within three days. For an optimization time of six days, the iteratively computed perturbations generate such transitions for almost all investigated cases.

The perturbations are compared with regional singular vectors, which are the linearly fastest-growing perturbations in the Euro-Atlantic area. In the linear range, the perturbations project mainly onto the leading regional singular vectors. In the nonlinear range the projection onto linearly slower-growing regional singular vectors is stronger. The method can easily be generalized to study the sensitivity for a transition to any weather regime or anomaly pattern. This approach can be useful to generate specific initial conditions for ensemble forecasting.

2.1 Introduction

The quality of a deterministic forecast depends, among other things, on the synoptic situation. Sometimes, small errors in the initial conditions or model errors can grow very fast so that the performance of the forecast is low already after a few days. In other situations, errors grow more slowly and a forecast may be reliable up to 10 days. The work of Lorenz (1965) made clear that infinitesimal errors in the atmosphere may grow so fast that the larger scales are influenced in a finite time. As a result, each atmospheric scale has its own predictability limit. It is estimated that, on average, infinitesimal errors may have an impact

on the largest scales (~ 1000 km) within 2 weeks. Therefore, a deterministic forecast beyond 2 weeks will be impossible.

Yet, the atmospheric circulation still exhibits some regular behaviour, even for periods longer than 2 weeks. In the extratropics, persistent large-scale atmospheric flow patterns are observed. Such quasi-stationary flow patterns are called weather regimes. The low-frequency variability of the extratropical atmosphere can be considered to be mainly due to the alternation between several weather regimes, interrupted by transition periods (Vautard 1990). The maintenance of weather regimes is fairly well investigated. On the other hand, onset and breakdown of weather regimes are still very poorly understood. In this chapter, we study the potential for the excitation of a weather regime or a weather regime transition by adding small perturbations to the initial conditions, which are optimal in some prescribed sense. We consider only two regimes. The first regime is the Euro-Atlantic blocking regime. During a blocking regime, the weather in Western Europe is characterized by nice and dry weather in summer and very cold periods in winter. Blocking flow patterns have been observed that persisted for months (Namias 1964). The second one is a strong zonal flow regime in the same geographical area. In this regime, the weather is highly variable because of the continuous advection of weather systems. Liu (1994) has studied both regimes and has showed that the probability distribution of the duration of the regimes is about the same.

In order to study the sensitivity of regime transitions to perturbations, we first need an objective tool to classify to what extent a particular flow pattern can be associated with a blocking or with a strong zonal flow regime. The two regimes are characterized by the same anomaly pattern but with opposite sign. Therefore a single index, called BINX (Liu 1994), measuring the amplitude of the anomaly pattern is used. Hereafter this index will be denoted by \mathcal{B} . The characterization of the regimes will be outlined in section 2.2. In section 2.3 we derive a method to compute initial perturbations that maximize the amplitude of this anomaly pattern at a given forecast time. So these perturbations cause a maximum change to one of the regimes. The computations are performed with a three-level quasigeostrophic model triangularly truncated at wavenumber 21 (T21QG), its tangent linear and adjoint versions. The tangent linear model describes the linear evolution of small errors along a reference orbit. With current analysis errors, the evolution is approximately linear during the first two or three days (Lacarra and Talagrand 1988; Vukićević 1991; Errico et al. 1993). Results for the quasi-linear range are presented in section 2.4. In section 2.5, a modification of the method is presented that enables extension of the method into the nonlinear range. Results of the modification method are presented in this section. In section 2.6, the perturbations that optimally trigger a regime transition are compared with regional singular vectors, which maximize linear error growth over the Euro-Atlantic area. Finally, some concluding remarks are given in section 2.7.

2.2 Characterization of weather regimes

Many different criteria have been used to characterize an atmospheric blocking regime. The best known criteria are those by Rex (1950a,b) and Dole (1978). According to Rex the 500-hPa height field of a blocking phenomenon must show a split of the jet flow, a sharp transition

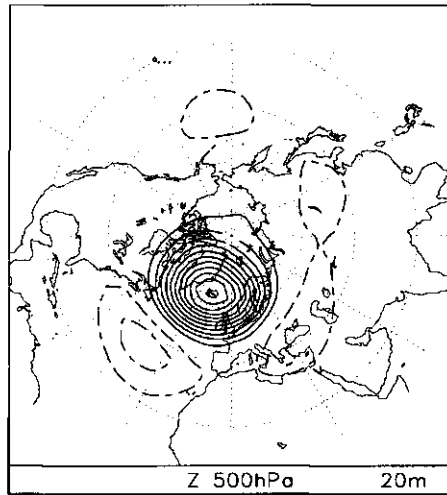


Figure 2.1: Blocking anomaly pattern: mean 500-hPa geopotential height anomaly pattern for the 10000 days out of 45000 days with the largest positive anomalies at 60° N. The contour interval is 20 m. Solid lines correspond to positive values, dotted lines to negative values.

from westerly to meridional flow and furthermore it must be persistent. Dole characterized blocking as a persistent flow with large 500-hPa geopotential height anomalies. It appears that there are two preferred regions for blocking in the Northern Hemisphere, over the central Pacific and over the eastern Atlantic. In this chapter we focus on the atmospheric flow over Europe. Using observations for 10 winter seasons (DJF) from 1982/83 to 1991/92, Liu (1994) found that a Euro-Atlantic blocking regime can be characterized by a dipolelike pattern, consisting of a very strong positive geopotential height anomaly with its center at about 60° N and a weaker negative anomaly south of it. Furthermore, he found that a strong zonal flow regime is characterized by the opposite dipole pattern. So, both regimes have approximately the same anomaly pattern with respect to the climatological mean but with opposite sign. Liu computed this anomaly pattern, which he called the blocking geopotential height anomaly pattern z_b , for the 10 winter season data set and for a data set simulated by a T21QG model (Liu and Opsteegh 1995). The latter z_b pattern is the mean 500-hPa geopotential height anomaly pattern for the 10000 days out of 45000 days of integration that had the largest positive anomalies at 60° N. This pattern is shown in Fig. 2.1 It is very similar to the pattern computed for the observed winter season data set. Based on these results Liu (1994) and Liu and Opsteegh (1995) defined a single index that measures the resemblance of a particular circulation pattern with the blocking regime or the strong zonal flow regime. This index is called the blocking index \mathcal{B} . The blocking index \mathcal{B} for a particular circulation pattern is defined by the projection of the daily geopotential height anomaly pattern z_d on the blocking geopotential height anomaly pattern z_b , weighted with the norm of z_b . The blocking index \mathcal{B} of a certain flow pattern with streamfunction ψ is

given by

$$\mathcal{B}(\psi) = \frac{\langle z_d(\psi), z_b \rangle}{\langle z_b, z_b \rangle}, \quad (2.1)$$

where the brackets denote a squared norm inner product on a sphere, integrated over height

$$\langle x, y \rangle = \frac{1}{4\pi} \int \int x y dV. \quad (2.2)$$

The relation between z_d and ψ is given by

$$z_d(\psi) = \mathbf{L}_z \psi - z_c, \quad (2.3)$$

where \mathbf{L}_z is the linear operator that transforms streamfunction into geopotential height and z_c is the climatological mean geopotential height. The geopotential height field is obtained from the streamfunction by solving the linear balance equation. A circulation pattern with $\mathcal{B} > 0.5$ ($\mathcal{B} < -0.5$) can be considered as a blocked flow (strong zonal flow). Furthermore, the larger \mathcal{B} the more pronounced the blocked flow is and the more negative the stronger the zonal flow is. Typical values of \mathcal{B} are between -2 and 2.

2.3 Perturbations triggering a regime onset

The blocking index \mathcal{B} characterizes in a simple manner whether a given atmospheric circulation pattern can be associated with a blocking regime or a strong zonal flow regime. Small perturbations in the initial conditions of a forecast will alter \mathcal{B} at a certain forecast time. It may even occur that \mathcal{B} changes sign, corresponding to a regime transition. We want to investigate how the predicted value of \mathcal{B} depends on small changes in the initial conditions. Therefore, we will determine the initial perturbation that maximizes the difference in \mathcal{B} between the reference forecast and the forecast made from the perturbed initial state at a prescribed forecast time. In other words, this perturbation maximizes the difference between the two forecasts in the direction of phase space corresponding to the dipolelike anomaly pattern. This approach could be considered as an extension of the study of Barkmeijer (1993), where error growth is maximized in a single grid point.

For the computations, we use a three-level quasi-geostrophic model triangularly truncated at wavenumber 21 (T21QG), its tangent linear and adjoint versions. The levels of the model are at 200, 500, and 800 hPa. The T21QG model is described in Marshall and Molteni (1993). The tangent linear model describes, to first order, the evolution of a small perturbation. The evolution of a streamfunction perturbation is given by the tangent linear operator \mathbf{R} , such that

$$\varepsilon(T) = \mathbf{R}(0, T) \varepsilon(0). \quad (2.4)$$

The period for which error growth is linear will be called the linear range. With realistic analysis errors the evolution of errors is linear up to about 72 hours (Errico et al. 1993). However, the length of the linear range is not constant but depends on the synoptic situation, the model used and the amplitude and structure of the error itself (Lacarra and Talagrand 1988). In this work, we will assume that the T21QG model describes the evolution of the

atmospheric flow perfectly, i.e. a perfect model approach. So, all forecast errors are due to errors in the initial conditions.

We are interested in the perturbation ε , keeping

$$\|\varepsilon(0)\|^2 = \langle \varepsilon(0), \varepsilon(0) \rangle \quad (2.5)$$

fixed, that maximizes

$$\mathcal{B}(\psi(T) + \varepsilon(T)) - \mathcal{B}(\psi(T)) \quad (2.6)$$

at prescribed forecast time T . Using the definition of \mathcal{B} (Eq. (2.1)), this can be written as

$$\frac{\langle \mathbf{L}_z(\psi(T) + \varepsilon(T)) - z_c, z_b \rangle - \langle \mathbf{L}_z\psi(T) - z_c, z_b \rangle}{\langle z_b, z_b \rangle} \quad (2.7)$$

The denominator is just a constant. Therefore, maximizing Eq. (2.7) is equivalent to maximizing

$$\langle \mathbf{L}_z\varepsilon(T), z_b \rangle = \langle \mathbf{L}_z\mathbf{R}\varepsilon(0), z_b \rangle = \langle \varepsilon(0), \mathbf{R}^*\mathbf{L}_z^*z_b \rangle, \quad (2.8)$$

where we introduced the operators \mathbf{R}^* and \mathbf{L}_z^* which are the adjoint operators of \mathbf{R} and \mathbf{L}_z , respectively, with respect to the squared norm inner product. A good introduction to the application of adjoint methods in meteorology has been given by Talagrand and Courtier (1987). It follows that the initial perturbation that maximizes the difference in \mathcal{B} at optimization time T is given by

$$\varepsilon_L(0) = \lambda \mathbf{R}^*\mathbf{L}_z^*z_b. \quad (2.9)$$

Here, the subscript L denotes that the perturbation is optimal when the error growth is linear. The absolute value of the scaling factor λ is fixed through the constraint, Eq. (2.5). When choosing λ positive, ε_L maximizes the change towards a blocking regime. A negative value of λ results in a maximum change towards a strong zonal flow regime. The pattern ε_L is related to the sensitivity fields described in Rabier et al. (1993, 1996). The difference in \mathcal{B} at optimization time between the reference and perturbed forecast for this optimal perturbation is

$$\Delta\mathcal{B} = \frac{\|\mathbf{R}^*\mathbf{L}_z^*z_b\|}{\|z_b\|^2} \|\varepsilon_L(0)\|. \quad (2.10)$$

The difference $\Delta\mathcal{B}$ is a linear function of the length of the optimal perturbation vector, provided the maximization period is kept constant. This linear relation is only valid for small initial perturbations, i.e. in the linear range.

We now define the linear sensitivity by

$$S_L = \frac{\|\mathbf{R}^*\mathbf{L}_z^*z_b\|}{\|z_b\|^2}. \quad (2.11)$$

The linear sensitivity S_L is a measure for the (maximum) divergence of trajectories when projected onto the direction of phase space that corresponds to the blocking anomaly pattern. We will see that adding a moderate perturbation $\varepsilon_L(0)$, as given in Eq. (2.9), to the reference forecast that has high linear sensitivity S_L for a forecast period of 3 days may already lead to a substantial difference in \mathcal{B} . Even a regime transition is sometimes triggered within 3 days.

We want to emphasize that the numerical costs of the computation of $\varepsilon_L(0)$ are very low. Only one backward integration of the adjoint model is needed.

2.4 Results for the quasi-linear range

Starting from an arbitrary initial condition, we integrated T21QG for 1200 days. We computed the blocking index each day for the final 1000 days. A part of the time series obtained in this way is presented in Figs 2.2a and 2.2c. The single valued function \mathcal{B} is a very useful quantity to decide whether the model state is in a blocking regime ($\mathcal{B} > 0.5$) or in a strong zonal flow regime ($\mathcal{B} < -0.5$). The T21QG model is capable of entering both regimes in a realistic way (Marshall and Molteni 1993; Liu and Opsteegh 1995). From Figs 2.2a and 2.2c it is clear that regime transitions extend over some days but sometimes are very abrupt (e.g. around day 60).

The linear sensitivity S_L for a forecast period of 3 days is computed for the same time series and is given in Figs 2.2b and 2.2d. A high linear sensitivity for a certain day d means that \mathcal{B} will differ substantially at day $d + 3$ if the optimal perturbation is added to the reference flow at day d . The real difference $\Delta\mathcal{B}$ depends on the amplitude of the initial perturbation and on the degree of linearity of the error growth. Low values of S_L indicate that moderate initial perturbations are not capable of changing the flow substantially towards a blocking or strong zonal flow regime. This indicates a high predictability of the regime type, because no initial perturbation can change the evolution of the flow strongly towards one of the regimes. So, if there is not a regime transition in the forecast, it is likely that the flow remains in the same regime. However, when there is a regime transition forecasted, it is likely that it will occur. In cases of high values of S_L moderate initial perturbations that change the flow towards one of the regimes can be found, so that the predictability is less in this situation. However, the difference $\Delta\mathcal{B}$ due to ε_L is an upper bound. So, a low sensitivity to the optimal error $\varepsilon_L(0)$ implies a relatively high predictability of the regime type. On the other hand, a high sensitivity to $\varepsilon_L(0)$ does not necessarily imply a high sensitivity to a (random) analysis error.

Figures 2.2b and 2.2d show that there are days where the sensitivity is small and days where the sensitivity is high. Differences may be as large as a factor of 5. Notice that the periods with a high linear sensitivity are very short. The majority of the days are insensitive to regime transitions.

Considering together Figs 2.2a and 2.2b and also Figs 2.2c and 2.2d, we see that low sensitivity can go along with quasi-stationary flows (e.g. around day 80) or with situations where a transition takes place (e.g. around day 60), which means that this particular transition is highly predictable. During periods with a high sensitivity, transitions may (e.g. around days 130,730) or may not occur (e.g. around day 690). The same characteristic features are shown in the simple three-variable Lorenz convective model that contains two regimes (Lorenz 1963; Palmer 1993).

To illustrate what the impact of ε_L is during a period with a high sensitivity we select day 687. The initial optimal perturbation $\varepsilon_L(0)$ for the period 687 to 690 is added to the reference field at day 687. The scaling factor λ , see Eq. (2.9), is chosen in such a way that the error in the geopotential height of the 500-hPa level is everywhere smaller than 15 m. The reference field and the perturbed field are integrated with the nonlinear T21QG model for a forecast period of four days. The evolution of these fields and of the error growth are given in Fig. 2.3 for the 500-hPa level. The reference field (Fig. 2.3b) and perturbed field (Fig. 2.3c)

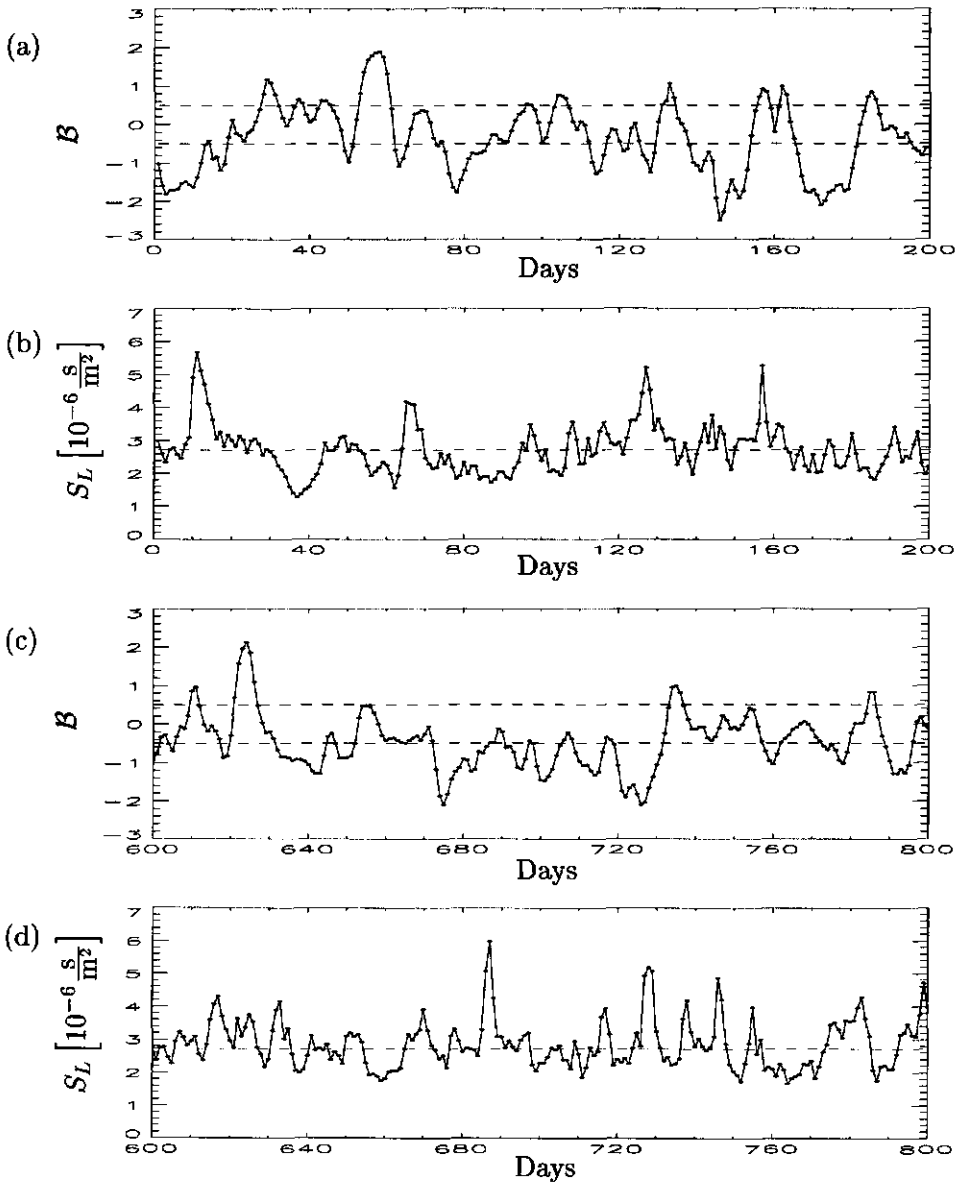


Figure 2.2: (a) B in the period 1 to 200. (b) Linear sensitivity S_L for a forecast period of 3 days during the same period. (c) B and (d) S_L in the period 600 to 800.

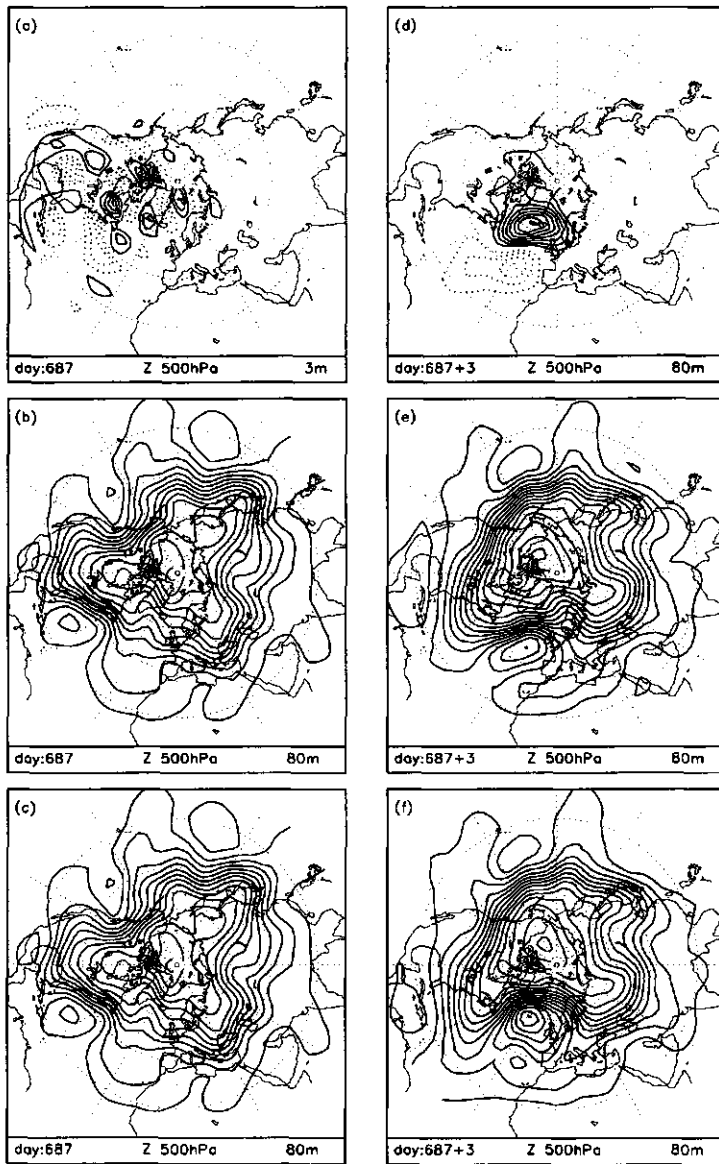


Figure 2.3: Geopotential height field at 500 hPa of (a) initial optimal perturbation $\epsilon_L(0)$ for day 687 with optimization time of 3 days, (b) the reference orbit at day 687 and (c) the perturbed orbit at day 687. Nonlinear integration for 3 days of (a), (b) and (c) results in (d), (e) and (f) respectively. Contour intervals are given in each panel.

at day 687 are both zonal above the Atlantic Ocean and Western Europe: for both fields $\mathcal{B} = -0.49$. The initial optimal perturbation, as is shown in Fig. 2.3a, is mainly located above North America with its energy distribution concentrated in the smaller scales and has a baroclinic structure (200- and 800-hPa levels are not shown in the Figures). Within one day the perturbation energy is transferred to lower wave numbers (i.e. the larger scales are more pronounced) and the error pattern reveals an equivalent barotropic dipolelike structure. The flow in the western part of the Atlantic Ocean becomes more meridional in the perturbed run. In the consecutive days the error grows very fast and is advected to the east (Fig. 2.3d). The difference $\Delta\mathcal{B}$ between the perturbed and reference field after three days (optimization time) is 2.02 and the transition to a blocked flow in the perturbed run is clearly visible ($\mathcal{B} = 1.78$), see (Fig. 2.3f). After four days $\Delta\mathcal{B}$ is even 2.52, the reference field is in the strong zonal flow regime ($\mathcal{B} = -0.63$) and the perturbed field in the blocking regime ($\mathcal{B} = 1.89$).

2.5 Extension to the nonlinear range

The operational predictability of blocking has been studied by Tibaldi and Molteni (1990). They have investigated the ability of the ECMWF General Circulation Model (GCM) to represent Euro-Atlantic and Pacific blocking onset and maintenance. It was found that blocking frequency and duration were underestimated in medium-range forecasts. New versions of the model have now been improved on these points (Palmer et al. 1990). Furthermore, it appeared that when the initial conditions were already blocked, the duration was reasonably well predicted. Very short-range forecasts of blocking onset are fairly successful but the onset is almost consistently missed beyond day 3 to 4. The inability to predict blocking onset in the medium range still has a substantial impact on the systematic model error. As a result of this inability, it is relevant to study the sensitivity of regime transitions to initial conditions beyond day 3. However, by then the evolution of analysis errors can not be considered linear anymore (Lacarra and Talagrand 1988; Vukićević 1991; Errico et al. 1993). As a consequence, the problem arises that the error growth can not be described with a tangent linear model. Therefore, it is proposed to modify the technique as outlined in section 2.3 for the nonlinear range.

The modification is based on the following idea. If we add a perturbation with a realistic amplitude to the initial conditions, the error growth is, on average, linear up to day 2 or 3. However, if we add a perturbation that is a few times smaller this period will be longer. The amplitude of the initial perturbation and the growth rate determine the length of the linear range. In a finite dimensional model this range can always be extended by decreasing the amplitude of the initial perturbation. On the other hand, in the real atmosphere, an infinite dimensional system, infinitesimal perturbations may evolve into large perturbations in a finite time. It follows that although in the model context the optimization time can be made very large, the modification only makes sense, a priori, for a maximum optimization time between one and two weeks.

The modification of ε_L is achieved by applying an iterative procedure. We start the modification by adding a perturbation $\varepsilon_0(0)$, which lies in the direction of ε_L , to the reference orbit ψ_0 . The amplitude of $\varepsilon_0(0)$ must be taken such that its error growth is linear during

a period of time T ($T < 10$ days). A new reference orbit ψ_1 is obtained by performing a nonlinear integration with T21QG starting from $\psi_0(0) + \varepsilon_0(0)$. Iteratively we now determine $\varepsilon_i(0)$, keeping $\langle \varepsilon_i(0), \varepsilon_i(0) \rangle$ fixed, so that

$$\mathcal{B}(\psi_i(T) + \varepsilon_i(T)) - \mathcal{B}(\psi_0(T)) \quad (2.12)$$

is maximal or minimal, corresponding to a maximum change towards the blocking regime or strong zonal flow regime, respectively. In this, we assume that the evolution of $\varepsilon_i(0)$ is linear with respect to ψ_i for time T . After each iteration, a new reference orbit ψ_{i+1} is determined by starting a model integration from $\psi_i(0) + \varepsilon_i(0)$. Because we only need to maximize the difference between two scalar quantities, it is easy to derive that $\varepsilon_i(0)$ can also be obtained by maximizing or minimizing the expression

$$\mathcal{B}(\psi_i(T) + \varepsilon_i(T)) - \mathcal{B}(\psi_i(T)) . \quad (2.13)$$

Thus, the computation of $\varepsilon_i(0)$ is reduced to the determination of ε_L with respect to ψ_i , which is described in section 2.3. The iterative procedure ends when the sum of the perturbations

$$\varepsilon_{NL}(0) = \sum_{i=0}^N (\varepsilon_L(0))_i \quad (2.14)$$

has the same norm as the original linearly optimal perturbation. In the linear case, the perturbation that is optimal for inducing a blocking regime differs only in sign from the one that optimally induces a strong zonal flow regime. However, in the nonlinear case the perturbations, obtained by either maximizing or minimizing each step (Eq. (2.13)), have different structures after at least two steps. Namely, after adding or subtracting $\varepsilon_1(0)$ to the reference flow two different new reference orbits will be obtained, so that the maximization and minimization procedures are not symmetric anymore. The procedure does not guarantee that of all perturbations with this norm, ε_{NL} results in the maximum difference in \mathcal{B} because the iterative algorithm can also lead to a local maximum. However, as we will see, the procedure works well in the sense that when ε_{NL} is added to the initial streamfunction $\psi(0)$, $\Delta\mathcal{B}$ is substantially larger than obtained with ε_L when the nonlinear range is entered.

In this way, we can still use the linear adjoint operator \mathbf{R}^* in each iteration but are able to extend the method into the nonlinear range. The same procedure is described in Barkmeijer (1996). Maximizing the cost function $\Delta\mathcal{B}$ iteratively is related to the iterative minimization procedures used in the context of four-dimensional data assimilation, see Thépaut et al. (1991) and Zupanski (1993).

In Fig. 2.4 results are given for an optimization time T of 6 days (144h) for all days between days 701 and 800. This period is chosen because of the high variability in the linear sensitivity S_L (see Fig. 2.2d). The difference in \mathcal{B} of the perturbed flow towards a blocking regime is computed in three ways. The dashed line gives $\Delta\mathcal{B}$ when the perturbation ε_L is added to the reference flow and error growth would be purely linear. The perturbation ε_L has every day the same initial amplitude such that the maximum value is about 15 m at 500 hPa. One way to compute $\Delta\mathcal{B}$ is to integrate ε_L with the tangent linear operator R , but it can also be calculated directly using Eq. (2.10). The dotted line gives for the same

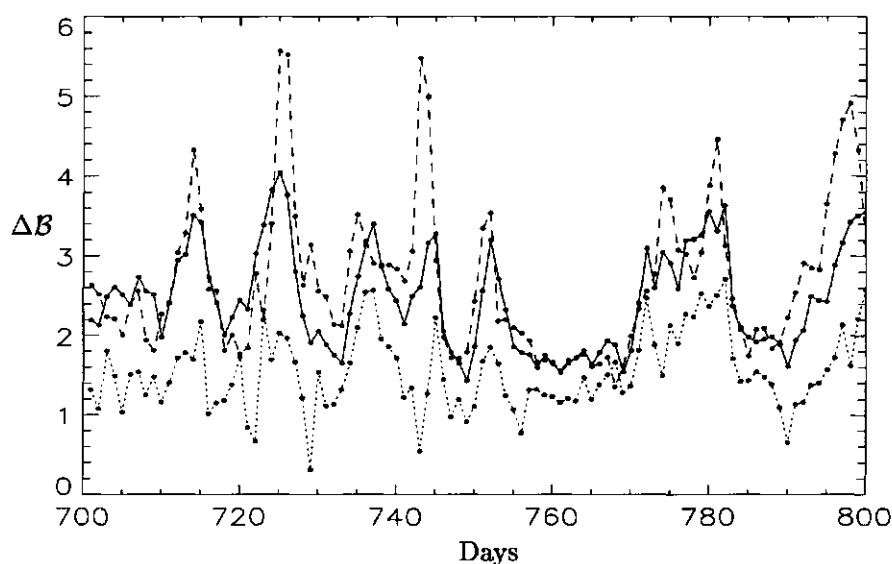


Figure 2.4: The difference ΔB when the optimal perturbation is added to the reference field with an optimization time T of 6 days for all days in the period 701 to 800. The dashed line corresponds with a linear integration of ε_L and a nonlinear integration of ε_L is given by a dotted line. The solid line shows ΔB when ε_{NL} is nonlinearly integrated.

perturbation ε_L the results for the nonlinear integrations. Comparing these two lines, one can see that in general the linear theory overestimates the error growth for ε_L . The solid line gives ΔB when ε_{NL} with the same length as ε_L is added to the initial conditions and integrated in time with the nonlinear model. The number of iterations used varies between 5 and 7. It appeared that the iterative procedure is not very sensitive to the size of the amplitude of the optimal perturbation that is computed in each iteration. Comparing the results of the nonlinear integrations with the orbits perturbed with ε_L and ε_{NL} one can conclude that modification of the initial error pattern ε_L can compensate for the saturation of the error growth due to nonlinear interactions.

We point out that the dashed line and solid line are correlated ($\rho = 0.76$). It seems that linear theory still enables one to determine if a circulation pattern is sensitive to small changes in the initial conditions or not, although nonlinear error growth is non-negligible. So, S_L still might provide an efficient predictor for possible transitions. If one is interested in the fastest-growing error patterns, the computation of ε_{NL} is needed.

An example that clearly illustrates the impact of the modification procedure is shown in Figs 2.5 and 2.6. In Fig. 2.5, the evolutions of ε_L and ε_{NL} are compared. The reference forecast orbit starts from day 722 and the optimization time T is 6 days. Both perturbations ε_L and ε_{NL} have initially the same length and their geopotential height fields at 500 hPa are given in Figs 2.5a and 2.5b respectively. In computing ε_{NL} we used 7 iterations. The main

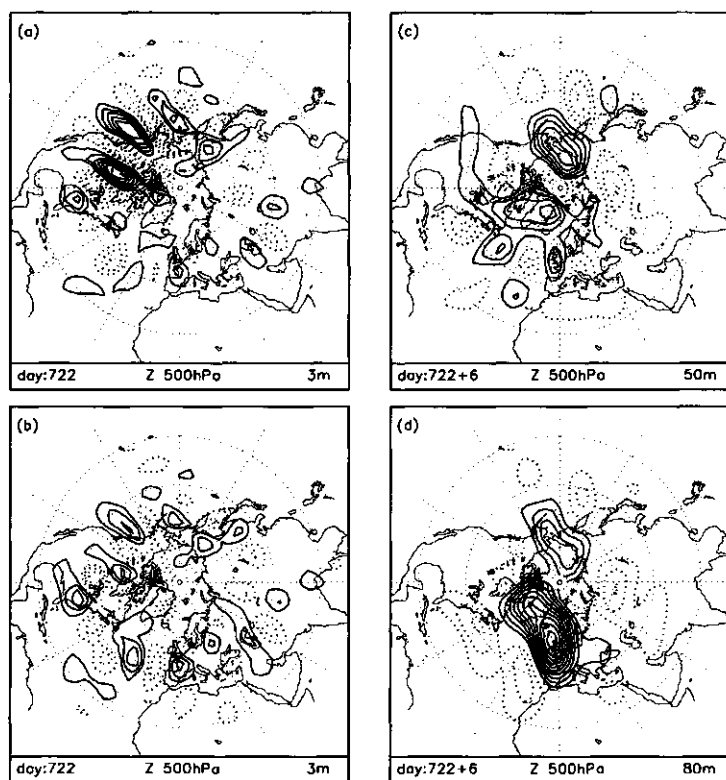


Figure 2.5: Geopotential height field of the initial optimal perturbations (a) ε_L and (b) ε_{NL} at 500 hPa for day 722 with optimization time of 6 days. Final errors (c) ε_L and (d) ε_{NL} after 6 days. Contour intervals are given in each panel.

part of ε_L is located in the North American area, whereas ε_{NL} has a more global nature. The nonlinear modification of ε_{NL} mainly affects the strength of the local structures, not the position. After two days the initially small-scale, baroclinic structures have evolved into larger scale and equivalent barotropic structures. The error growth of ε_L is faster than ε_{NL} during the first four days, which can be expected, but decreases after day five. At optimization time, ε_L has only a small amplitude over the Atlantic and Western European areas (Fig. 2.5c). The modified perturbation ε_{NL} continues to grow in the nonlinear range and shows a large projection onto the blocking anomaly pattern at optimization time (Fig. 2.5d). Figure 2.6 gives the geopotential height patterns at optimization time of the reference orbit (Fig. 2.6a), the orbit perturbed with ε_L (Fig. 2.6b) and ε_{NL} (Fig. 2.6c). The perturbation ε_L was not able to initiate a transition. Both Figs 2.6a and 2.6b are in the strong zonal flow regime with B values of -2.1 and -1.4 respectively. Figure 2.6c indicates that a transition to a blocked flow is possible at day 6 ($B = 0.86$) by using ε_{NL} as initial perturbation.

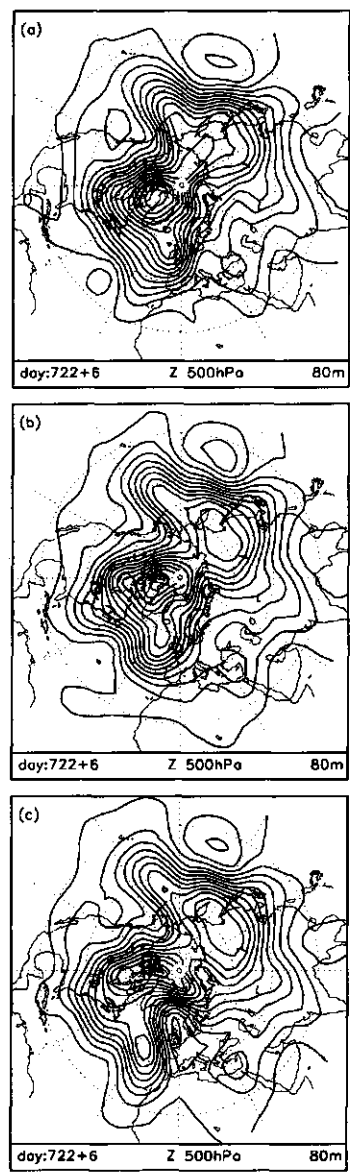


Figure 2.6: Geopotential height patterns at optimization time (6 days) of (a) reference orbit and the reference orbit perturbed with (b) ϵ_L and (c) ϵ_{NL} after nonlinear integration. Contour interval is 80 m.

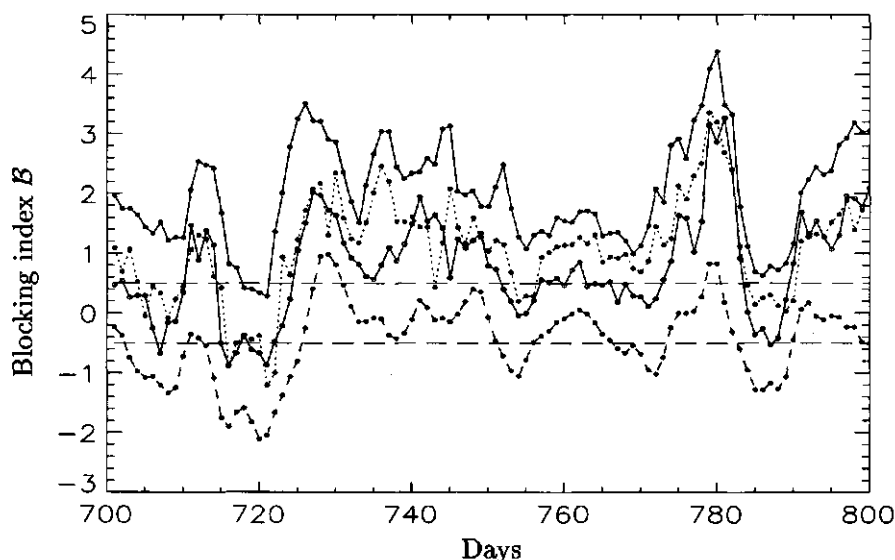


Figure 2.7: B in the period 701 to 800 for the reference flow (dashed line) and the flows perturbed with ε_L (dotted line), ε_{NL} (solid line) or ε_{NL}^\perp with $M=5$ (line marked with open circles), all nonlinearly integrated.

In Fig. 2.7, ΔB at forecast day 6 computed with T21QG for all days between days 701 and 800, shown in Fig. 2.4, is added to B of the reference flow. It turns out that 74 days perturbed with ε_L and 94 days perturbed with ε_{NL} are blocked ($B > 0.5$). This means that almost always a block can develop within six days. In the same way, we have tried to decrease B every day to get a strong zonal flow. We have found the same high rates of transitions to this regime (not presented). Of course, the probability of such a development is important. Results show that by adding ε_{NL} to the reference flow, the model is forced into regions on the attractor that it rarely visits. In almost 50% of all cases B becomes larger than 2, a value that is exceeded only a few times in the 1000-day data set (Figs 2.2a and 2.2c). From this, it can be concluded that the optimal error is very special. Probably, an analysis error usually has a low projection on it.

2.6 Relation to singular vectors

The eigenfunctions of the product operator $\mathbf{R}^*\mathbf{R}$ are usually called the singular vectors of \mathbf{R} and the square roots of the eigenvalues the singular values of \mathbf{R} . The singular vectors with the largest singular values maximize linear error growth, measured with a chosen norm, at a certain forecast time. The product operator $\mathbf{R}^*\mathbf{R}$ is always symmetric, by construction. The singular vectors are therefore complete and orthogonal (in contrast to the eigenvectors of \mathbf{R} , which are for stationary reference flows the familiar normal modes). We will now define regional singular vectors to be the fastest-growing errors over the Atlantic-European area,

which corresponds to the blocking anomaly area. We can compare ε_L with these patterns. In order to define the regional singular vectors, we make use of a linear projection operator \mathbf{P} . A similar projection operator has been used by Barkmeijer (1993) and Buizza (1994). The projection operator \mathbf{P} acts on a global error field $\varepsilon(T)$ and projects onto the Atlantic-European area (10°N - 85°N , 90°W - 60°E) at the 500-hPa level. The regional streamfunction error vector $\mathbf{P}\varepsilon(T)$ is a N -dimensional vector, with $N = 231$. The regional squared forecast error will be defined as $(\mathbf{P}\varepsilon(T), \mathbf{P}\varepsilon(T))$, where (\cdot, \cdot) denotes the Euclidean inner product running over all 3 levels:

$$(x, y) = \sum_{j=1}^3 \sum_{i=1}^N x_{ij} y_{ij}. \quad (2.15)$$

Although the field $\mathbf{P}\varepsilon(T)$ is only non-zero at the 500-hPa level, the singular vectors that maximize the regional squared forecast error are in general a field at all levels. The definition of the Euclidean inner product is therefore extended to all three levels. By using adjoint operators we can put the regional squared forecast error in the form:

$$(\mathbf{S}\mathbf{R}^*\mathbf{P}^*\mathbf{P}\mathbf{R}\mathbf{S}^{-1}\xi, \xi), \quad \text{with} \quad \xi = \mathbf{S}\varepsilon(0). \quad (2.16)$$

The coordinate transformation \mathbf{S} has the property that the Euclidean norm of ξ equals the squared norm of $\varepsilon(0)$:

$$(\xi, \xi) = (\mathbf{S}\varepsilon(0), \mathbf{S}\varepsilon(0)) = \langle \varepsilon(0), \varepsilon(0) \rangle. \quad (2.17)$$

In maximizing Eq. (2.16) with constraint (Eq. (2.17)) (fixed $\|\varepsilon(0)\|$), we employ the Lanczos algorithm. In order to apply this algorithm, which makes use of matrix algebra, it was necessary to define the regional squared forecast error using an Euclidean inner product instead of a squared norm inner product. A description of the Lanczos algorithm can be found in Parlett (1980). With the Lanczos algorithm the eigenvalues and their corresponding eigenvectors are found iteratively in decreasing magnitude of the eigenvalues. The advantage of this algorithm is that an explicit form of the product operator $\mathbf{S}\mathbf{R}^*\mathbf{P}^*\mathbf{P}\mathbf{R}\mathbf{S}^{-1}$ is not required. The first eigenvector maximizes Eq. (2.16) with constraint (Eq. (2.17)). The second one maximizes Eq. (2.16) in the subspace orthogonal to the first eigenvector, also with constraint (Eq. (2.17)) and so on *.

The regional singular vectors are complete and span the N -dimensional vector space. The error ε is thus a linear combination of the N regional singular vectors:

$$\varepsilon(0) = \sum_{i=1}^N \eta_i v_i, \quad (2.18)$$

where v_i denotes the i 'th regional singular vector. We will now investigate which regional singular vectors are important for the construction of ε_L and ε_{NL} . Therefore, the (normalized)

The numerical computations are performed by first computing the eigenvalues of the lower-dimensional product operator $\mathbf{P}\mathbf{R}\mathbf{R}^\mathbf{P}^*$ which are identical to the non-zero eigenvalues of the higher-dimensional product operator $\mathbf{S}\mathbf{R}^*\mathbf{P}^*\mathbf{P}\mathbf{R}\mathbf{S}^{-1}$. The eigenvectors w_i of $\mathbf{S}\mathbf{R}^*\mathbf{P}^*\mathbf{P}\mathbf{R}\mathbf{S}^{-1}$ are related to the eigenvectors u_i of $\mathbf{P}\mathbf{R}\mathbf{R}^*\mathbf{P}^*$ through the relation $w_i = \frac{1}{\lambda_i} \mathbf{S}\mathbf{R}^*\mathbf{P}^* u_i$. In this way, we find a set of independent regional singular vectors $v_i = \mathbf{S}^{-1} w_i$.

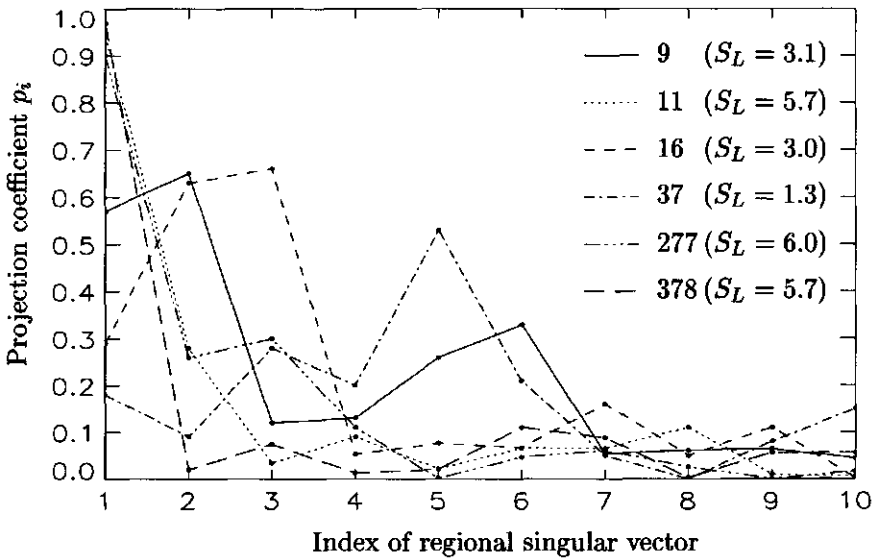


Figure 2.8: Projection of ε_L onto the 10 fastest regional singular vectors, for six initial conditions (optimization time of 3 days). The linear sensitivity S_L is given in units of 10^{-6} .

projection p_i of ε on v_i is defined by

$$p_i = \frac{|\eta_i|}{\|\varepsilon\|}, \quad (2.19)$$

so that the sum over all squared projections $\sum_{i=1}^N p_i^2$ equals 1.

We will first show some results for the quasi-linear range (optimization time $T=3$ days). The projection of ε_L on the corresponding 10 fastest regional singular vectors for an integration period of 3 days is given in Fig. 2.8 for six initial conditions. For each case, the regional singular vectors have been computed. Days 11, 277, 378 are very sensitive to a change of the initial conditions. Days 9 and 16 have a mean sensitivity and day 37 is the most insensitive day of the data set (Fig. 2.2b). The initial perturbations ε_L of the 3 most sensitive days all have a very high projection on the first regional singular vector. For day 9 the most important regional singular vectors are the first and second, for day 16 the second and third. Day 37 has the highest projection on the fifth regional singular vector. It appears that in the linear range the optimal perturbations ε_L can be constructed in nearly all cases from combinations of the first five regional singular vectors. Especially in the case of very sensitive days ε_L grows very fast and so the leading regional singular vectors must dominate the error pattern. For less sensitive days usually more regional singular vectors further down in the spectrum are needed.

In the nonlinear range many more regional singular vectors are involved in the construction of ε_{NL} . The regional singular vectors are ordered according to their linear growth rate.

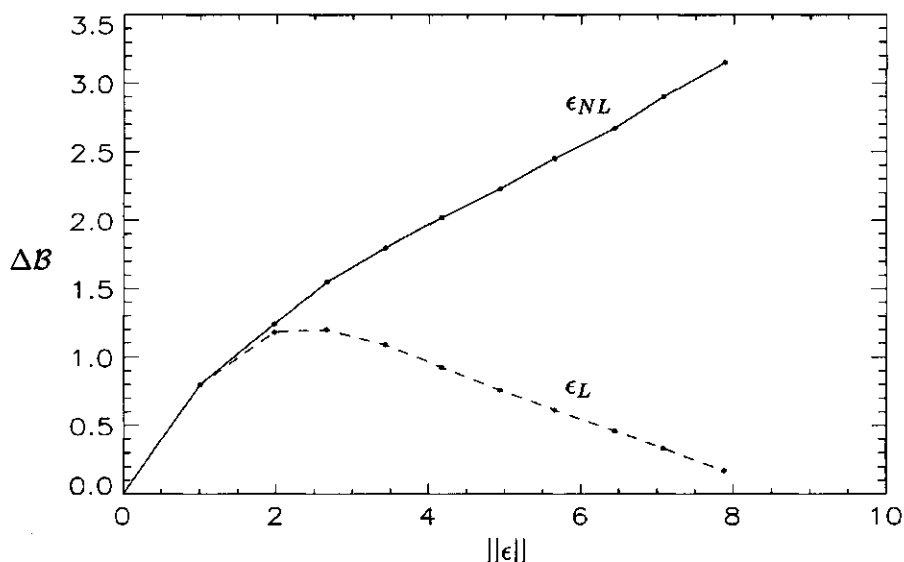


Figure 2.9: ΔB as a function of the length of the initial error using ϵ_{NL} (solid line) or ϵ_L (dashed line). The length of the perturbation vectors are given in units of 10^{-5} .

Nonlinear interactions can change this order. In the following, we will focus on perturbations that trigger the onset of a blocking regime. The conclusions from these experiments are comparable to the ones obtained for the onset of a strong zonal flow regime. We have computed ϵ_{NL} and the regional singular vectors for an optimization time of $T=6$ days, starting at day 743. After 10 iterations, the maximum amplitude of ϵ_{NL} is less than 20 m at 500 hPa. For each iteration ΔB at forecast time T is calculated. The results are given in Fig. 2.9. We can see that ϵ_{NL} computed with 10 iterations changes B by 3.1. In Fig. 2.9 is also given ΔB as a result of nonlinear integrations of the linearly optimal perturbation (obtained by 1 iteration) as a function of its amplitude. It is obvious that this perturbation is not optimal anymore for larger initial amplitudes of the perturbation, when nonlinear terms have a larger impact. Next, we have computed the projection of ϵ_{NL} onto the regional singular vectors. The projection p_1 onto the first regional singular vector is 0.2. When we integrate only this part of ϵ_{NL} , ΔB is 0.85. So, the contribution of the first regional singular vector to the total difference in B is 26 %. In Fig. 2.10, the percentage of the total difference in B when using only the first M regional singular vectors is shown as a function of M . A combination of the first five regional singular vectors explains 61% of ΔB . Many more regional singular vectors are required to increase the change substantially. We therefore conclude that relatively slow-growing singular vectors (in a linear sense) are contributing to error growth in the nonlinear range. In Fig. 2.10 are also given the results of the linearly optimal perturbation for a period of $T=3$ days, starting again at day 743. The contribution of higher-index regional singular vectors to ΔB is very low.

In most cases, the projection of ϵ_{NL} on one (or more) leading regional singular vectors

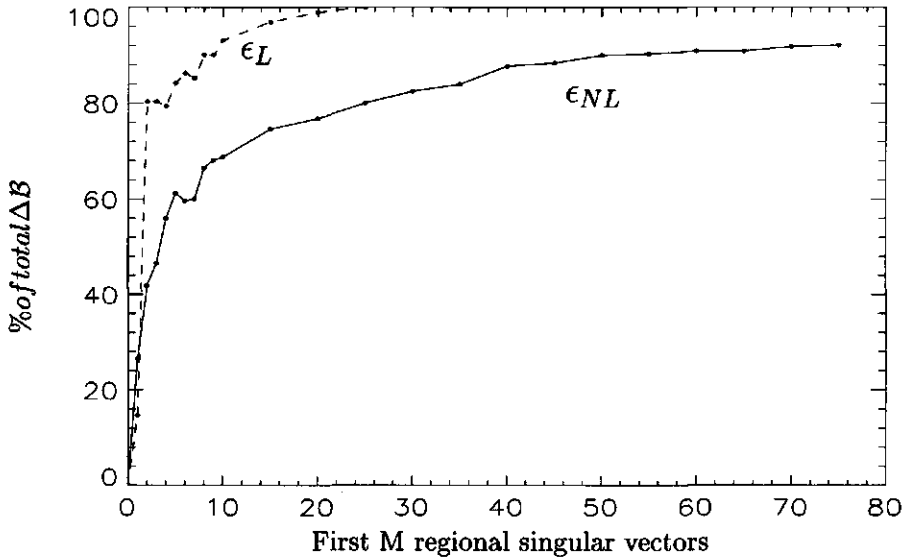


Figure 2.10: The percentage of the total difference ΔB explained by the pattern by ϵ_L (optimization time 3 days) and ϵ_{NL} (optimization time 6 days) when only the first M regional singular vectors are used in the expansion.

is high. However, it does not imply that for those cases combinations of the linearly slower-growing singular vectors can not produce a large difference in B too. Although the differences will normally be less than the difference induced by ϵ_{NL} , it can still be very large. To study the impact of the linearly slower-growing regional singular vectors, we will now compute the perturbation that is optimal for a regime onset with the additional constraint that it is orthogonal to the first M regional singular vectors. This perturbation will be denoted by ϵ^\perp . We then have to maximize, see Eq. (2.8),

$$\langle \epsilon(0), b \rangle, \quad (2.20)$$

where $b = \mathbf{R}^* \mathbf{L}_z^* z_b$, with the constraints that $\|\epsilon(0)\|$ is fixed and

$$\langle \epsilon(0), v_i \rangle = 0, \quad i = 1, \dots, M. \quad (2.21)$$

The regional singular vectors form a complete and orthogonal set. Therefore, we can expand $\epsilon(0)$ and the backward integrated blocking pattern b :

$$\begin{aligned} \epsilon^\perp(0) &= \sum_{i>M}^N \eta_i v_i, \\ b &= \sum_{j=1}^N b_j v_j. \end{aligned} \quad (2.22)$$

In this way we have directly fulfilled the constraint that $\varepsilon^\perp(0)$ is orthogonal to the first M singular vectors. Now,

$$\begin{aligned}
 \langle \varepsilon^\perp(0), b \rangle &= \langle \sum_{i>M}^N \eta_i v_i, \sum_{j=1}^N b_j v_j \rangle \\
 &= \langle \sum_{i=1}^N \eta_i v_i, \sum_{j>M}^N b_j v_j \rangle \\
 &= \langle \varepsilon(0), \mathbf{P}^\perp b \rangle,
 \end{aligned} \tag{2.23}$$

where \mathbf{P}^\perp is the projection on the subspace that is orthogonal to the vector space spanned by the first M regional singular vectors. The solution of Eq. (2.20) with constraints Eq. (2.21) is given by

$$\varepsilon^\perp(0) = \lambda \mathbf{P}^\perp \mathbf{R}^* \mathbf{L}_z^* z_b, \tag{2.24}$$

with scaling factor λ .

We have computed ε_{NL}^\perp with the constraint that ε_{NL}^\perp is orthogonal to the first 5 regional singular vectors ($M=5$). The optimization time T is again 6 days and the computations are performed for all days between days 701 and 800. The length of ε_{NL}^\perp is fixed and equal to the lengths of ε_L and ε_{NL} , which have been computed for the same period, see section 2.5. The results are shown in Fig. 2.7. One can see that $\Delta\mathcal{B}$, on average, is about as large as the change of \mathcal{B} as a result of the growth of the linearly optimal perturbations ε_L . In 56 cases (of the unperturbed non-blocked cases) ε_{NL}^\perp triggers a blocking onset. From this, we can conclude that high index regional singular vectors become important in the nonlinear range. They may even trigger a regime transition.

2.7 Concluding remarks

We have studied the sensitivity to initial conditions of the onset of blocking and strong zonal flow regimes in the Atlantic-European area. Both regimes are characterized by the same anomaly pattern but with opposite sign. Experiments have been performed with a three-level T21QG model with its tangent linear and adjoint versions. To classify the flow, the blocking index \mathcal{B} introduced by Liu (1994) has been used. Values of $\mathcal{B} > 0.5$ correspond to blocked flows, $\mathcal{B} < -0.5$ to strong zonal flows. Perturbations have been computed that maximize the difference in \mathcal{B} between the perturbed and reference forecasts; i.e. that maximize the change towards one of the regimes at a prescribed forecast time. These perturbations show the largest projection onto the anomaly pattern at optimization time. For short optimization times and perturbations with realistic amplitudes, error growth is almost linear. The sensitivity varies from day to day with only short periods of a high sensitivity. Results show that during these short periods, perturbations may initiate a regime transition. It might be useful to investigate the periods with high sensitivity in order to obtain some insight in the mechanisms that play a role in triggering a blocking onset or breakdown. For an optimization time of two or three days, the perturbations mainly originate from North America and are subsequently advected to the east. Thus, features of the jet stream like

its strength, geographical position and diffuent character are likely to be important. The sensitive periods may show some general properties that could possibly be revealed in this way. These issues will be made part of future studies.

For longer optimization times error growth is affected by nonlinear interactions. For the medium range, perturbations have been computed that maximize the difference in B using an iterative procedure. Experiments have been performed for an optimization time of 6 days with a time series of 100 days. For each day, the optimal perturbation was computed. It was found that almost all these perturbations (with realistic initial length) were able to trigger a blocking regime or a strong zonal flow regime. Furthermore, very large values of B could be obtained, which did not occur in the 1000 days reference orbit. So, it seems that these perturbations can induce extreme events. The special character of the perturbations can be caused because they do not lie, initially, in the hyper plane given by the local directions of the trajectories on the attractor. Another possibility is that the optimal perturbation lies on the attractor but in a region in phase space, given by all points that trigger extreme events, that is only a small part of the error region around the reference point. In the latter case a transition could be possible, but is still very unlikely because the probability that the initial condition lies in this particular region is very small. Because the iteratively computed perturbations maximize the same cost function as the linearly optimal perturbations, one may assume that there is no substantial difference between them. In turn, the linearly optimal perturbations are strongly related to the fastest-growing singular vectors, so that the special character could hold for the latter ones too. This is consistent with the results of Anderson (1994), who found that in the three-variable Lorenz convective model the fastest-growing singular vectors do not lie in the local tangent plane to the attractor.

Our approach of constructing perturbations, conceivably with additional constraints concerning above remarks, could be useful for ensemble forecasting. In the ensemble forecasting technique, a set of slightly different initial states is integrated in time (Lorenz 1965; Leith 1974). The divergence of the trajectories is taken as a measure of the predictability. In this way, the skill of the operational numerical weather prediction (NWP) models could be estimated in advance. It is, however, difficult to define a set of initial states that is not too large and at the same time gives statistically representative information about the divergence of the flow. Current methods to compute a set of initial states are making use of singular vectors (Mureau et al. 1993; Buizza et al. 1993) or bred perturbations (Toth and Kalnay 1993). Results of these methods show that the spread in the medium range is usually too small (ECMWF 1993). The iterative algorithm presented in this chapter can be useful to compose an ensemble that produces a significantly larger spread in the nonlinear range.

Moreover, one of the main objectives of ensemble forecasting is to detect possible regime transitions that are not forecasted by a general circulation model (GCM). The GCM's are in many cases not able to predict regime transitions beyond a few days correctly (Tibaldi and Molteni 1990). An important reason for this inability is that the onset of a regime is usually very sensitive to the initial conditions (as is shown in this chapter for a blocking and strong zonal flow regime). An ensemble prediction system (EPS) should warn for a possible regime transition. However, none of the present ensemble prediction systems specifically selects initial perturbations that trigger the onset of a particular weather regime. An alternative set of initial perturbations for an EPS can be computed by a generalization of our method.

In the same way as for the blocking anomaly pattern, perturbations can be computed that maximize the projection onto the anomaly patterns of other regimes or flow patterns. A possible choice of patterns are the first N empirical orthogonal functions (EOF's), which explain most of the variance of the flow. Then, the ensemble of initial perturbations consists of perturbations that maximize the projection on these EOF's at a prescribed forecast time. Thus, the sensitivity of an evolution towards specific regimes or patterns can be determined more directly.

Our results for the medium range indicated that a large spread could be produced for almost every initial condition. When such large spreads can be produced in more realistic models, one must doubt the usefulness of a predictability estimate obtained with a small ensemble. By this we mean that almost all weather scenarios (i.e. Großwetterlage patterns) may occur in the medium range for initial perturbations comparable to current analysis errors. In section 2.6, we have shown that linearly slow-growing perturbations can have a large impact on the flow in the medium range. These perturbations may be of equal importance, or even more important for nonlinear error growth than the linearly fastest-growing perturbations. The growth of the latter perturbations is usually strongly reduced by nonlinear interactions. So, in order to obtain a representative probability distribution of the occurrence of certain weather scenarios linearly fastest-growing as well as linearly slower-growing perturbations must be taken into account in the EPS. It is our intention to validate the above method and remarks for larger models.

Acknowledgements. Many thanks are due to the members of the predictability section at KNMI for the stimulating discussions and their assistance during the preparation of this chapter. This work has been supported by the foundation for Geological, Oceanographic and Atmospheric Research (GOA) of the Netherlands Organisation for Scientific Research (NWO). Computing facilities are provided by the Royal Netherlands Meteorological Institute.

Chapter 3

Onset of blocking and strong zonal flow regimes

Flows with high and low sensitivity with respect to the initial conditions for onset of blocking (BL) and strong zonal flow (SZF) regimes have been analysed. We have considered BL and SZF regimes at 20° W (Atlantic region) and at 150° W (Pacific region). The BL and SZF regimes are characterized by the same dipolelike anomaly pattern but with opposite signs. Experiments have been performed with a three-level quasigeostrophic model triangularly truncated at wavenumber 21 (T21QG), and its tangent linear and adjoint versions. The sensitivity is calculated by perturbing the reference flow with perturbations that optimally trigger the onset of a BL or SZF regime after a prescribed forecast time. For forecast times larger than 3 days an iterative technique is used to take into account nonlinear growth of the perturbations.

The flows with a high sensitivity show an intensified jet stream to the west of a diffluent flow. The strong jet stream by itself results in large perturbation growth. The presence of a diffluent flow amplifies the growth and gives perturbations a typical dipolelike character. Idealized experiments with a barotropic T21 model confirm that these properties increase the sensitivity. Sensitive flows are also characterized by an enhanced ridge upstream of the intensified jet stream. This does not directly influence the sensitivity but is associated with an intensified jet stream. The flow patterns of periods with low sensitivity are more zonal and weaker.

The diffuence of the flow also results in an asymmetry between sensitivity for BL and SZF onset in the medium range. Nonlinear feedback mechanisms increase sensitivity toward BL and decrease sensitivity toward SZF. Finally, it is shown that, on average, a transition toward BL corresponds with a larger than average sensitivity and that the sensitivity is larger when the transition is stronger. Transitions toward an SZF correspond, on average, with an average sensitivity independent of the strength of the transition. The precursor patterns of transitions toward BL have similar characteristics as the flows with a high sensitivity. Thus, blocking onset is likely to be in many cases an inherently sensitive phenomenon.

3.1 Introduction

Extratropical atmospheric dynamics is dominated by a chaotic succession of transient synoptic weather systems. Still, periods of quasi-stationary behavior in the large scales can be observed. Such persistent quasi-stable flow configurations are called weather regimes (Reinhold and Pierrehumbert 1982). In such weather regimes, the quasi equilibrium of the larger scales is a result of a subtle balance between the planetary scales and the synoptic scales, which may be disrupted by a single disturbance. No dynamically preferred time scale seems evident (Reinhold 1987). The regimes may persist for several days or even for several weeks. Onset and, to a lesser degree, also breakdown of regimes are usually very rapid and difficult to predict. The development and transitions of regimes mainly seem to be due to internal chaotic dynamics. Anomalous external forcing, like sea surface temperature anomalies (Hoskins and Karoly 1981), may cause a shift in the preference of particular weather regime events to occur (Palmer 1993). Also particular phases of low-frequency oscillations may favor the occurrence of weather regimes (Plaut and Vautard 1994).

Despite the random time-dependent character it seems that still some regularity can be found in the flow patterns that precede and succeed weather regime events (Dole 1986; Mo and Ghil 1987, 1988; Vautard 1990; Plaut and Vautard 1994). Vautard identified, using observations of 37 winters, preferred precursor and successor patterns of a set of four Euro-Atlantic regimes. Among these are a blocking (BL) and strong zonal flow (SZF) regime. A striking result is that of all regimes, the BL precursors are most successfully classifiable, whereas the prediction of blocking appears to be difficult. Studies on the predictability of blocking with the European Centre for Medium-Range Weather Forecasts (ECMWF) operational model (Tibaldi and Molteni 1990; Palmer et al. 1990; Tibaldi et al. 1994) showed that blocking onset is almost consistently missed beyond day 3-4. Also, the ECMWF ensemble prediction system (EPS) underpredicts the ensemble spread during periods in which blocking regimes develop (Molteni et al. 1996). This inability is partly due to model errors, such as the loss of eddy kinetic energy during the course of the integration (improved in more recent versions of the ECMWF model), which results in an underestimation of blocking frequency. Another reason for the difficulty in predicting blocking is the inherently sensitive and explosive nature of blocking onset. The dynamics in such periods is very sensitive to small changes in the flow patterns.

In this chapter periods will be investigated that show high and low sensitivity for regime transitions. The regimes studied are Pacific and Atlantic BL and SZF regimes. Here, sensitivity is defined with respect to the initial conditions, for regime transitions after a prescribed forecast time. Optimization times of 3 and 5 days are used. We want to emphasize that the sensitivity does not indicate whether a regime transition will occur, but that it is a measure of the predictability of such an event. High sensitivity indicates a low predictability and low sensitivity a high predictability of the forecast evolution. The following questions will be addressed. Can we identify general properties of sensitive and insensitive circulation patterns? Can we understand why these flows are so sensitive or insensitive? Are these flows (in)sensitive only for specific regime transitions or are they (in)sensitive for all kinds of perturbation growth? Can we relate sensitive circulation patterns to precursor patterns of transitions? Why is the onset of blocking so difficult to predict?

Experiments are performed with a three-level quasigeostrophic model triangularly truncated at wavenumber 21 (T21QG), described in Marshall and Molteni (1993), and its tangent linear and adjoint model. The tangent linear model describes the linear evolution of (small) perturbations along a reference forecast (updated every 40 min). Using an adjoint model the gradient of an output parameter at a specific optimization time with respect to all input parameters at initial time can be efficiently computed (Le Dimet and Talagrand 1986). In chapter 2 (Oortwijn and Barkmeijer 1995), a method is described that selects initial perturbations that optimally trigger the onset of a BL or an SZF regime after a prescribed optimization time. This method will be used and allows us to analyse a large number of onsets at relatively low numerical costs.

In section 3.2 we will briefly describe the performance of the T21QG model in simulating blocking and strong zonal flow regimes, and an index is given to determine whether a circulation pattern is blocked, zonal, or in some intermediate state. The method to select initial perturbations that optimally trigger the onset of a BL or SZF regime is described. The method also contains a technique to take into account nonlinear growth of perturbations, so that sensitivity for the medium range can be calculated. In section 3.3, results of three model experiments are presented. Two experiments are performed for the short range for the Atlantic and Pacific regions. The third experiment is done for the medium range and the Atlantic region. Mean (anomaly) patterns of sensitive and insensitive flows are identified. In section 3.4 we will discuss features of (in)sensitive flows and analyse results with a barotropic model. Furthermore, the relation between sensitive flows and precursors of transitions is studied. Finally, in section 3.5 we give a summary and some concluding remarks.

3.2 Blocking and strong zonal flow regimes

3.2.1 Model performance

Experiments are performed with a three-level quasigeostrophic model triangularly truncated at wavenumber 21 (T21QG). The levels of the model are at 200, 500, and 800 hPa. The T21QG model is described in detail in Marshall and Molteni (1993). Liu and Opsteegh (1995) have made a comparison between model simulations and observations. Figure 3.1a shows the mean 500-hPa geopotential height field for the Northern Hemisphere, calculated from a 50000-day perpetual winter integration. The mean field resembles the observed winter mean 500-hPa geopotential height field quite well. Figure 3.1b shows the simulated standard deviation of the daily 500-hPa fields. The global structure of the variance fields is also in general agreement with the observations. The model is able to simulate the maxima over the Pacific and Atlantic oceans, which are preferred regions for blocking. The structure of the blocking pattern, its effect on transient weather systems, the number of blocking days, and the interannual variability in blocking activity are all similar to the observations. However, in the model, the duration of a blocking regime is, on average, one day shorter. The same conclusions can be made for the model statistics of strong zonal flow regimes. For a more extensive description of the model performance (with respect to blocking and strong zonal flow regimes) we refer to Liu and Opsteegh (1995).

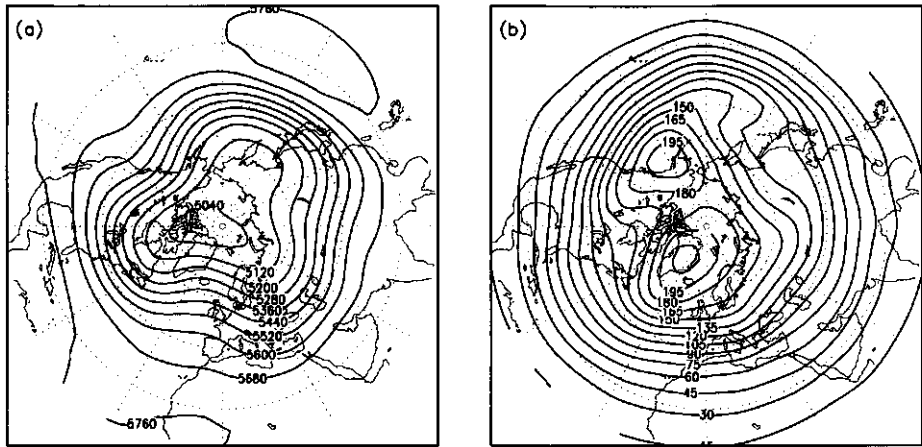


Figure 3.1: (a) Mean and (b) standard deviation of 500-hPa geopotential height field. Contour interval is 80 m in (a) and 15 m in (b).

3.2.2 Characterization of the regimes

Liu (1994) characterized a Euro-Atlantic BL regime by a dipolelike pattern, consisting of a very strong positive geopotential height anomaly with its center at about 60° N and a weaker negative anomaly to the south of it. This pattern was calculated by applying a blocking criterion to a 10-winter season data set of observations. By applying the same blocking criterion to a data set of the T21QG model, Liu and Opsteegh (1995) found a similar pattern. Furthermore, it was found that an SZF regime is characterized by the same pattern with opposite sign. This anomaly pattern, which is called the blocking geopotential height anomaly pattern z_b , is shown in Fig. 3.2a. In order to study Pacific blocking and strong zonal flow regimes, we have shifted the blocking anomaly pattern to the Pacific area (Fig. 3.2b). In section 3.3.2 we will describe this latter approach more extensively.

A blocking index, \mathcal{B} , can now be defined that measures the resemblance of a particular circulation pattern with the BL or SZF regimes. The blocking index \mathcal{B} for a particular circulation pattern with streamfunction ψ is defined by the projection of the daily geopotential height anomaly pattern z_d on the blocking geopotential height anomaly pattern z_b , weighted with the norm of z_b :

$$\mathcal{B}(\psi) = \frac{\langle z_d(\psi), z_b \rangle}{\langle z_b, z_b \rangle}. \quad (3.1)$$

Here the brackets denote a squared norm inner product on a sphere, S , integrated over height

$$\langle x, y \rangle = \frac{1}{4\pi} \int \int \int x y dS dh. \quad (3.2)$$

Here, the integration over height is performed by a summation over the three levels. The relation between z_d and ψ is given by

$$z_d(\psi) = \mathbf{L}_z \psi - z_c, \quad (3.3)$$

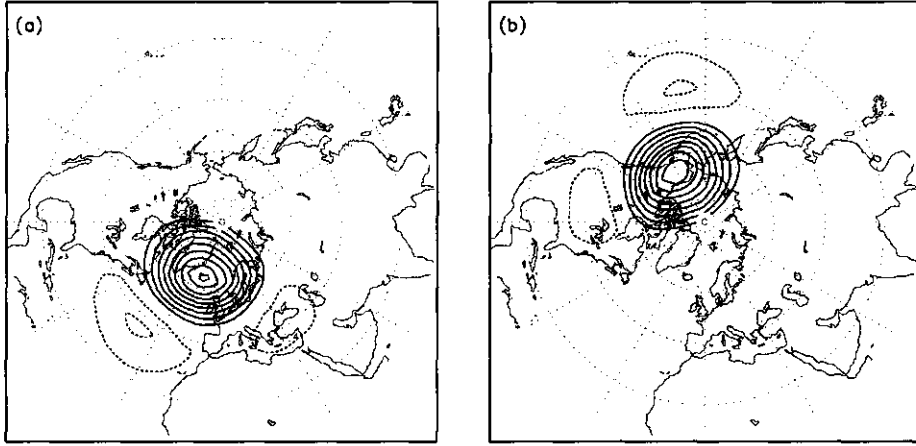


Figure 3.2: Blocking anomaly pattern: mean 500-hPa geopotential height anomaly pattern. (b) Pacific blocking anomaly pattern (rotated). The contour interval is 20 m. Solid lines correspond to positive values, dotted lines to negative values.

where L_z is the linear operator that transforms streamfunction into geopotential height and z_c is the simulated climatological mean geopotential height. The geopotential height field is obtained from the streamfunction by solving the linear balance equation. A circulation pattern with $B > 0.5$ ($B < -0.5$) is considered as a blocked flow (strong zonal flow). Furthermore, a larger value of B indicates a more pronounced blocked flow, and a smaller negative value of B indicates a stronger zonal flow.

3.2.3 Inducing transitions

The intent here is to obtain some insight in the sensitivity for the onset of BL and SZF regimes to perturbations in the initial conditions. In chapter 2 (Oortwijn and Barkmeijer 1995) a method is described to obtain the initial perturbation that maximizes the difference in B between reference and perturbed flow after a prescribed forecast time:

$$\Delta B = B(\psi(T) + \varepsilon(T)) - B(\psi(T)) \quad (3.4)$$

at prescribed forecast time T , keeping

$$\|\varepsilon(0)\|^2 = \langle \varepsilon(0), \varepsilon(0) \rangle \quad (3.5)$$

fixed. Using the definition of B (Eq. (3.1)), one can show that, for the limit of small perturbations, maximizing Eq. (3.4) is equivalent to maximizing

$$\langle L_z \varepsilon(T), z_b \rangle = \langle L_z R \varepsilon(0), z_b \rangle = \langle \varepsilon(0), R^* L_z^* z_b \rangle, \quad (3.6)$$

where we introduced the tangent linear operator R , which describes the linear evolution of a streamfunction perturbation. The operators R^* and L_z^* are the adjoint operators of R and

\mathbf{L}_z , respectively, with respect to the squared norm inner product. It follows that the initial perturbation that maximizes the difference in \mathcal{B} at optimization time T is given by

$$\varepsilon_L(0) = \lambda \mathbf{R}^* \mathbf{L}_z^* z_b. \quad (3.7)$$

Here, the subscript L denotes that the perturbation is optimal when the perturbation growth is linear. The absolute value of the scaling factor λ is fixed through constraint Eq. (3.5). For positive λ , the perturbation ε_L maximizes the change toward a blocking regime. A negative value of λ results in a maximum change toward a strong zonal flow regime. The difference in \mathcal{B} at optimization time between the reference and perturbed forecast for this optimal perturbation is

$$\Delta \mathcal{B} = \frac{||\mathbf{R}^* \mathbf{L}_z^* z_b||}{||z_b||^2} ||\varepsilon_L(0)||. \quad (3.8)$$

The linear sensitivity is now defined by

$$S_L = \frac{||\mathbf{R}^* \mathbf{L}_z^* z_b||}{||z_b||^2}. \quad (3.9)$$

In chapter 2 is shown that for periods with a high sensitivity S_L for a forecast period of 3 days, a small perturbation $\varepsilon_L(0)$ can lead within just 3 days to a regime transition. For longer optimization periods, the assumption of linearity is not valid anymore. Also an iterative technique is described to take into account nonlinear perturbation growth (see section 2.5). With this procedure, ε_L is modified to $\varepsilon_{NL,BL}$ such that the perturbed flow makes a transition toward a BL regime. In the same way, $\varepsilon_{NL,SZF}$ induces an SZF regime for the perturbed flow in the medium range. In contrast to the linear case, $\varepsilon_{NL,BL}$ and $\varepsilon_{NL,SZF}$ have a different structure. In chapter 2 is shown that for almost all circulation patterns small initial perturbations can be found such that the perturbed flow makes a regime transition, toward BL as well as SZF regimes, within 6 days. These methods, both for the short range and the medium range, will be used to compute the sensitivity of flows. Results are presented in section 3.3.

In section 3.4, we will investigate if periods with a high or low sensitivity have specific perturbation growth dynamics. Perturbations optimized to trigger regime transitions will be compared with perturbations that maximize growth in a particular region. Therefore, regional singular vectors (RSVs) will be calculated. These RSVs maximize the rms error over the Atlantic-European area; that is, they maximize the following Euclidean inner product:

$$(\mathbf{P}\varepsilon(T), \mathbf{P}\varepsilon(T)) = (\mathbf{R}^* \mathbf{P}^* \mathbf{P} \mathbf{R} \varepsilon(0), \varepsilon(0)) \quad (3.10)$$

Here \mathbf{P} is a linear projection operator, which acts on a global streamfunction field $\varepsilon(T)$ and projects onto the Atlantic-European area (10° - 85° N, 90° W- 60° E) at the 500-hPa level. The operators \mathbf{R}^* and \mathbf{P}^* are the adjoint operators of \mathbf{R} and \mathbf{P} with respect to the Euclidean inner product. The RSVs are the eigenvectors of $\mathbf{R}^* \mathbf{P}^* \mathbf{P} \mathbf{R}$. Employing a Lanczos algorithm, successive approximations for the eigenvalues and eigenvectors of this symmetric matrix operator are found iteratively (Parlett 1980).

3.3 Identification of sensitivity patterns

In this section results are presented of three model experiments. For each experiment a long integration of 20000 days with the T21QG model is performed. In the first experiment the linear sensitivity S_L for a forecast period of 3 days is calculated each day. In the second experiment, the same is done with respect to transitions to BL or SZF in the Pacific region. In the third experiment the sensitivity during a forecast period of 5 days is calculated each day (Euro-Atlantic region only). Next, circulation patterns with high or low sensitivity are selected. Here the term sensitivity refers to sensitivity for regime onsets with respect to the initial conditions as calculated with the methods described in the previous section. It appears that the sensitive circulation patterns as well as the insensitive circulation patterns have common properties. Anomaly patterns will be identified that are representative for sensitive and insensitive flows. These anomaly patterns will be referred to as sensitivity patterns.

3.3.1 Euro-Atlantic region, short range

The distribution of the linear sensitivity for a forecast period of 3 days is given in Fig. 3.3. The mean linear sensitivity \bar{S}_L equals $2.69 \times 10^{-6} \text{ s m}^{-2}$. The standard deviation of the distribution is $0.68 \times 10^{-6} \text{ s m}^{-2}$. The minimum value of the linear sensitivity S_L is $1.23 \times 10^{-6} \text{ s m}^{-2}$, whereas the maximum value equals $7.98 \times 10^{-6} \text{ s m}^{-2}$. Hence the maximum difference in ΔB between reference and perturbed flow after 3 days may be as large as a factor of 6.4 between low- and high-sensitivity periods. In many high-sensitivity cases transitions can be triggered by adding or subtracting the optimal perturbation to the initial conditions with a realistic amplitude. From the distribution, the 1000 most sensitive and 1000 most insensitive 3-day periods are selected. Each period is characterized by the flow at initial time. The sensitive periods all have $S_L > 3.94 \times 10^{-6} \text{ s m}^{-2}$ and the insensitive days $S_L < 1.79 \times 10^{-6} \text{ s m}^{-2}$.

In Figs. 3.4a and 3.4b the mean fields of the high- and low-sensitivity sets are shown. In Figs. 3.4c and 3.4d the anomalies with respect to the climatological mean (Fig. 3.1a) are shown. The zonal wavenumber 3 structure has a large amplitude for the high-sensitivity patterns and a small amplitude for the low-sensitivity patterns. For the high-sensitivity patterns the ridge in the eastern Pacific is enhanced and the trough above the North American continent has deepened. This trough, together with the positive anomaly to its southeast, intensifies the jet stream. The low-sensitivity flows are characterized by the opposite anomaly patterns; the flow is more zonal and the jet stream is weaker. The anomaly patterns resemble a PNA-like structure. The vertical structure of the sensitivity patterns (not shown) is equivalent barotropic, with largest values of the anomalies at the 200-hPa level and smallest at the 800-hPa level. In Figs. 3.4e and 3.4f the same plots as in Figs. 3.4c and 3.4d are given, but now for the 100 most sensitive and insensitive patterns. The anomalies are even larger. Note that also the positive anomaly over Europe is larger for the high-sensitivity set, which enhances the ridge over Europe. The negative anomaly over Europe in the low-sensitivity set is even larger and causes the ridge over Europe to disappear completely.

There is a clear signal in the sensitivity patterns above the North American continent and

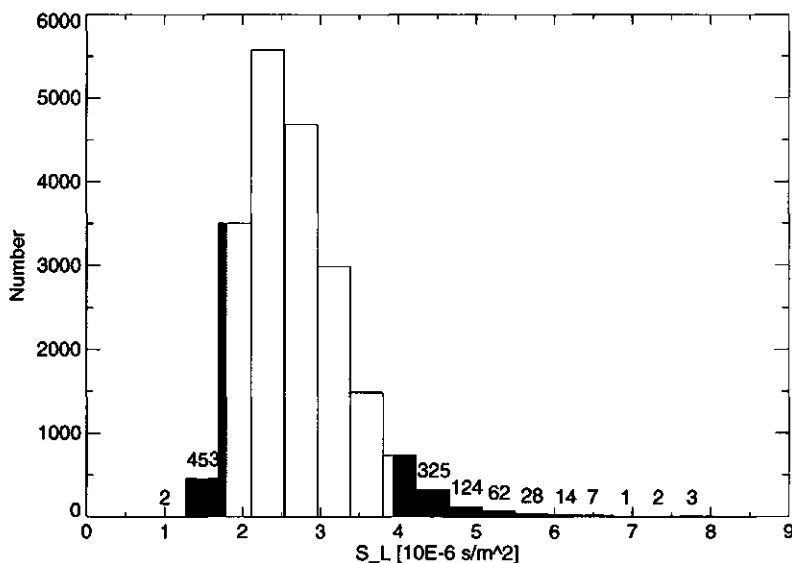


Figure 3.3: Distribution of S_L taken from a set of 20000 periods. Shaded are the low- and high-sensitivity sets, each consisting of 1000 periods.

North-East Pacific. This region is upstream of the target area (Euro-Atlantic region) and the initial optimal perturbation has much of its amplitude here. The signal in the target area is much weaker in the mean patterns. Does this imply that the local circulation in the target area has no influence on the sensitivity? Or is it still possible to identify differences between high- and low-sensitivity flows that are not revealed in the mean sensitivity patterns?

First of all, we have compared the variance fields of the two sets (not shown) with the climatological variance (Fig. 3.1b). The variance in the high-sensitivity set is somewhat less over Europe (standard deviation maximum 195 m), about the same over the Pacific, and less over the North American continent (less variance in the deepness of the trough over North America). For the low-sensitivity set the variance over the Atlantic is larger than for the climatological variance (the maximum of the standard deviation is more than 225 m), whereas over the Pacific area the standard deviation is much less (maximum about 165 m). Thus, the individual flow patterns of the low-sensitivity set may differ strongly over the target area. The differences are smaller but still substantial for the high-sensitivity set.

In order to reveal the differences between flow patterns in both sets, a cluster algorithm was applied as described in Michelangeli et al. (1995). The clustering is done for North America and the Euro-Atlantic region (130°W-40°E, 30°-70°N). For each of the 2000 flow patterns the anomaly field with respect to the climatological mean for this region is calculated. The anomaly patterns are projected in a five-dimensional space by projecting the patterns onto the first five empirical orthogonal functions (EOFs). The EOFs are computed from a 30000-day run of the T21QG model. The cluster analysis localizes high concentrations of points, the clusters, in the five-dimensional space. The iterative algorithm computes

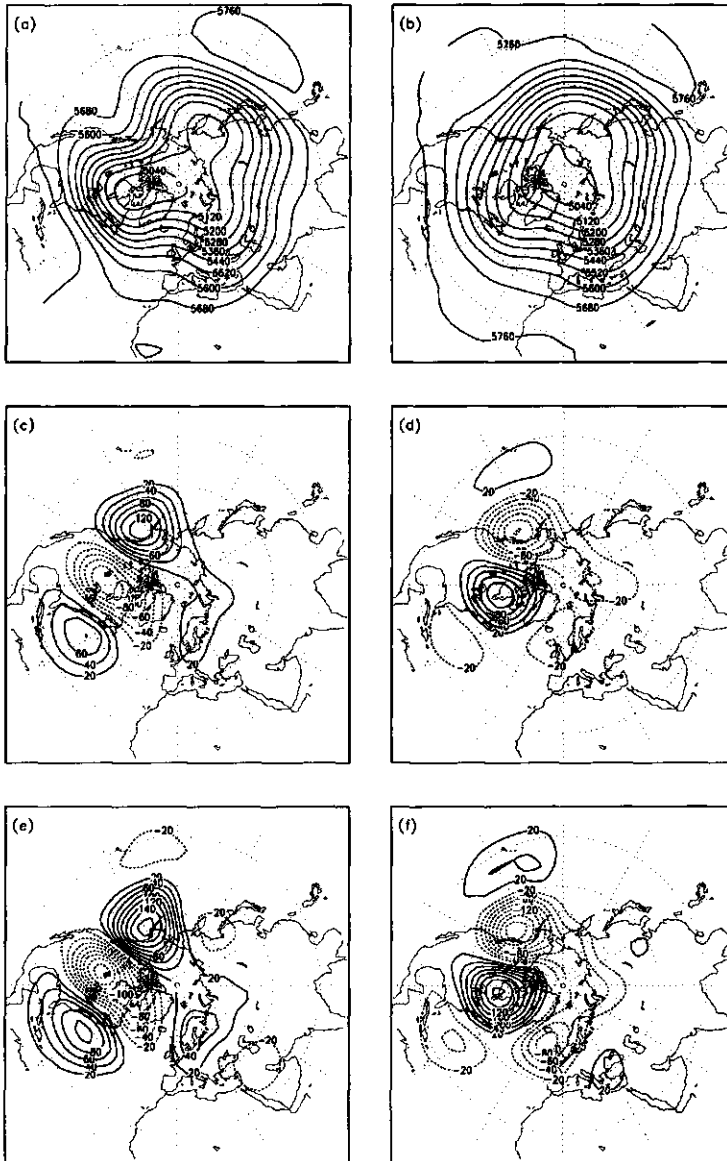


Figure 3.4: (a) Mean geopotential height field at 500 hPa of the high sensitive set and (b) of the low-sensitivity set (1000 patterns). (c) and (d): As (a) and (b) but for the anomaly fields. (e),(f): Anomalies of 100 most (in)sensitive patterns. Contour intervals are given in meters.

the best partition of the points for a prescribed number of clusters, k , with k between 2 and 10. Cluster means, which are the mean patterns of each cluster, are used to characterize the cluster. The best partition minimizes the sum of variances within the clusters. To test the robustness of the method, classifiability and reproducibility tests are performed. A high classifiability indicates that the final partition is nearly independent of the initial starting points. Reproducibility is tested by splitting the data in two samples and recomputing the cluster means.

For the low-sensitivity set we find four clusters. The cluster means, shown in Figs. 3.5a-d, are well classifiable and reproducible. Clusters 1 and 3 are zonal regimes, cluster 2 corresponds to a Greenland anticyclone with a very weak circulation, and cluster 4 with a ridge over England. The number of patterns that correspond to each cluster are given in the panels. The optimal number of clusters for the high-sensitivity set is less clear. The best number seems to be for only two clusters, which are well classifiable and reproducible. In Figs. 3.5e and 3.5f the high sensitivity cluster means are given. The difference between the means is that in the first cluster the ridge over the Euro-Atlantic region is shifted more to the east, whereas in the second cluster the ridge is shifted more to the west. When this set is divided into three or four clusters the differences are small. The ridges are more pronounced and one cluster is found with a weak zonal flow over Europe but that is still slightly diffluent. So, most of the high-sensitivity patterns are characterized by a ridge, which involves diffluent flow, over the Euro-Atlantic region. The phase of this ridge shows large variability, which causes a cancellation of its signal in the mean sensitivity patterns. The low-sensitivity patterns show all kinds of circulation patterns in the target area, from strong zonal flows to blocking highs.

3.3.2 Pacific region, short range

In order to find sensitivity patterns for Pacific BL or SZF flows, a similar experiment as the previous one has been performed. As anomaly pattern the Atlantic blocking anomaly pattern (Fig. 3.2a) is used but is then shifted to the Pacific area (Fig. 3.2b). A study by Tibaldi et al. (1994) about blocking frequency and predictability, using ECMWF analyses and forecasts of 500-hPa geopotential height fields, showed that Pacific blocking has a similar dipolelike anomaly field (see their Fig. 3e). The Pacific blocking anomaly pattern (Fig. 3.6b) also resembles quite well the Alaska blocking pattern given in Hsu and Wallace (1985, their Fig. 2) and Nakamura and Wallace (1990). The exact amplitude of the blocking anomaly pattern is not essential for this study, because we only compute the projection of the optimal perturbation on this pattern.

Again, the T21QG model is integrated for 20000 days, and for each day the linear sensitivity for a forecast period of 3 days is calculated. The mean linear sensitivity $\overline{S_L}$ equals $3.31 \times 10^{-6} \text{ s m}^{-2}$ and the standard deviation of the distribution is $0.95 \times 10^{-7} \text{ s m}^{-2}$. Both the mean value and standard deviation are somewhat larger for the Pacific region than for the Euro-Atlantic region, indicating a slightly higher sensitivity in the Pacific region. The minimum value of the linear sensitivity S_L is $1.40 \times 10^{-6} \text{ s m}^{-2}$, whereas the maximum value equals $11.0 \times 10^{-6} \text{ s m}^{-2}$.

In Figs. 3.6a and 3.6b the mean geopotential height fields at 500 hPa of the 1000 most

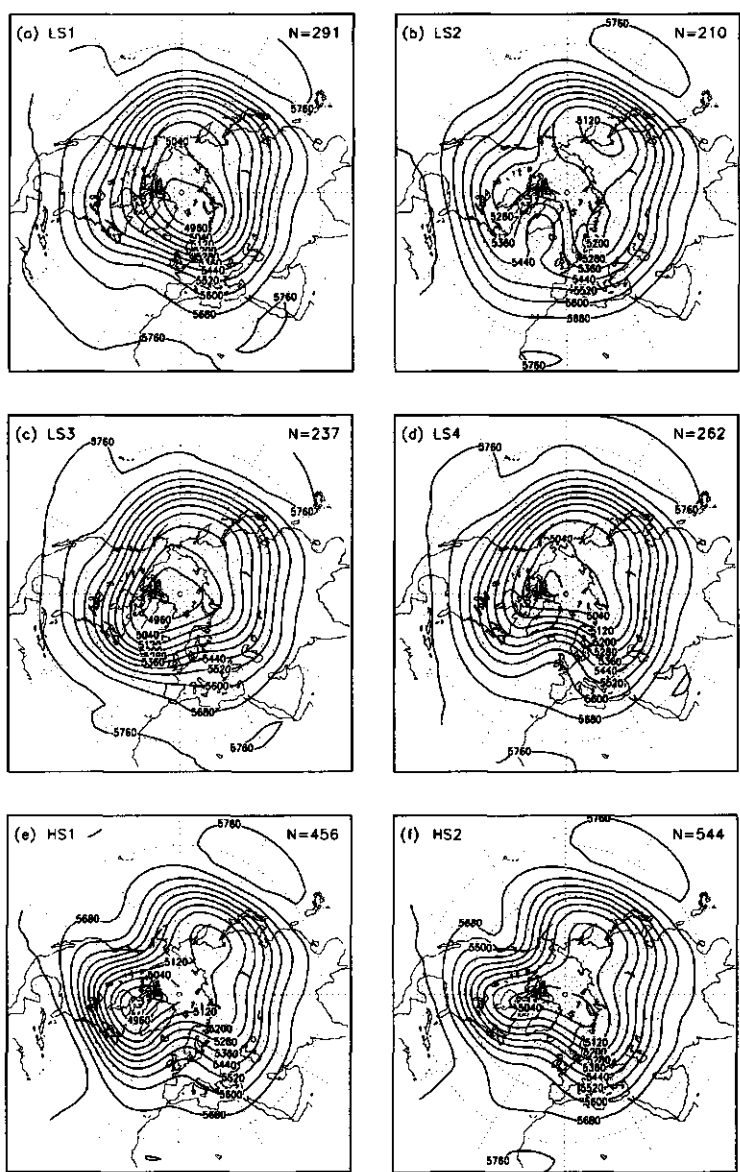


Figure 3.5: Cluster means of low-sensitivity (LS) and high-sensitivity (HS) set. (a) LS1, (b) LS2, (c) LS3, (d) LS4, (e) HS1, and (f) HS2. Shown is the geopotential height field at 500 hPa. Contour interval is 80 m.

sensitive and the 1000 most insensitive initial flow patterns are shown, respectively. In Figs. 3.6c and 3.6d the anomalies with respect to the climate mean geopotential height field are given for the 500-hPa level. The sensitive flows have, as in the former case, an intensified jet stream, but now it is over the western Pacific. The insensitive flows are more zonal and can be characterized by the opposite anomaly patterns. The intensification (weakening) of the wave with zonal wavenumber $m = 3$ in the high-sensitivity (low-sensitivity) patterns is clearly visible. In comparison to the Atlantic sensitivity patterns, there is slightly more amplitude in the target area and less amplitude far upstream. The high-sensitivity periods are characterized by enhancement and the low-sensitivity periods by weakening of the ridge in the target area. Possibly because of the geographical differences (like the Rocky Mountains, which interact strongly with the flow) between the Pacific and Atlantic regions, the phase of the Pacific ridge shows less variability. The signal in the mean sensitivity patterns is therefore larger in the Pacific region. A cluster analysis technique is not necessary to reveal the differences between low- and high-sensitivity flows in the target area. We conclude that the sensitivity for Pacific BL or SZF onset is mainly influenced by the same features as sensitivity for Atlantic BL or SZF onset.

3.3.3 Euro-Atlantic region, medium range

The third experiment concerns medium-range sensitivity for the Euro-Atlantic region. For 20000 days, perturbations ε_L , $\varepsilon_{NL,BL}$, and $\varepsilon_{NL,SZF}$ have been computed for a 5-day optimization period. The perturbation ε_L is the linearly optimal perturbation. Modifications of ε_L to take nonlinear growth into account are denoted by $\varepsilon_{NL,BL}$ and $\varepsilon_{NL,SZF}$ [see sections 2.5, also 3.2.3]. The perturbation $\varepsilon_{NL,BL}$ is maximized to induce a blocking regime, whereas $\varepsilon_{NL,SZF}$ is maximized to induce a strong zonal flow at day 5. On average, four iterations are adequate in order to compute $\varepsilon_{NL,BL}$ and $\varepsilon_{NL,SZF}$.

Three high- and low-sensitivity sets have been created, each set consisting of 1000 initial flow patterns. Integrating $\varepsilon_L(0)$ linearly along the reference orbit yields a value for the linearly expected difference ΔB . The first high- and low-sensitivity sets are obtained on the basis of this linear prediction. Another two pairs of high- and low-sensitivity sets are obtained by nonlinear integrations of $\varepsilon_{NL,BL}$ and $\varepsilon_{NL,SZF}$. Comparison of the three different sets may yield some knowledge about the validity of the linearity assumption in the calculation of the sensitivity patterns (as defined in section 3.3). But first, we will discuss results of the computations when the full nonlinear interactions are taken into account (the latter two pairs of sets).

Figures 3.7a-d show the mean geopotential height anomaly fields at 500 hPa that have been obtained from the nonlinear integrations of $\varepsilon_{NL,BL}$ and $\varepsilon_{NL,SZF}$. In Fig. 3.7a the mean anomaly pattern is given of the 1000 initial flow patterns that are most sensitive for a change toward a BL regime. Figures 3.7b-d show the mean anomaly patterns that are most insensitive toward a BL regime (Fig. 3.7b), most sensitive toward an SZF regime (Fig. 3.7c) and most insensitive toward an SZF regime (Fig. 3.7d). The difference in sensitivity to BL and SZF is due to the nonlinearity of the evolution of the perturbations. Figure 3.8a shows the difference between the mean patterns that are sensitive for changes to BL and SZF (Fig. 3.7a-c). In Fig. 3.8b the difference between Figs. 3.7b and 3.7d

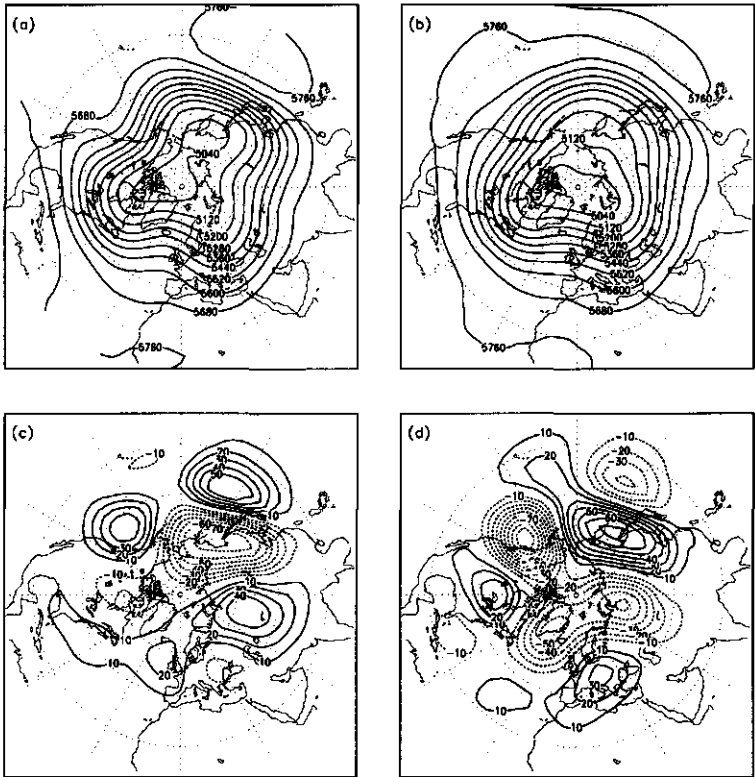


Figure 3.6: Mean geopotential height fields at 500 hPa for (a) high-sensitivity set and (b) low-sensitivity set (Pacific region). (c) and (d): As (a) and (b) but for the anomaly fields. Contour lines are in meters.

is plotted. Both patterns resemble the blocking anomaly pattern (Fig. 3.2) with negative and positive amplitude, respectively. In order to explain that the differences resemble the blocking anomaly pattern, one must realize that an initially blocked flow (positive \mathcal{B}) has, on average, also at optimization time a larger value of \mathcal{B} than a flow at optimization time, which at initial time is strong zonally. Here, the periods that are most sensitive toward a BL regime ($\Delta\mathcal{B} > 2.4$) have at optimization time, on average, a value of $\mathcal{B} = -0.8$. The periods that are most sensitive toward an SZF regime ($\Delta\mathcal{B} < -2.2$) have at optimization time, on average, a value of $\mathcal{B} = 0.4$. Next, one may argue that when the blocking index of a flow is already positive (negative) at optimization time, the flow can be more easily perturbed toward a lower (higher) value of the blocking index. (The values of \mathcal{B} of a data set obtained by integrating the T21QG model vary within a range of about -2.5 to 2.5; only rarely does \mathcal{B} exceed this range.) Thus, a flow that is initially blocked can more easily be perturbed toward a zonal flow than toward an even stronger blocked flow (and vice versa). The opposite is the case for the low-sensitivity patterns. If a flow is already blocked it is more insensitive for

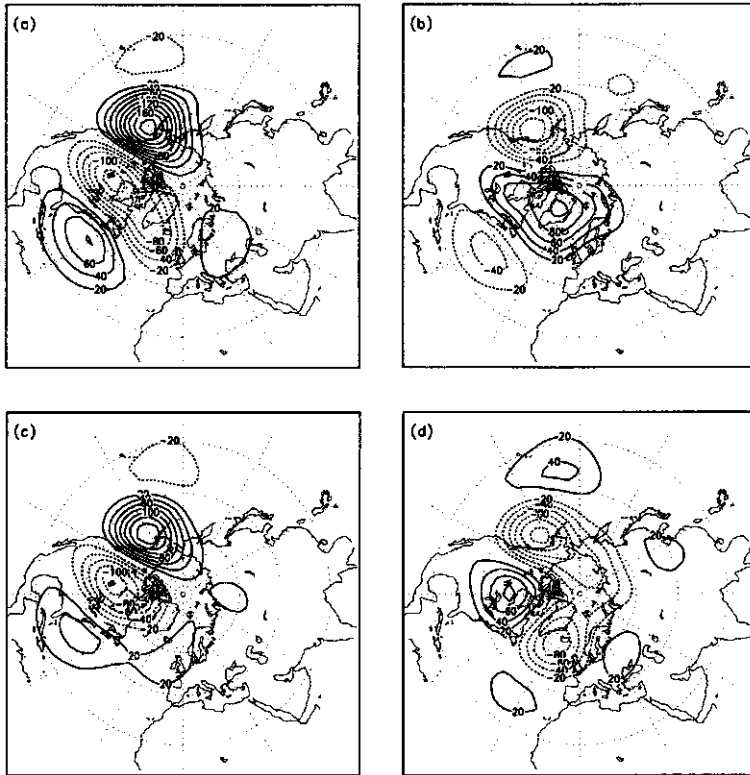


Figure 3.7: Sensitivity patterns, in geopotential height at 500 hPa, for the 5 day period. (a) High sensitivity to BL, (b) low sensitivity to BL, (c) high sensitivity to SZF, and (d) low sensitivity to SZF. Contour interval is 20 m.

changes toward a blocked flow than toward a strong zonal flow. The nonlinear terms cause, on average, a tendency of the flow to return to the climate mean.

In Fig. 3.8c is shown the mean geopotential height anomaly field of the 500 most sensitive initial flow patterns toward BL and the 500 most sensitive toward SZF (thus, the anomaly field is again the mean of 1000 patterns as in Figs. 3.4c and 3.4d). In Fig. 3.8d this field is shown for the 500 most insensitive initial flow patterns toward BL and the 500 most insensitive toward SZF. So Figs. 3.8c and 3.8d show the initial anomaly patterns for sensitive and insensitive 5-day periods, with respect to transitions toward BL and SZF. Again, larger amplitudes of these anomaly patterns are related to a higher (in)sensitivity. Nearly the same sensitivity patterns (not shown) are obtained as in Figs. 3.8c and 3.8d by taking into account only linear interactions, that is, by computing linear integrations of ϵ_L . We conclude that linear theory still enables one to determine the sensitivity of circulation patterns for a 5-day forecast period (at least in this model), although nonlinear perturbation growth is nonnegligible (see also chapter 2).

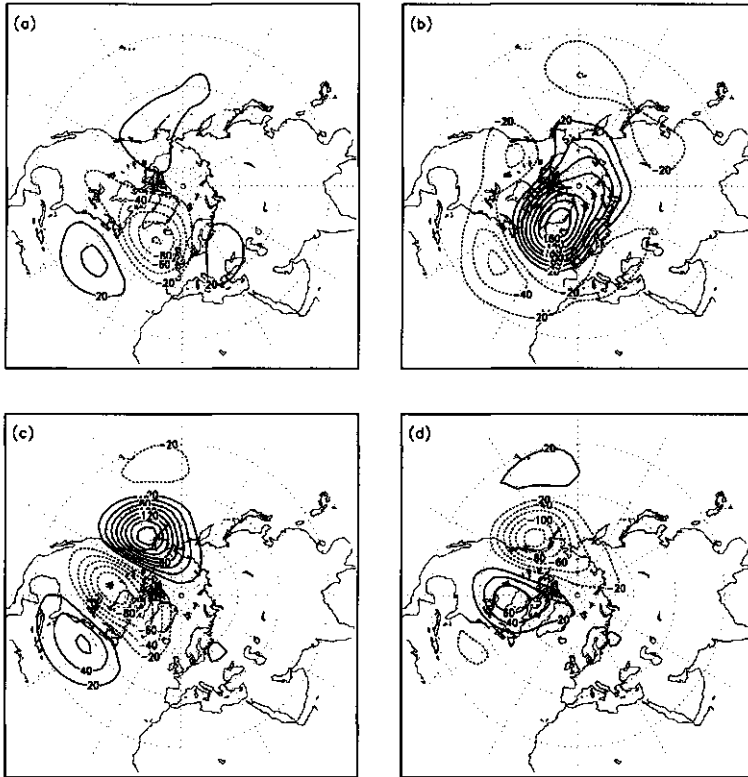


Figure 3.8: (a) High sensitivity to BL - high sensitivity to SZF, (b) low sensitivity to BL - low sensitivity to SZF. (c) High-sensitivity patterns for sensitivity to BL or SZF and (d) low-sensitivity patterns for sensitivity to BL or SZF. Contour interval is 20 m.

Comparing Figs. 3.8c and 3.8d with Figs. 3.4c and 3.4d, respectively, one notices that the sensitivity patterns are nearly equal; even the phases are the same. Only the amplitudes of the cells differ slightly and the low-amplitude structures over Europe in the 3-day sensitivity patterns have almost completely disappeared in the 5-day sensitivity patterns.

Why are the differences in the 3- and 5-day sensitivity patterns so small? One may expect that the initial perturbations for the 5-day periods are located farther upstream, resulting in a westward shift of the sensitivity patterns. In order to study the mean locations of the initial perturbations, 500 perturbations for high-sensitivity periods and 500 for low-sensitivity periods are calculated for 3-day and 5-day periods. Mean and standard deviations are calculated. Because of the small-scale structures that depend strongly on the reference flow, the mean patterns have very low amplitudes which are nearly insignificant. (Significance is tested by averaging different sets of initial perturbations.) Only for the mean initial perturbation patterns of the 3-day low-sensitivity periods can a significant pattern be identified.

In this case the initial perturbation tends to be a dipolelike pattern located upstream of the target area. For low-sensitivity periods, which are characterized by a weak zonal westerly flow, the initial dipolelike perturbations will be mainly advected to the target area. During this evolution the amplitude of the perturbation will amplify slowly.

Plots of the standard deviations for the initial perturbations of high-sensitivity 3- and 5-day periods are shown in Fig. 3.9. In Figs. 3.9a, 3.9b, and 3.9c the standard deviations in geopotential height at 200, 500 and 800 hPa, respectively, are given for the 3-day period. In Figs. 3.9d, 3.9e, and 3.9f the standard deviations are given for the 5-day period. Comparing the 3- and 5-day standard deviations, one can see that at all levels the maximum values are shifted to the west in the latter plots, so the initial perturbations for a 5-day period are located farther upstream. An explanation for the small differences between the 3- and 5-day sensitivity patterns could be that these special flow patterns are easily perturbed in the direction of blocking/strong zonal flow. The position and structure of these patterns are not affected by the length of the integration. However, for 5-day integrations it is necessary to start off initial perturbations farther upstream than for 3-day integrations in order to trigger the transitions. Notice also the westward tilt with height in the standard deviation, suggesting that baroclinic mechanisms play a role in the amplification of the perturbations.

3.4 Features of high-sensitivity and low-sensitivity flows

So far, we have identified mean anomaly patterns that give an indication of the sensitivity of a circulation pattern to make a transition toward a BL or SZF regime after 3 or 5 days. The difference between the mean sensitivity patterns for short and medium range was small. The sensitivity patterns consisted mainly of upstream wave trains of positive and negative anomalies (Figs. 3.4c, 3.4d, 3.6c, 3.6d, 3.8c, and 3.8d). It was also shown that the flow properties in the target area, that is, the area where the blocking anomaly pattern is located, contribute to the sensitivity, although the signal in the mean patterns is weak. Sensitive periods are characterized by an enhanced ridge in the target area. In this section we will address the following questions. What causes these flows to have a high or low sensitivity? Are these periods sensitive for regime transitions only, or are they sensitive for all kinds of perturbation growth? What are the basic mechanisms involved in the growth of the perturbations? Can the high-sensitivity patterns be related to precursor patterns for BL and SZF regimes? In order to answer these questions we have computed regional singular vectors (RSVs) for both sensitive and insensitive periods. The RSVs are computed for the Atlantic-European region (10° - 85° N, 90° W- 60° E). Also, a comparison with barotropic experiments will be made. Because of the small differences between the 3- and 5-day sensitivity patterns, experiments are performed for the short range only.

In Fig. 3.10 an example is given of a high-sensitivity period. Figure 3.10a shows the initial flow pattern in geopotential height at 500 hPa and Figs. 3.10b-f show the first five evolved regional singular vectors. The RSVs maximize the expression in Eq. (3.10) and are subsequently integrated linearly for 3 days to get the evolved patterns at optimization time. Figure 3.11 shows these patterns for a low-sensitivity period. The eigenvalues of the RSVs of the high-sensitivity period are much larger than those of the low-sensitivity period.

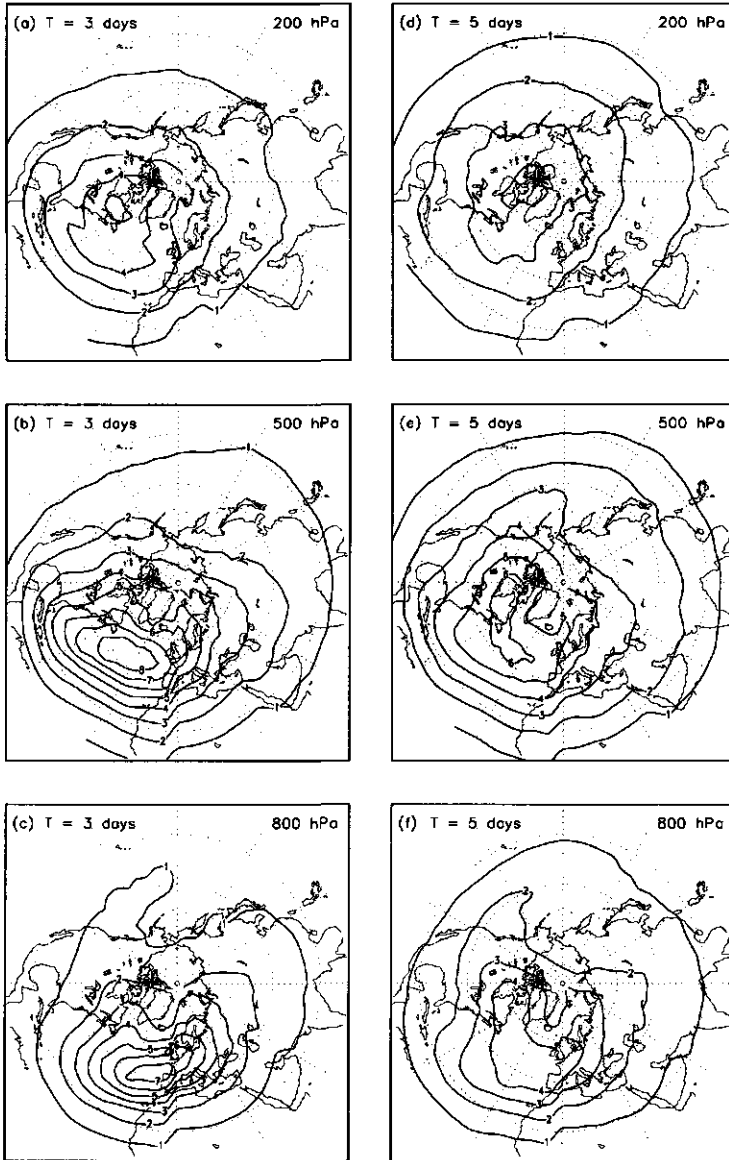


Figure 3.9: Standard deviations in geopotential height of $\varepsilon_L(0)$ ($T_{\text{opt}} = 3$ days) at (a) 200 hPa, (b) 500 hPa, and (c) 800 hPa; and of $\varepsilon_{NL}(0)$ ($T_{\text{opt}} = 5$ days) at (d) 200 hPa, (e) 500 hPa, and (f) 800 hPa.

Furthermore, the structure of the RSV patterns are different. The evolved RSVs in the low sensitivity period consist of wave trains aligned zonally, whereas the evolved RSVs (mainly RSV1, RSV2, and RSV5) have more meridional structure in the high-sensitivity period. Especially the first evolved RSV has a large projection onto the blocking anomaly pattern (Fig. 3.2a). To measure the projection of the i th evolved RSV v_i onto the blocking anomaly pattern z_b , we define the (normalized) projection coefficient p_i by

$$p_i = \frac{|\langle v_i, z_b \rangle|}{\|v_i\| \|z_b\|}, \quad (3.11)$$

so that p_i varies between 0 and 1. The projection coefficient p_1 of the first evolved RSV v_1 , shown in Fig. 3.10b, equals 0.52.

Table 3.1 gives the mean values of the eigenvalues and projection coefficients calculated from 100 high-sensitivity periods and 100 low-sensitivity periods for the first five RSVs. Comparing the mean values of the 100 sensitive and insensitive periods one notices that for sensitive periods the eigenvalues are larger. In particular, the first two eigenvalues have much larger values. Also, the projection coefficients of the first two evolved RSVs are on average much larger. So perturbation growth is not only larger during sensitive periods but also the perturbations have a preference to evolve into meridional dipolelike structures. Furthermore, the occurrence of large eigenvalues and dipolelike perturbation patterns seem to coincide.

ith RSV	High-sensitivity set		Low-sensitivity set	
	Mean p_i	Mean eigenvalue	Mean p_i	Mean eigenvalue
$i=1$	0.47	168	0.11	56
$i=2$	0.28	101	0.11	45
$i=3$	0.13	66	0.14	36
$i=4$	0.14	51	0.13	30
$i=5$	0.13	41	0.13	24

Table 3.1: Mean projection coefficient p_i and mean eigenvalue of 100 high and 100 low sensitive periods as a function of the i th RSV.

3.4.1 Barotropic versus baroclinic mechanisms

To study the relative importance of barotropic and baroclinic mechanisms some integrations are performed with the barotropic vorticity equation on a rotating sphere without orography. The barotropic model is triangularly truncated at T21. The radius of the earth is used as a unit of length and the inverse of the angular speed of rotation of the earth as the unit of time.

Optimal transition perturbations and RSVs are computed with the adjoint of the barotropic model for the same high- and low-sensitivity periods used in the computation of the baroclinic RSVs. The reference flow in the adjoint computation is taken to be the 500-hPa streamfunction evolution of the baroclinic model. The evolved barotropic RSV patterns for the sensitive periods have again a more meridional dipolelike structure and the

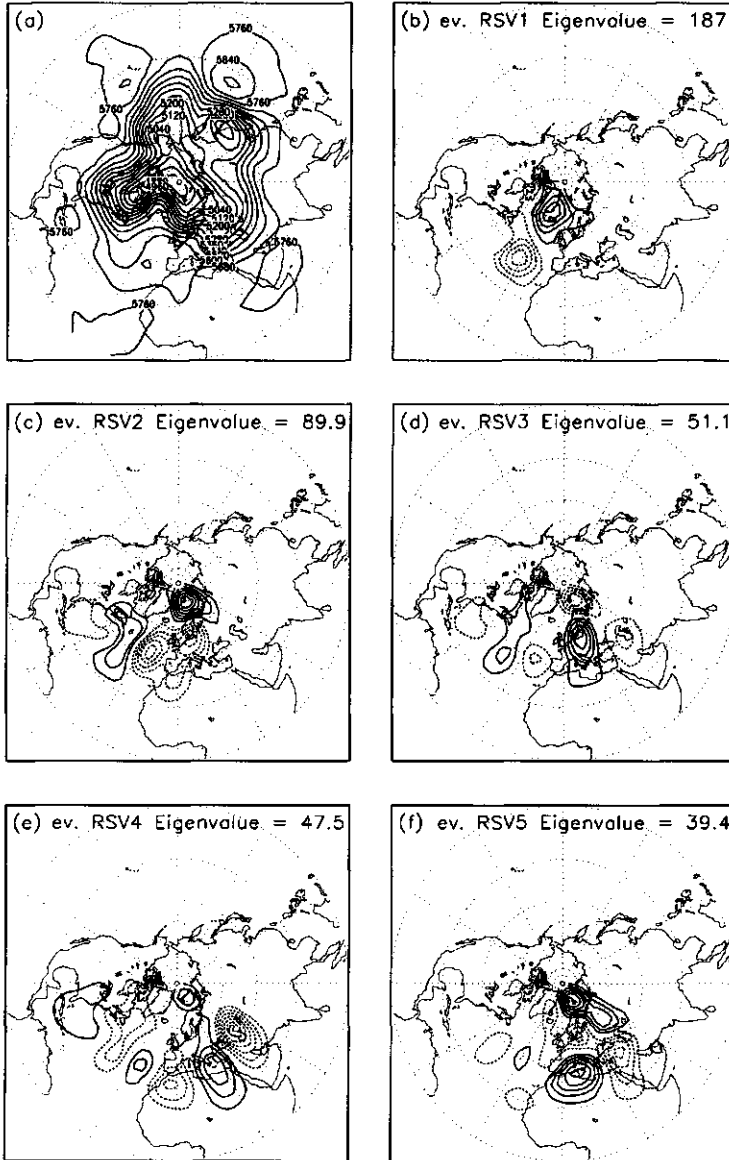


Figure 3.10: (a) Initial flow pattern in geopotential height at 500 hPa for a high-sensitivity 3-day period. Contour interval is 80 m. (b)-(f): The first five evolved RSV patterns.

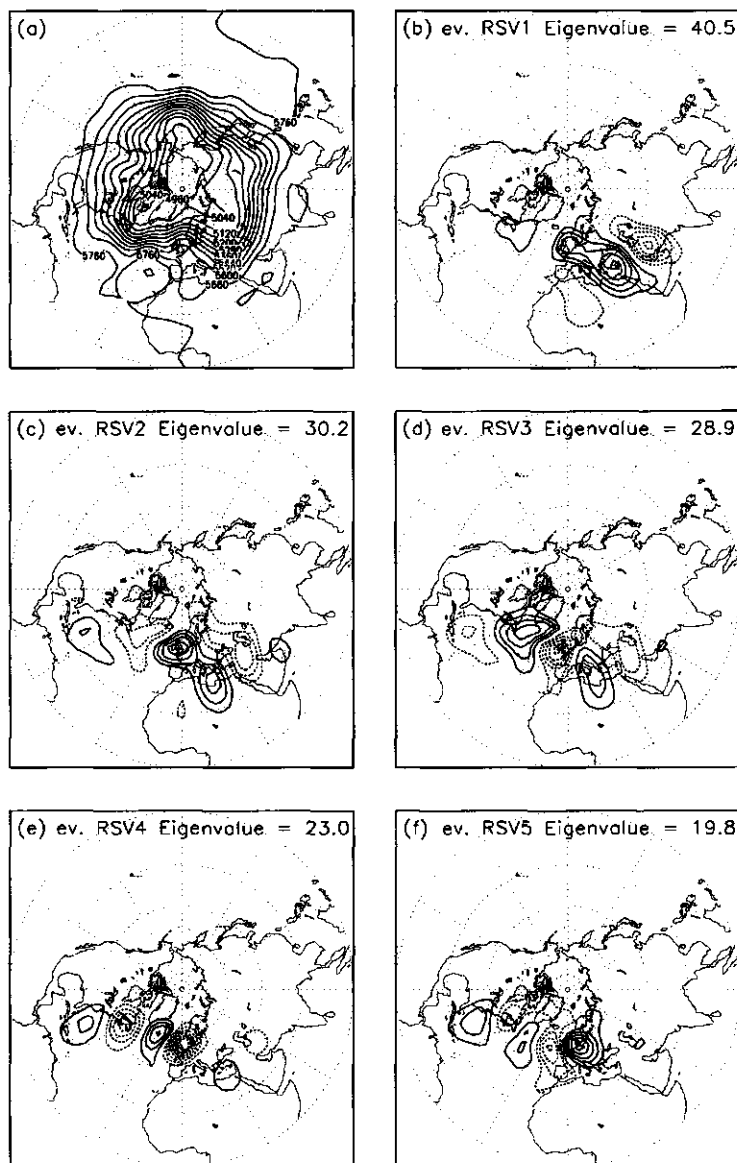


Figure 3.11: (a) Initial flow pattern in geopotential height at 500 hPa for a low sensitive 3 day period. Contour interval is 80 m. (b)-(f): The first five evolved RSV patterns.

low-sensitivity periods show zonally oriented wave trains. The growth of the barotropic perturbations/RSVs is on average a factor of 2 smaller than the growth of the baroclinic perturbations/RSVs. Baroclinic mechanisms are thus essential for realizing much larger amplifications of the perturbations.

In this chapter, we study the mean characteristics of high-sensitivity and low-sensitivity flows. The averaging process yields mean characteristics that can be understood mainly by barotropic mechanisms. Therefore, in the following we will focus on the barotropic properties of these periods. However, one must keep in mind that the baroclinic properties of the perturbations and reference flows are important and have a large contribution to the growth of the perturbations, as is shown in the last paragraph.

The sensitivity patterns (Figs. 3.4a and 3.4b, 3.6a and 3.6b) correspond to flows with a larger (high sensitivity) or smaller (low sensitivity) amplitude of the zonal wavenumber 3 mode. In order to perform an idealized experiment these flows are simulated by stationary Rossby-Haurwitz waves ($m=3$, $n=6$) with varying amplitude. Rossby-Haurwitz waves are of the form

$$\psi_{m,n}(\lambda, \mu, t) = A \operatorname{Re}\{Y_{m,n}(\lambda - \omega t, \mu)\} - C\mu, \quad (3.12)$$

where $Y_{m,n}$ is the spherical harmonic function of order m and degree n . Here, $\operatorname{Re}(z)$ denotes the real part of the complex number z , and

$$\omega = \frac{2 + C(2 - n(n+1))}{-n(n+1)}. \quad (3.13)$$

In the T21QG model, orography fixes the phase of the planetary waves. Here, the addition of a zonal velocity (proportional to C) fixes the phase of the Rossby-Haurwitz waves. The Rossby-Haurwitz wave with $m=3$ and $n=6$ is stationary ($\omega=0$) when $C=0.05$. Figure 3.12 shows two Rossby-Haurwitz waves for two different amplitudes A . Increasing the amplitude causes local areas with a stronger flow and areas with diffuent flow. Eigenvalues of the RSVs are calculated for an optimization time of 3 days.

Table 3.2 summarizes the results for the largest eigenvalue for different amplitudes of the Rossby-Haurwitz wave. Increasing the amplitude yields a higher sensitivity. In Fig. 3.13 the first 2 initial and evolved RSVs are plotted for the Rossby-Haurwitz wave shown in Fig. 3.12b. Note that the second RSV is orthogonal to the first RSV. When propagating into the diffuent area, the initial zonally elongated anomalies shrink in the zonal direction, because the westerly flow is stronger to the west than to the east (convergence of the flow in zonal direction), and become meridionally stretched, due to the diffuent flow. This process yields dipolelike blocking anomaly patterns. Similar processes can also be seen in the evolution of the baroclinic RSVs. Furthermore, the combination of a strong westerly flow upstream of a diffuent flow, where the westerly flow is much weaker, results in the convergence of the RSV patterns near the entrance of the diffuent area. This effect contributes also to larger eigenvalues of the RSVs. Namely, a locally confined large amplitude pattern yields a larger value of the norm [because of the squared weighting, like in Eq. (3.10)] than a broader low-amplitude pattern. One may wonder then why the cluster mean shown in Fig. 3.5b, a blocked circulation pattern, is low sensitive. This flow pattern yields dipolelike RSV patterns in the target area. However, because of the weak flow to the west of the block, the eigenvalues are small. Thus the flow is low sensitive.

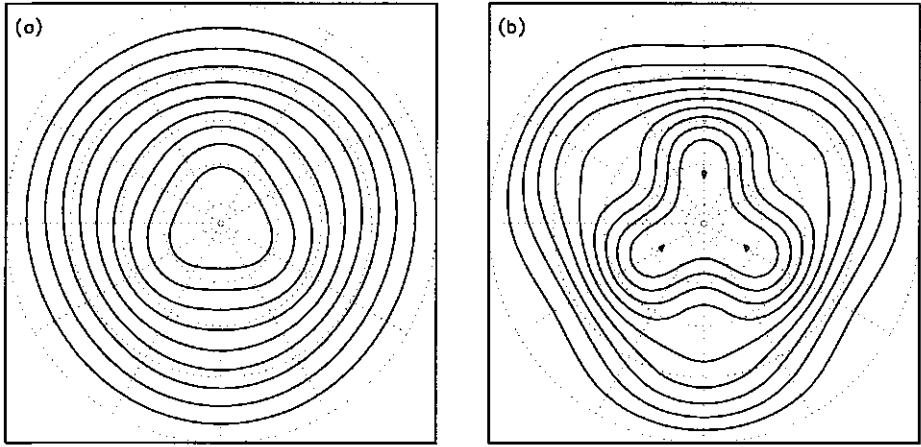


Figure 3.12: Stationary Rossby-Haurwitz waves with amplitudes equal to (a) 0.5 and (b) 2.5.

Amplitude A	0.1	0.5	1.0	2.5	5.0
Eigenvalue of first RSV	1.21	2.54	5.86	33.3	226.5

Table 3.2: Eigenvalue of the first RSV as a function of the amplitude of the Rossby-Haurwitz wave.

A look at the perturbation kinetic energy tendency equation elucidates why the RSVs have a zonally elongated structure as shown in Figs. 3.13a and 3.13c. The perturbation kinetic energy is defined by $KE' = u'^2 + v'^2$, with u' and v' the zonal and meridional perturbation velocity components, respectively. The growth of the perturbation kinetic energy can be approximated by (Simmons et al. 1983)

$$\frac{\partial KE'}{\partial t} = \overline{\mathbf{E} \cdot \nabla u_b}. \quad (3.14)$$

Here, u_b is the velocity field of the reference flow, the bar denotes an area-weighted integration and

$$\mathbf{E} = -(u'^2 - v'^2, u'v'), \quad (3.15)$$

is the so-called \mathbf{E} -vector (Hoskins et al. 1983). In the regions of diffuence, the velocity gradient $\partial u_b / \partial x < 0$. Thus, zonally elongated eddies, with $u'^2 > v'^2$, will increase the perturbation kinetic energy KE' . In case of a zonal reference flow with horizontal shear (like for many low-sensitivity periods), the second term of the basic velocity gradient $\partial u_b / \partial y$ is largest. [For a pure zonal reference flow and linear perturbation growth Eq. (3.14) is exact.] Growth of KE' is implied by eddies whose axes are tilted from northeast to southwest ($\partial u_b / \partial y < 0, u'v' > 0$) or from northwest to southeast ($\partial u_b / \partial y > 0, u'v' < 0$). The perturbation Reynolds stresses extract energy from the horizontal shear of the reference flow, decelerating the zonal flow and increasing perturbation kinetic energy. During the growth,

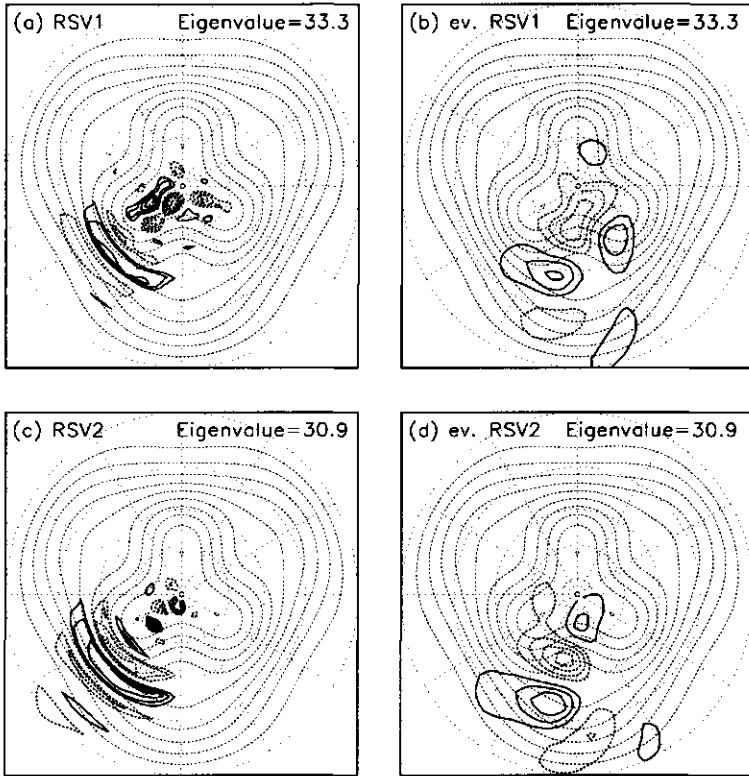


Figure 3.13: The first 2 RSVs and evolved RSVs at optimization time equal to 3 days. Contour interval in (a) and (c) is 1 (with the -1, 0, and 1 contour lines omitted) and in (b) and (d) 10 (with the 0 contour line omitted). The reference flow is the stationary Rossby-Haurwitz wave with amplitude equal to 2.5.

the tilt of the perturbations decreases and the perturbations form wave trains along a latitude. Such RSV structures are observed for the low-sensitivity periods (Fig. 3.11). A similar analysis is performed in a paper by Buizza and Molteni (1996).

The eddies are not only modified by shrinking and stretching processes but also by wave dispersion. When, for instance, only the largest eddy of the second RSV (Fig. 3.13c) is integrated linearly, the evolved pattern resembles the wave train structure of the second evolved RSV (Fig. 3.13d). However, its amplitude is much lower. The fastest growing RSVs divide their energy over several eddies instead of one large eddy. This is apparently the most optimal partition of energy. The initial amplitude of an RSV, η , is fixed through the constraint $(\eta, \eta) = n$, with n a fixed scalar. The constraint is quadratic in η . Therefore, by spreading the energy over several eddies instead of one, more energy can be 'hidden' in the initial perturbation field. Next, the process of wave dispersion yields that at optimization time the patterns add up to large-amplitude eddies, resulting in a large value of the squared

norm and thus in a large eigenvalue.

Finally, we want to discuss the impact of the diffuence of the flow on the sensitivity with respect to the medium range computations in the T21QG model. The diffuence of the flow probably results in an asymmetry between sensitivity for BL and SZF onset in the medium range. Increasing the diffuence of the flow through amplification of the blocking pattern intensifies the sensitivity of the flow, whereas weakening of the blocking anomaly pattern decreases the sensitivity. Nonlinear feedback mechanisms will therefore cause a larger sensitivity for transitions toward blocking than toward an SZF regime. This asymmetry in the sensitivity becomes important in the iterative modification procedure. After one iteration, the new reference flow for a transition toward a blocking regime is stronger diffuent than the former and so, on average, more sensitive. For transitions toward an SZF regime the opposite is true. By this asymmetry, the mean difference in \mathcal{B} between the reference and perturbed flow for high-sensitivity 5-day periods toward BL turns out to be about 10% larger than the mean difference in \mathcal{B} for high-sensitivity 5-day periods toward SZF in the T21QG model.

In summary, we have found that an upstream strong flow (characterized by the sensitivity patterns) and a local diffuent flow (only partly contained in the sensitivity patterns) are properties that make a circulation pattern sensitive for transitions to BL or SZF. The diffuent flow not only causes dipolelike RSV patterns, but also contributes to larger eigenvalues. Furthermore, the diffuence of the flow affects the medium-range sensitivity such that the sensitivity for transitions to BL is slightly larger than sensitivity for transitions to SZF. The barotropic evolution can explain a large part of the sensitivity although baroclinic growth is necessary to get much larger amplifications of the perturbations. This is also reflected in the initial baroclinic structure of the optimal perturbations (Fig. 3.9).

3.4.2 Precursors of transitions

In this subsection, we investigate whether transitions toward BL and toward SZF are more or less sensitive to initial conditions than the mean sensitivity to initial conditions. In addition, we also study whether the high or low sensitivity patterns can be related to precursor patterns of transitions. We therefore performed a 100000-day model integration.

First of all, the linear sensitivity S_L for a forecast period of 3 days is calculated each day. Again, the distribution of S_L is similar to the distribution shown in Fig. 3.3, with a mean sensitivity $\bar{S}_L = 2.69 \times 10^{-6} \text{ s m}^{-2}$. Secondly, the blocking index \mathcal{B} is calculated each day. In order to identify transitions toward BL or SZF we calculate the change $\Delta\mathcal{B} = \mathcal{B}(T = 3) - \mathcal{B}(T = 0)$. A transition is stronger when the change $\Delta\mathcal{B}$ is larger. We are now able to relate the change $\Delta\mathcal{B}$ to the sensitivity S_L . Figure 3.14a gives the mean sensitivity of all transition cases toward BL with a strength larger than the corresponding value of $\Delta\mathcal{B}$ on the x axis. So, all transitions with $\Delta\mathcal{B} > 0.25$ have a mean sensitivity of $\bar{S}_L = 2.72 \times 10^{-6} \text{ s m}^{-2}$, whereas all transitions with $\Delta\mathcal{B} > 2.75$ have a mean sensitivity of $\bar{S}_L = 3.52 \times 10^{-6} \text{ s m}^{-2}$. Hence, when the transition toward BL is stronger, the sensitivity S_L is, on average, larger. However, the standard deviation of the sensitivity values varies between $0.70 \times 10^{-6} \text{ s m}^{-2}$ and $0.95 \times 10^{-6} \text{ s m}^{-2}$. This means that individual transitions can still have a rather low-sensitivity S_L . Figure 3.14b gives the mean sensitivity of all transition cases toward SZF.

One can see that the sensitivity S_L is independent of the strength of the transition. The standard deviations vary between $0.55 \times 10^{-6} \text{ s m}^{-2}$ and $0.73 \times 10^{-6} \text{ s m}^{-2}$, indicating that individual transitions may have a high- or low-sensitivity S_L .

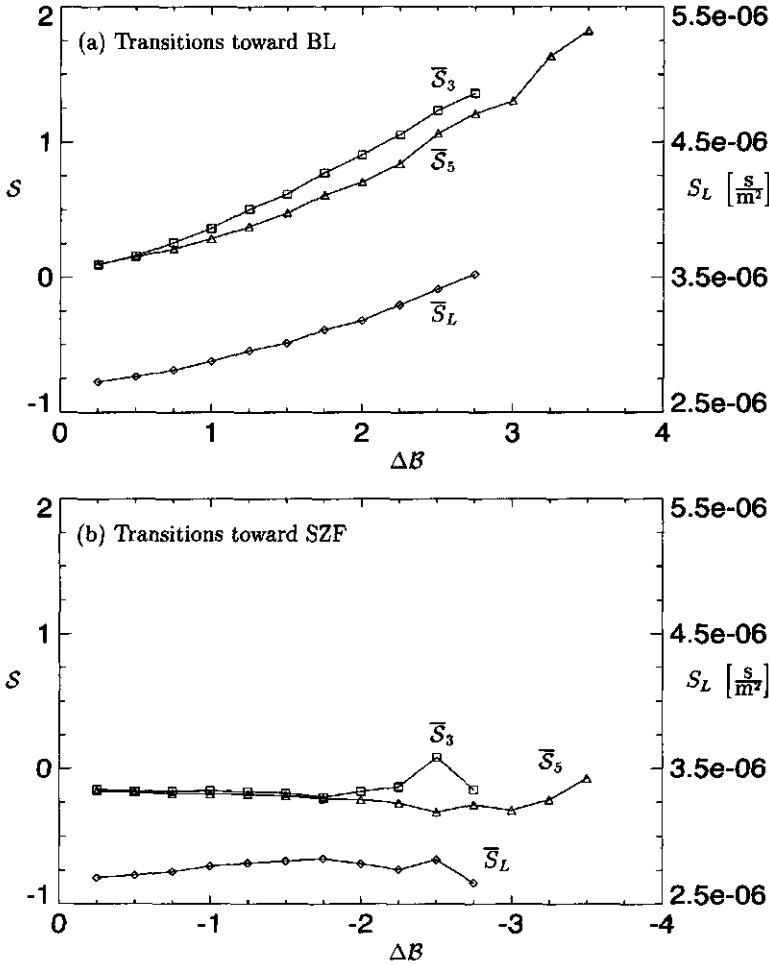


Figure 3.14: (a) Mean \bar{S}_L , \bar{S}_3 , and \bar{S}_5 of all transition cases toward BL with a strength larger than the corresponding value of ΔB on the x axis. (b) Same as in (a) but for transitions toward SZF.

Next, we look at precursor patterns of transitions. We define the n -day precursor pattern of a transition of strength ΔB as that pattern at $T = 0$ that precedes an n -day period in which B changes an amount $\Delta B = B(T = n) - B(T = 0)$. Transitions toward BL are, on average, rather sensitive and, as shown in section 3.3, high-sensitivity periods are characterized by the upstream anomaly patterns given in Fig. 3.4. So we expect that the precursor patterns

of BL have a projection on these anomaly patterns. In order to facilitate the comparison between precursor patterns and sensitivity patterns we now define a sensitivity index, \mathcal{S} . The sensitivity index \mathcal{S} is similar to the blocking index \mathcal{B} ,

$$\mathcal{S}(\psi) = \frac{\langle z_d(\psi), z_s \rangle}{\langle z_s, z_s \rangle}. \quad (3.16)$$

Here, z_s is the geopotential height anomaly field given in Fig. 3.4c and z_d is the daily geopotential height anomaly pattern. A large and positive value of \mathcal{S} indicates a large projection on z_s , so that the flow over North America is similar to that shown in Fig. 3.4a. A large negative value indicates that the flow projects on the anomaly pattern of Fig. 3.4d (which is almost the opposite of z_s) and is such more similar to the flow of Fig. 3.4b.

In Fig. 3.14a is plotted the mean $\bar{\mathcal{S}}_n$ of all transition cases toward BL within n days for n is 3 and 5, with a strength larger than the corresponding value of $\Delta\mathcal{B}$ on the x axis. One can see that the 3-day precursors as well as the 5-day precursors have, on average, a larger projection on z_s when the strength of the transition is larger. The standard deviations of $\bar{\mathcal{S}}_n$ vary barely around 1.0. In Fig. 3.14b it is shown that in the case of transitions to SZF the mean $\bar{\mathcal{S}}_n$ is independent of the strength of the transition and equal to about -0.15.

Figure 3.15a shows the mean geopotential height anomaly pattern at 500 hPa of all 3-day precursors of transitions with a strength larger than $\Delta\mathcal{B} > 1.5$. The mean projection on z_s equals 0.62. Over the North Atlantic and Europe the anomaly pattern has a negative projection on the blocking anomaly pattern, $\mathcal{B} = -0.9$, so that 3 days later the mean projection will be at least $\mathcal{B} = 0.6$. Figure 3.15b shows the same anomaly pattern but for transitions toward an SZF ($\Delta\mathcal{B} < -1.5$). The projection indices are, respectively, $\mathcal{S} = -0.18$ and $\mathcal{B} = 0.79$. Precursors of stronger or weaker transitions differ mainly in the amplitude of the precursor anomaly pattern. An analysis of precursors and successors of transitions could make clear how BL and SZF regimes are established in the T21QG model. That analysis is out of the scope of this study.

When we correlate the time series of the daily values of S_L and \mathcal{S} , the correlation coefficient is equal to 0.45. This means that in the T21QG model the flow structure above North America at 500 hPa at a particular day already explains a substantial part of the sensitivity S_L for transitions toward BL or SZF regimes 3 days later. That is remarkable since the calculation of \mathcal{S} requires no computer time, whereas for the calculation of S_L an adjoint calculation (a time-consuming process) must be performed.

So far, we measured sensitivity by calculating the linear sensitivity S_L . The underlying assumption is that S_L gives a good estimate of the local divergence of orbits in phase space. Here divergence is with respect to variations in \mathcal{B} . In order to test the validity of the assumption we performed the following calculations. When the sensitivity is high we expect a large divergence of orbits: some orbits will evolve into blocked flows, others into zonal flows. Thus for high-sensitivity periods the variations in $\Delta\mathcal{B}$ must be large, resulting in a large standard deviation σ of the distribution of $\Delta\mathcal{B}$. Orbits in low-sensitivity phase space regions will have a smaller divergence. This does not necessarily result in a lower standard deviation $\sigma(\Delta\mathcal{B})$, because nearby orbits in low-sensitivity regions can all evolve to large \mathcal{B} or to large $-\mathcal{B}$. However, when the standard deviation of low-sensitivity periods turns out to be lower than the one of high-sensitivity periods, the assumption must certainly be partly

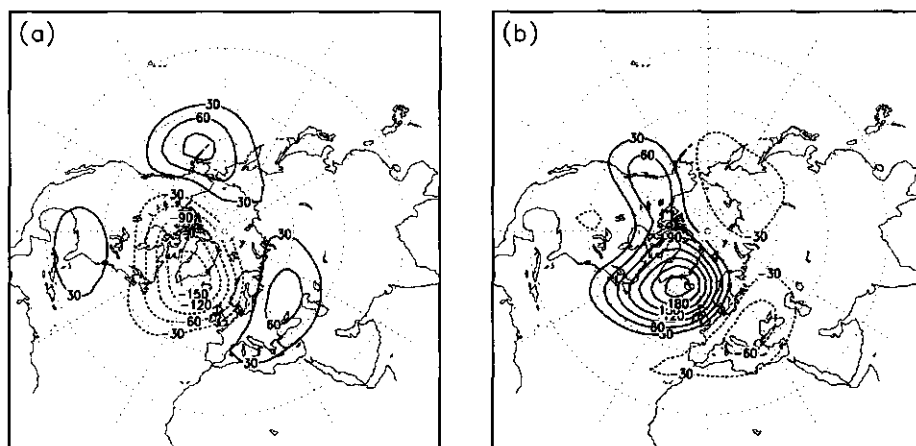


Figure 3.15: Mean geopotential height anomaly patterns of all 3-day precursor patterns with a strength (a) $\Delta B > 1.5$ and (b) $\Delta B < -1.5$.

valid. In Fig. 3.16 is plotted $\sigma(\Delta B)$ for all cases within a bin of width $0.5 \times 10^{-6} \text{ s m}^{-2}$ with a mean sensitivity, S_L , at the corresponding value on the x axis. It is clear that the standard deviation increases with increasing sensitivity. We conclude that the sensitivity S_L is a reflection of the divergence of orbits in phase space.

That transitions toward BL are on average sensitive to initial conditions and sensitivity depends on the strength of the jet stream and on the diffuence of the flow in the blocking area is in agreement with results of Nakamura and Wallace (1990, 1993) and Vautard (1990). Nakamura and Wallace (1990) found that during blocking onset periods the upstream storm track is highly active and Nakamura and Wallace (1993) found that the flow exhibits strong diffuence ahead of a deep trough. Vautard (1990) identified three precursor mechanisms for Euro-Atlantic blocking. The flow patterns of all three precursors are characterized by an intensified jet stream and/or a diffuent flow over the Atlantic region. Thus, blocking onset is likely to be in many cases an inherently sensitive phenomenon.

3.5 Summary and discussion

In this chapter, flows with high and low sensitivity with respect to initial conditions for the onset of blocking (BL) and strong zonal flow (SZF) regimes have been analysed. We have considered BL and SZF regimes at 20° W (Atlantic region) and at 150° W (Pacific region). The BL and SZF regimes are characterized by the same dipolelike anomaly pattern but with opposite signs. Experiments have been performed with a quasigeostrophic model triangularly truncated at wavenumber 21 (T21QG), and its tangent linear and adjoint versions. Using the method described in chapter 2 (Oortwijn and Barkmeijer 1995) initial perturbations can be computed that optimally trigger the onset of a BL or SZF regime after a prescribed forecast time. For optimization times larger than 3 days an iterative technique is used to take into

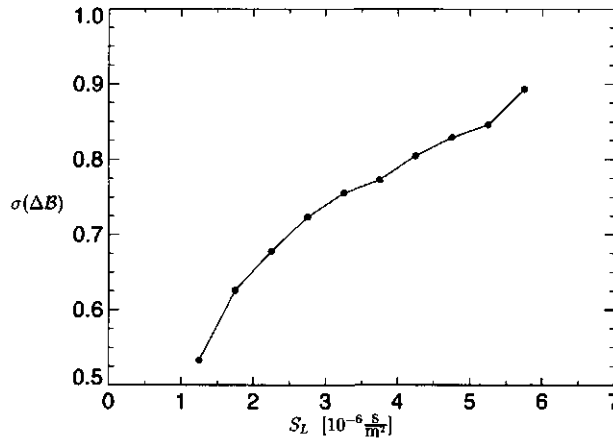


Figure 3.16: Standard deviation $\sigma(\Delta B)$ for all cases within a bin of $0.5 \times 10^{-6} \text{ s m}^{-2}$ with a mean sensitivity, S_L , at the corresponding value on the x axis.

account nonlinear growth of the perturbations.

We have performed three large experiments. For each experiment the T21QG model is integrated for 20000 days, and for each day the sensitivity for onset of BL and SZF regimes is calculated. The first two experiments concern sensitivity for periods of a 3-day duration in the Atlantic and Pacific regions, respectively. In the third experiment the sensitivity for 5-day periods is calculated for the Atlantic region. The initial reference flow patterns of the 1000 most sensitive and 1000 most insensitive periods have been selected. The mean fields of these patterns have been analysed. It appears that the initial patterns of the high-sensitivity periods of the first and third experiment (Atlantic region) are characterized by an enhanced ridge over the northeast Pacific, and a deep trough over the North American continent with a ridge to its southeast. Low sensitivity flows are more zonal and weaker in this area (characterized by the opposite anomaly patterns with respect to the climate mean). The mean patterns for the Pacific regimes have similar features. High-sensitivity periods are characterized by an enhanced ridge over Asia, a deep trough at the east coast of the Asian continent, and a ridge to its southeast. The flow of low-sensitivity periods is again more zonal and weaker. In the target regions (Atlantic and Pacific regions) many of the high-sensitivity flows are characterized by a ridge and by diffuent flow for 3-day optimization times. The projection of a 500-hPa daily flow pattern on the mean sensitivity patterns already explains a substantial part of the 3-day sensitivity for transitions toward BL or SZF. The influence of nonlinear terms, causing an asymmetry between sensitivity for BL and SZF onset in the medium range, is noticeable in the patterns in the target area of high or low sensitivity flows. Flows that are sensitive for BL onset and insensitive for SZF onset are more zonal in this region. Flows that are insensitive for BL onset and sensitive for SZF onset show an enhanced ridge in this region. Thus, the nonlinear terms cause, on average, a tendency of the flow to return to the climate mean.

Simulations with the T21QG model and a T21 barotropic model reveal that two main

properties of a flow are responsible for the sensitivity of the flow: a diffuent flow in the target area and an intensified jet stream upstream of it. The strong jet stream by itself results in large perturbation growth. The presence of a diffuent flow amplifies the growth and gives perturbations a typical dipolelike character. These properties can be understood with barotropic mechanisms. However, baroclinic properties of the perturbations and reference flows are important and have a large contribution to the growth of the perturbations. The relation between a diffuent flow and sensitivity also results in an asymmetry between sensitivity for BL and SZF onset in the medium range. Nonlinear feedback mechanisms increase sensitivity toward BL and decrease sensitivity toward SZF. High-sensitivity flow patterns are also characterized by enhancement of the ridges over the northeast Pacific (sensitivity for Euro-Atlantic BL and SZF regimes) and Asian continent (sensitivity for Pacific BL and SZF regimes). These enhancements do not directly influence the sensitivity. The figures of the variance in the initial perturbation fields (Fig. 3.9) show that the perturbations do not have much amplitude in these regions. The dynamics of the model (and atmosphere?) causes a deep trough associated with an upstream ridge; that is, the combination is a preferred flow pattern. The variability over the northeast Pacific is larger than over the Asian continent, so that the amplitude of the ridge is larger over the northeast Pacific.

It is shown that, on average, a transition toward BL corresponds with a larger than average sensitivity and that the sensitivity is larger when the transition is stronger. Transitions toward an SZF correspond, on average, with an average sensitivity independent of the strength of the transition. The precursor patterns of transitions toward BL have similar characteristics as the flows with a high sensitivity. These results are in general agreement with characteristics of precursors of blocking onset found in observations by Nakamura and Wallace (1990, 1993) and Vautard (1990). Thus, blocking onset is likely to be in many cases an inherently sensitive phenomenon.

Finally, we point out the importance of the norm. It is well known that the structures of the optimal perturbations depend strongly on the norm. In this chapter a squared norm is used to calculate the optimal perturbations. Some calculations are performed with a kinetic energy norm to study whether the sensitivity patterns also depend strongly on the norm. Using a kinetic energy norm, the correlation between the wind fields of the blocking anomaly pattern and the perturbation pattern is maximized rather than the correlation between the geopotential height anomaly fields. A kinetic energy norm puts more weight on larger scales of the initial perturbation. It is found that with a kinetic energy norm the high-sensitivity circulation patterns found with the streamfunction norm are also more sensitive than the low-sensitivity circulation patterns found with the streamfunction norm. So, the sensitivity patterns depend much less on the norm than the initial perturbations. Besides, a test that shows that the sensitivity reflects the divergence of orbits in phase space of the model with respect to BL and SZF transitions has also been performed. Therefore, the conclusions made in this chapter are probably rather general and not specific for the perturbations used here to compute the sensitivity of circulation patterns.

Acknowledgements. Many thanks are due to Jan Barkmeijer, Robert Mureau, Theo Opsteegh, and Wim Verkley of the predictability section at the Royal Netherlands Meteorological Institute for the stimulating discussions and their assistance during the preparation

of this chapter. I would like to thank Paul-Antoine Michelangeli and Robert Vautard, who were so kind to perform computations with their cluster analysis algorithm. I am also indebted to three anonymous reviewers for their comments that led to the improvement of the original manuscript. This work has partly been supported by the foundation for Geological, Oceanographic, and Atmospheric Research (GOA) of the Netherlands Organization for Scientific Research (NWO). Computing facilities were provided by the Royal Netherlands Meteorological Institute.

Chapter 4

Growth properties of optimal transition perturbations

The development of perturbations that optimally trigger the onset of Euro-Atlantic blocking (BL) and strong zonal flow (SZF) regimes has been investigated. These perturbations will be called optimal transition perturbations. First, a T21 three-level quasigeostrophic model (T21QG) including a forward and adjoint tangent propagator, is used to compute the sensitivity in the initial conditions for onset of BL and SZF regimes. The evolution of an optimal transition perturbation during a sensitive 72-hour period is extensively studied. Barotropic and baroclinic mechanisms are distinguished by displaying the results in terms of the barotropic and baroclinic modes of the system. Next, the perturbation is decomposed in normal modes. The evolution can be divided in two phases. During the first rapid phase, the growth is strongly non-modal and baroclinic. After that, the growth is still non-modal but not as strong and almost barotropic.

In the second part of this chapter, the barotropic evolution is studied using a leading order WKB approximation adopted for non-zonal smooth background flows. This approach is based on the assumptions that the perturbations may be represented by wave packets and that a scale separation between the perturbations and the background flow can be made. The WKB approach is used as a diagnostic tool to interpret the evolution of the optimal perturbations qualitatively. We have focused on the evolution of zonally elongated wave packets that are located in or near the jet stream, and propagate into a diffluent area. Because the background flow is non-zonal, total wave action of a packet is not conserved. However, under certain conditions total wave enstrophy of a packet is conserved. The WKB equations predict reasonably well the evolution of the perturbations, although the assumptions are violated in the final stage of the integration period.

4.1 Introduction

Regime transitions are often very rapid and difficult to predict (Tibaldi and Molteni 1990; Palmer et al. 1990; Kimoto et al. 1992; Tibaldi et al. 1994). It is therefore of great practical importance to investigate these transitions. In chapter 3, periods with a high- and low-sensitivity in the initial conditions for onset of blocking (BL) and strong zonal flow (SZF) regimes were studied (Oortwijn 1998a). One of the results was that flows with a high

sensitivity are characterized, on average, by an intensified jet stream to the west of a diffuent flow. The sensitivity was calculated by computing perturbations that maximize transitions between BL and SZF regimes within a fixed period of time. In this chapter we will focus on the structure and evolution of such optimal perturbations in terms of vertical modes, normal modes and with a WKB approach.

It is now generally known that perturbations may grow much faster, although for a finite time, than the fastest growing normal mode (e.g. Orr 1907; Farrell 1982; Trefethen et al. 1993; Buizza and Palmer 1995). Normal modes (NMs) are eigenvectors of the linear evolution operator. They grow exponentially and preserve their shape. The linear evolution operator is generally non-normal. The set of eigenvectors is therefore not orthogonal. As a consequence, a small perturbation written as a sum of NMs can have large components with respect to this basis. This may result in a larger than exponential growth of the perturbation with a shape that evolves in time. Here we will actually compute, within the context of a T21 three-level quasigeostrophic model (hereafter called T21QG model, described in Marshall and Molteni 1993), the complete set of 1449 NMs and show how the components of the optimal perturbation change as a function of time.

The optimal transition perturbations are directly related to singular vectors (SVs), which maximize RMS error growth over a fixed period of time (see chapter 2). Singular vectors have been applied to predictability and sensitivity studies (Farrell 1988 and 1989; Molteni and Palmer 1993) and are used in the ECMWF operational ensemble forecasting system (Mureau et al. 1993; Molteni et al. 1996). Optimal perturbations/SVs are able to grow fast by extracting energy very efficiently from the background flow. Both barotropic and baroclinic processes provide energy for the development of the perturbations. In this chapter, we will separate barotropic and baroclinic modes in order to determine their relative importance.

In order to interpret the growth of optimal perturbations, we will also look at it from another perspective. Many singular vectors are strongly localized and have a wave train-like structure. Therefore, they may be considered as Rossby wave packets propagating on a mean flow dominated by much larger scales. When the time mean flow \bar{u} is assumed to be zonal, a conservation law for the evolution of wave packets can be derived using a leading order WKB approach (after Wentzel, Kramers and Brillouin). Under these conditions total wave action is conserved (Bretherton and Garrett 1969). This conservation law states that for a packet, with local zonal wavenumber m and frequency ω , the ratio between the energy of the packet and its intrinsic frequency $\hat{\omega} = \omega - m\bar{u}$ is constant. Hence, a wave packet that propagates into a stronger zonal flow acquires a larger intrinsic frequency and therefore gains energy. Buizza and Palmer (1995) use this argument to interpret SV growth. However, as is already mentioned above, in periods with a large sensitivity to the onset of BL or SZF regimes the mean flow is diffuent and thus essentially non-zonal. Under certain conditions wave action conservation may be generalized to conservation of wave enstrophy for wave packets propagating on non-zonal mean flows (Young and Rhines 1980). It is investigated whether we can interpret the growth of the optimal perturbations by means of these concepts.

This chapter is organized as follows. In section 2 we will describe the time evolution of an optimal transition perturbation. In section 3 we will derive equations for the barotropic evolution of wave packets on non-zonal mean flows using a WKB approximation. In section 4 we will show the numerical results of this approach and compare them to results of com-

putations performed with a barotropic model. Finally, we will discuss the main results in section 5.

4.2 The evolution of an optimal transition perturbation

4.2.1 General remarks

In section 2.3 (Oortwijn and Barkmeijer 1995) we described a method to compute perturbations that optimally trigger a BL or SZF onset. It was shown that the streamfunction of the initial optimal perturbation is given by Eq. (2.9)

$$\varepsilon_L(0) = \lambda \mathbf{R}^* \mathbf{L}_z^* z_b. \quad (4.1)$$

Here, λ is a scaling factor. The pattern z_b is a dipolelike geopotential height anomaly field at 500 hPa, which characterizes a Euro-Atlantic BL regime (positive sign) or a SZF regime (negative sign). This so-called blocking anomaly pattern was computed by Liu (1994) using atmospheric observations and again by Liu and Opsteegh (1995) using data of the T21QG model. These patterns differ only slightly. The blocking anomaly pattern consists of a very strong positive geopotential height anomaly with its center at about 60°N and a weaker negative anomaly to the south of it (see Fig 2.1 of chapter 2). The operators \mathbf{L}_z^* and \mathbf{R}^* are the adjoint operators of \mathbf{L}_z and \mathbf{R} with respect to the squared norm inner product

$$\langle f, g \rangle = \sum_{i=1}^3 \int \int f_i g_i dS. \quad (4.2)$$

The sum runs over the three levels (200, 500, and 800 hPa, $i=1,2,3$) of the T21QG model and the integration is performed over the sphere. The linear operator \mathbf{L}_z transforms streamfunction into geopotential height, based on the linear balance equation. The linear evolution operator $\mathbf{R}(0, T)$ describes, to first order, the evolution of a small perturbation starting at initial time 0 to optimization time T . It follows, that the backward integration of the blocking anomaly pattern, using the adjoint operator \mathbf{R}^* , gives the optimal perturbation $\varepsilon_L(0)$, which at optimization time T has the largest projection onto this dipolelike pattern. A negative value of the scaling factor λ results in a change toward a SZF regime, a positive value toward a BL regime.

The linear sensitivity S_L was defined in Eq. (2.11) on the basis of the amplification rate of the perturbation,

$$S_L = \frac{\|\mathbf{R}^* \mathbf{L}_z^* z_b\|}{\|z_b\|^2}, \quad (4.3)$$

where $\|\varepsilon\|^2 = \langle \varepsilon, \varepsilon \rangle$. Hence, a large amplification corresponds to a high sensitivity. In chapters 2 and 3 it was shown that the linear sensitivity S_L for a three-day time interval fluctuates strongly. The mean linear sensitivity \bar{S}_L of the total distribution (20000 cases, Fig. 4.3) is $2.69 \times 10^{-6} \text{ s m}^{-2}$ with a standard deviation of $0.68 \times 10^{-6} \text{ s m}^{-2}$. The difference in the three-day sensitivity can be as large as a factor 6 between high- and low-sensitivity periods.

So far, we have focused on the mean properties of these sensitive and insensitive flows. In this chapter, we will pay attention to the evolution of the perturbations during a three-day time interval. We will describe the evolution of an optimal perturbation ε_L for one sensitive period of the T21QG model and show which mechanisms are involved in the amplification of the perturbation. In other sensitive cases the evolution of the optimal perturbations is very similar. In the discussion, we will distinguish between barotropic and baroclinic mechanisms. In order to determine their relative importance it will be useful to write the fields in the T21QG model as a sum over vertical modes. In the three-level T21QG model one barotropic mode and two baroclinic modes can be distinguished. The derivation of these modes is given in Appendix A.

4.2.2 A representative example of a sensitive period

Vertical mode decomposition

Figure 4.1 shows the streamfunction of a sensitive three-day period and its optimal perturbation at $T = 0$ and optimization time $T = 72$. The linear sensitivity S_L equals $4.35 \times 10^{-6} \text{ s m}^{-2}$, which is about 2 standard deviations larger than the mean value. Figure 4.1a shows the streamfunction of the perturbation superimposed on the streamfunction of the initial reference flow at 200 hPa. Figures 4.1b and 4.1c show the 500- and 800-hPa fields, respectively. The perturbation is located over the North-Atlantic ocean and North-American continent. It has a baroclinic structure and its energy is concentrated predominantly in small-scale eddies. The perturbation is mainly located in or near the jet stream, intensifying or moderating local gradients. In Fig. 4.2 the streamfunction of the corresponding barotropic mode and the first and second baroclinic modes are given. A large part of the energy is concentrated in the second baroclinic mode at initial time. We recall that the relation between the streamfunction at model levels and the vertical modes is derived in appendix A. The vertical structure of the streamfunction field of the barotropic mode is independent of height. The first baroclinic mode changes its sign once in the vertical, whereas the second baroclinic mode changes sign twice.

The evolved perturbation is obtained by integrating the tangent linear model. It is characterized by a large negative streamfunction anomaly surrounded by two positive streamfunction anomalies above the Atlantic and Western-Europe. The vertical structure has become equivalent barotropic (Figs. 4.1d, 4.1e and 4.1f) and so it is the barotropic mode (Fig. 4.2d) that dominates the structure. The first (Fig. 4.2e) and second (Fig. 4.2f) baroclinic modes contribute little to the perturbation field. The amplification of the perturbation is enormous. Measured in total kinetic energy (TKE) the amplification is a factor 93.5, with

$$TKE = \sum_{i=1}^3 KE_i = \sum_{i=1}^3 \int \int \nabla \psi_i \cdot \nabla \psi_i dS. \quad (4.4)$$

Here $\nabla \psi_i$ is the gradient of streamfunction at model level i .

In Fig. 4.3a the kinetic energy (KE) of the perturbation is plotted as a function of time. Shown are TKE and KE_i at 200, 500, and 800 hPa ($i=1,2,3$). The TKE is normalized to one initially. The growth is nearly uniform at all levels. The KE of the vertical modes is

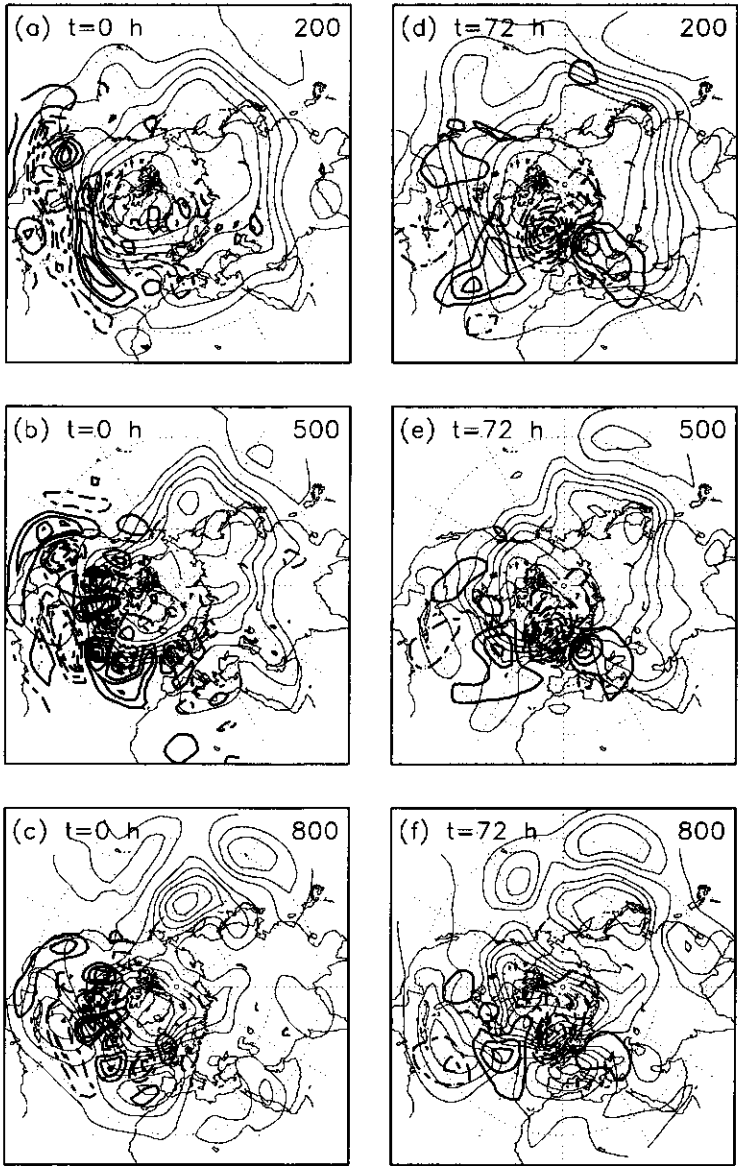


Figure 4.1: Streamfunction of the optimal perturbation superimposed on the reference flow at the three levels 200, 500, and 800 hPa at initial and optimization time. At optimization time the perturbation contour interval is 20 times larger than at initial time.

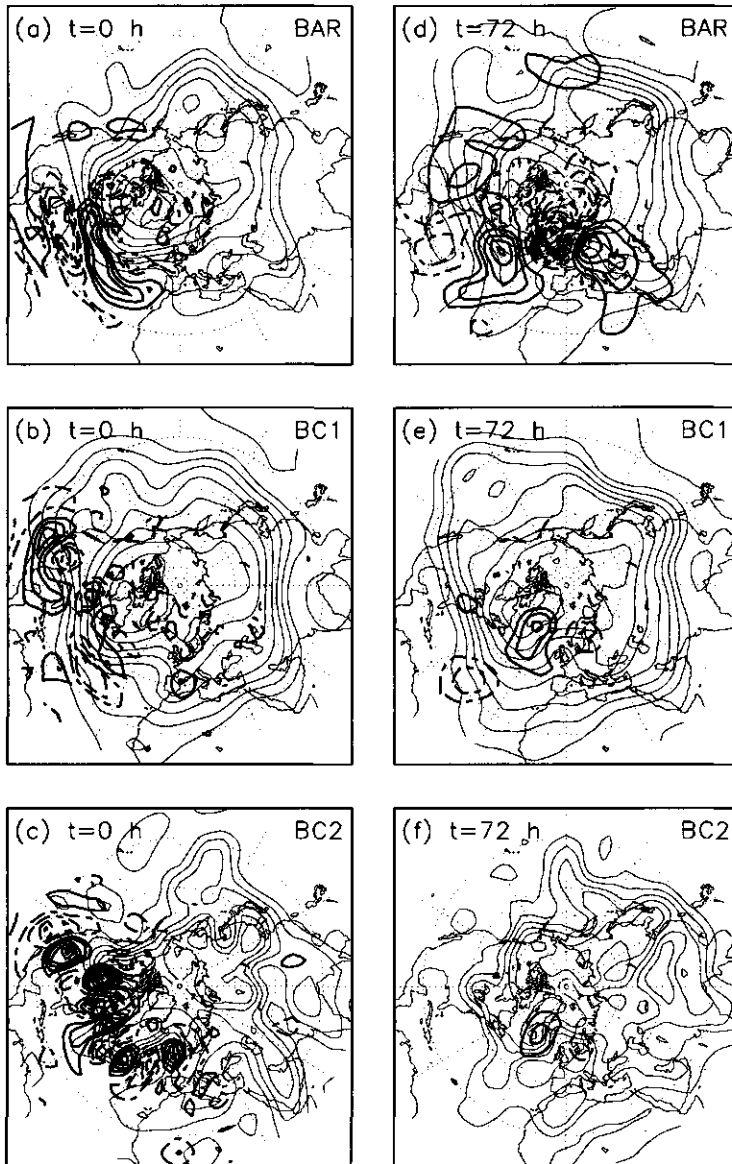


Figure 4.2: Streamfunction of the barotropic, first baroclinic and second baroclinic mode of the optimal perturbation and reference flow at initial and optimization time. At optimization time the perturbation contour interval is 20 times larger than at initial time.

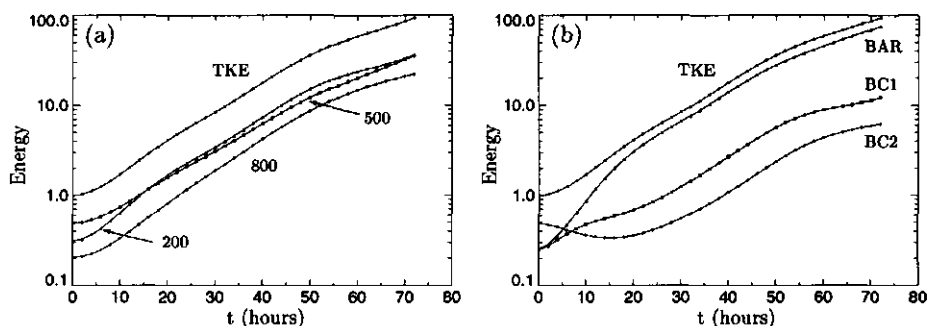


Figure 4.3: Kinetic energy of the optimal perturbation (log-scale) as a function of time. (a) TKE (normalized to 1 initial time) and KE of the three model levels, (b) TKE and KE of the three vertical modes.

given in Fig. 4.3b. It shows that initially the second baroclinic mode dominates the *KE*. The *KE* in the second baroclinic mode even decreases for a while and within 24 hours almost all *KE* is located in the barotropic mode. So, the perturbation gains energy by barotropic as well as by baroclinic processes during the first 24 hours. After that, barotropic processes dominate the evolution (see also Molteni and Palmer 1993; Mureau et al. 1993). The initial decrease of baroclinic energy, which is essentially available potential energy, implies that the latter is converted to barotropic kinetic energy. The available potential energy (*APE*) in the T21QG model consist of two terms that are proportional to the square of the 200-500 hPa thickness and the 500-800 hPa thickness,

$$APE = APE_{12} + APE_{23} = \iint \left(\frac{1}{2R_1^2} (\psi_1 - \psi_2)^2 + \frac{1}{2R_2^2} (\psi_2 - \psi_3)^2 \right) dS, \quad (4.5)$$

where R_1 (=700 km) and R_2 (=450 km) are Rossby radii of deformation. The difference between *APE* in the reference flow and perturbed flow increases particularly the first 12 hours very fast (not shown). Hence, *APE* has been converted to *KE* in the perturbed flow.

In order to test whether the evolution of the perturbation during the last 48 hours can indeed be described mainly by barotropic processes, the following experiment has been performed: the optimal perturbation is integrated with the tangent linear model of the barotropic vorticity equation along the 500-hPa reference orbit, which was obtained by integrating the baroclinic model. Then, the initial optimal perturbation at 500 hPa is integrated for 72 hours. The *KE* amplification of this perturbation is only 67 % of the *KE* amplification of the three-level perturbation at 500 hPa in the baroclinic T21QG model. Next, we take the 24-hour evolved optimal perturbation of the baroclinic model at 500 hPa and integrate this structure along the last 48 hours of the baroclinic reference flow using again the tangent linear model of the barotropic vorticity equation. In this case, the amplification is 93 % of the original amplification. In both cases the evolved structure is very similar to the original perturbation shown in Fig. 4.2b. We conclude that during the last 48 hours barotropic mechanisms dominate the evolution strongly.

The evolution can thus be separated into two phases. In the first phase, lasting about 24 hours, baroclinic and barotropic processes both cause growth of the perturbation, which has the structure of a baroclinic wave train. Its small-scale structure at 500 and 800 hPa disappears and an equivalent barotropic wave train arises, of which the growth during the second phase, from about 24 to 72 hours, can be understood in terms of barotropic processes.

Normal mode decomposition

In order to further study the two phases of the evolution, a comparison is made between the growth of the optimal perturbation and the growth of the normal modes. Normal modes (NMs) are eigenvectors of \mathbf{R} in case of a stationary reference flow. The tangent linear operator of a stationary reference flow can be written as $\mathbf{R} = \exp(\mathbf{F}' \cdot T)$, where \mathbf{F}' gives the linear evolution of a perturbation, $\frac{d\epsilon}{dt} = \mathbf{F}'\epsilon$. The eigenvectors of the real matrix operator \mathbf{F}' are generally complex with complex eigenvalues $\exp((\lambda + i\mu)T)$. The complex eigenvectors form complex conjugated pairs with complex conjugated pairs of eigenvalues. The real and imaginary parts of the eigenvector pairs form an invariant two-dimensional subspace in the real phase space. This causes modal growth of the NMs, which means that NMs grow exponentially and, in addition, oscillate in the invariant two-dimensional subspace. Apart from these periodic oscillations, NMs do not change their pattern in time. In general, the operator \mathbf{R} is non-normal. Because of this, the NMs form a nonorthogonal basis. As a result, a small perturbation written as a sum of NMs can have relatively large components with respect to this basis of NMs. The dimension of the phase space of the T21QG model is 1449, so that a given perturbation ϵ can be written as

$$\epsilon = \sum_{i=1}^{1449} \eta_i \mathbf{n}_i, \quad (4.6)$$

where η_i is the i th component of ϵ with respect to the basis \mathbf{n} . When the i 'th and $(i+1)$ 'th NM form a complex conjugated pair then we will take \mathbf{n}_i equal to the real part of these NMs and \mathbf{n}_{i+1} equal to the imaginary part of these NMs. In this way, \mathbf{n}_i and \mathbf{n}_{i+1} span the invariant two-dimensional subspace.

To compute the NMs, the mean flows of the first 0 to 24 hours and the last 24 to 72 hours are taken as time invariant basic states. In Fig. 4.4a the components of the initial optimal perturbation are shown with respect to the basis formed by the NMs of the 24-hour mean flow. The NMs are sorted by the real part of the eigenvalue so that the first NM amplifies fastest. Both NMs and optimal perturbation are normalized to 1 with respect to the squared norm inner product, Eq. (4.2). The first 465 NMs amplify, all other NMs decay. The fastest e -folding time is about 45 hours. The optimal perturbation is a combination of many amplifying and decaying NMs. As many as 69 components have a value larger than 1 with a maximum of 3.5. The sum of the squares of the components equals 213, which should be equal to 1 for an orthonormal basis. After 24 hours, this sum for the normalized evolved optimal perturbation reduces to 31, see Fig. 4.4b. This indicates that the initial perturbation is constructed in such a way that it is a combination of many NMs that point into directions that partially cancel each other. The different oscillation and amplification properties of the NMs cause a fast non-modal growth of the perturbation. The dispersion of the pattern

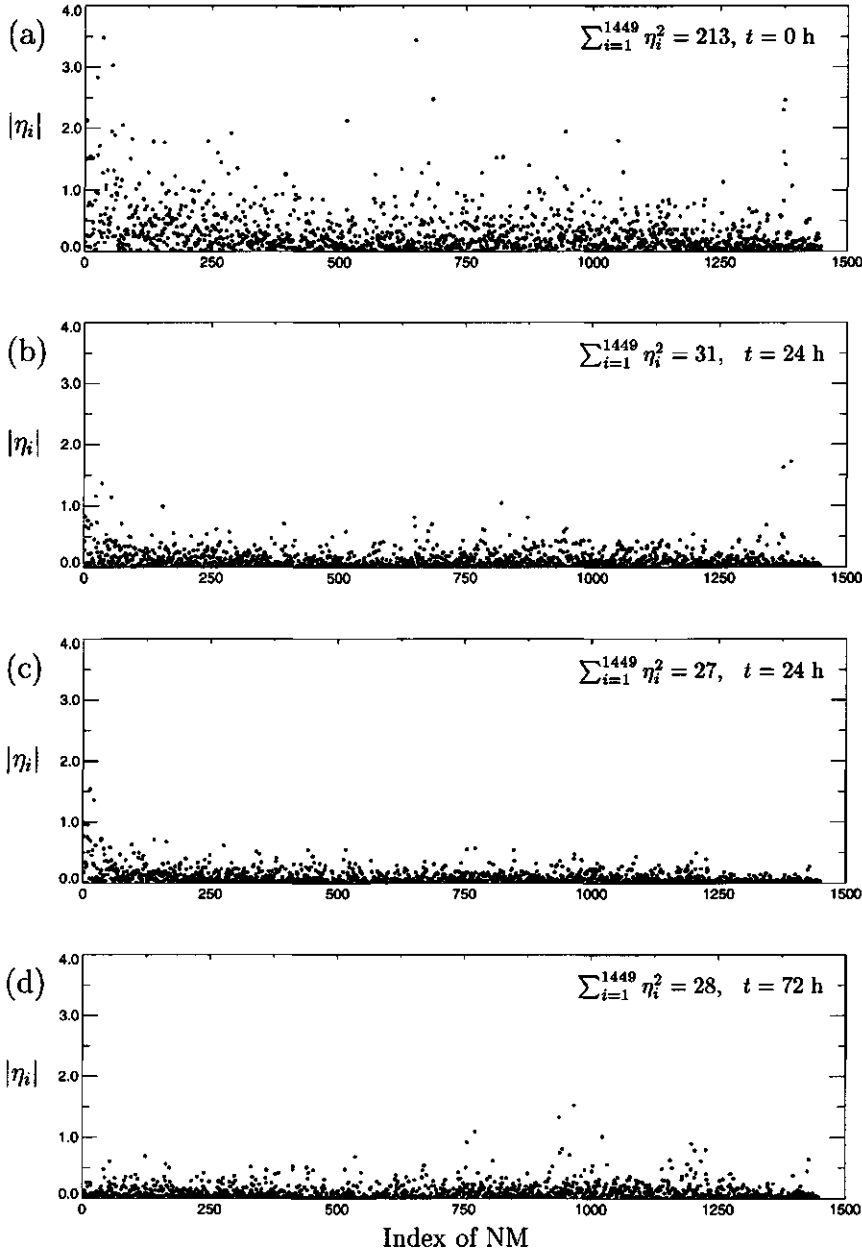


Figure 4.4: Absolute values of the components of the optimal perturbation with respect to a basis of NMs. The NMs are sorted by amplification rate such that the first NM grows fastest. (a) Optimal perturbation at initial time and (b) after 24 hours for NMs of first basic state. (c) After 24 hours and (d) 72 hours for second basic state.

is also clearly noticeable in the evolution of the optimal perturbation (see also Fig 4.5). Growth through rapidly amplifying NMs seems to be of secondary importance, although the initial components of the growing NMs are somewhat larger than of the decaying modes. In Fig. 4.4c the components of the normalized 24-hour evolved optimal perturbation are given with respect to the NMs of the average 24- to 72-hour basic state. Figure 4.4d shows the normalized optimal perturbation after 72 hours. In the last 48 hours the sum of the squared components does not reduce anymore. The perturbation mainly grows by amplification of NMs. In Fig. 4.4c one can see that the components that are significant are principally with respect to the fastest growing NMs. After 72 hours, significant components are also with respect to the decaying NMs. The perturbation will not show any substantial growth anymore.

Time evolution

As we have seen, the baroclinic modes cause an initial displacement from available potential energy of the reference flow to perturbation kinetic energy. This process occurs fast, within 12 to 24 hours. After that, the energy transfer is completely dominated by the kinetic energy transfer from the reference flow to the perturbation. It seems that the baroclinic growth principally results in an additional amplification of the structure. Without the baroclinic growth, the evolved perturbation would be weaker but its spatial structure, as is shown in the subsection 4.2.2, would be much less affected. This extra amplification is important though; it leads on average to a factor 2 larger growth (see section 3.4.1). The barotropic interactions between the reference flow and perturbation determine effectively the further evolution and structure of the perturbation. We shall therefore, in the remainder, focus on the barotropic part of the growth. First, the time evolution of the barotropic mode will be described. Then, in the next sections, we will consider at length the barotropic mechanisms that are involved using a WKB approach.

In Fig. 4.5 the barotropic mode of the optimal perturbation is plotted together with the associated reference state. The evolution is shown every 6 hours for the first 36 hours, after that, every 12 hours. The evolution can be characterized by a developing wave train. Especially in the first part the perturbation grows very fast. However, as shown in Fig. 4.3 this is partly due to baroclinic energy conversion. The wave train propagates with a group velocity that is larger than the phase velocity of the individual eddies; wave crests appear in front of the packet and disappear at the back. The energy propagates through the train, resulting in the downstream development of the eddies. Furthermore, it can be observed that wave energy propagates northward relative to the background velocity field. During the evolution, the zonal length scale of the eddies decreases and the meridional length scale increases. In addition, the horizontal tilt of the eddies from southeast to northwest vanishes. The development of the ridge in the reference flow near the end of the period causes the most downstream crest to move southward and the second crest to move northward so that, together with the third crest, a typical blocking anomaly pattern arises.

In summary, the optimal transition perturbation has initially the structure of a rather small-scale baroclinic wave train. However, the barotropic part of the perturbation has initially already a large-scale zonally elongated structure. Baroclinic energy sources supply

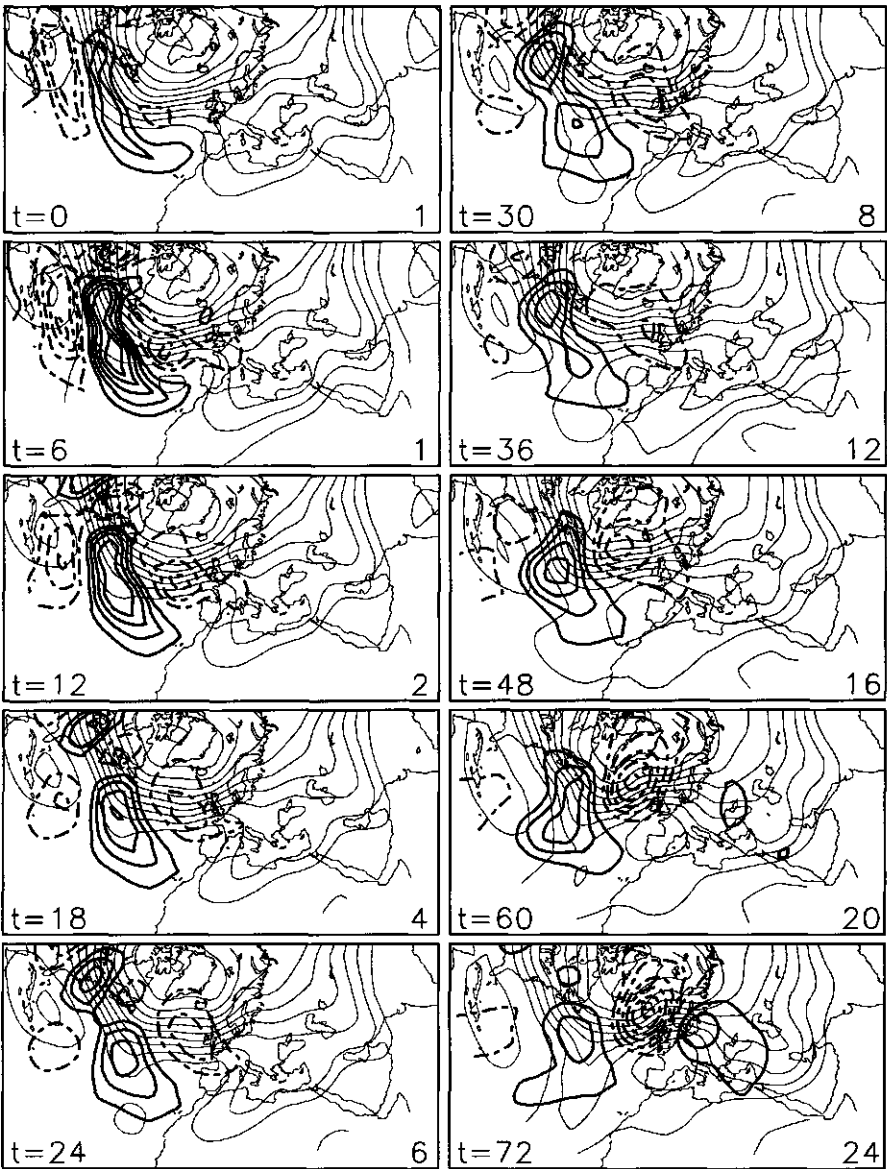


Figure 4.5: The evolution of the barotropic mode of the optimal perturbation is shown together with the associated reference flow. The first 36 hours the time interval is 6 hours, after that the time interval is 12 hours. The contour interval of the perturbation is set to 1 initially. In each panel the contour interval of the perturbation is given.

energy for the barotropic kinetic energy growth. In the first hours available potential energy is converted into kinetic energy. After this first rapid phase of strongly non-modal growth, the perturbation evolves mainly by barotropic mechanisms, with it extracting kinetic energy from the reference flow. In the next sections, we will study these dominant barotropic processes by considering wave packets on a barotropic basic state.

4.3 A WKB approximation

We will now study the barotropic interaction between the perturbation and the mean background flow. In chapter 3 it was shown that in sensitive periods the mean 500-hPa flow is characterized by an intensified jet stream upstream of a diffuent flow associated with a ridge over the Euro-Atlantic area. This leads to a picture in which the larger scales already have a ridge structure that can be reinforced or weakened by the smaller-scale perturbations. The mean flow is thus essentially non-zonal. This makes the analysis much more complex than in the case of a zonal background flow. In the following, we will assume that a scale separation can be made between the perturbations and mean flow. In addition, it is explicitly assumed that the perturbations can be represented in the form of wave packets.

In addition to the results for the sensitivity on the 500-hPa level, we have described in section 4.2.2 that the barotropic evolution on the 500-hPa level was a reasonable approximation to the latter part of the evolution of the optimal perturbation. We shall therefore compare the evolution of the barotropic mode to barotropic experiments on the 500-hPa level. We consider the non-divergent barotropic vorticity equation,

$$\frac{D}{Dt}(\zeta + f) = \left(\frac{\partial}{\partial t} + u \frac{\partial}{\partial x} + v \frac{\partial}{\partial y} \right) \zeta + \frac{\partial f}{\partial y} v = 0. \quad (4.7)$$

This equation expresses conservation of the vertical component of absolute vorticity $\zeta + f$ following the horizontal motion, where ζ is the relative vorticity and f the planetary vorticity. The variation of the Coriolis parameter with latitude is given by $\frac{df}{dy}$. The non-divergent velocity field $\mathbf{v} = (u, v)$ can be written in terms of the streamfunction ψ by letting $\mathbf{v} = \mathbf{e}_3 \times \nabla \psi$, where \mathbf{e}_3 is the unit vector pointing vertically upward. The relative vorticity is then given by $\zeta = \nabla^2 \psi$.

We are interested in the evolution of small perturbations superimposed on a mean flow, which itself is a solution of the barotropic vorticity equation (Eq. (4.7)). Linearisation of Eq. (4.7) about the mean flow yields

$$\left(\frac{\partial}{\partial t} + \bar{u} \frac{\partial}{\partial x} + \bar{v} \frac{\partial}{\partial y} \right) \zeta' + \frac{\partial \bar{\zeta}}{\partial x} u' + \frac{\partial \bar{\zeta}}{\partial y} v' = 0. \quad (4.8)$$

Here mean flow variables are marked by an overbar and perturbation variables by a prime.

The next step in the derivation is the representation of perturbations by wave packets. A wave packet is the envelope of a train of propagating waves. The wave train is assumed to have a dominant local frequency ω , vector wavenumber $\mathbf{k} = (m, n)$ and amplitude a , which are all slowly varying functions of space and time. That is, the derivatives of these variables

to x , y and t are $O(\varepsilon)$, with ε a small parameter. The local frequency and wavenumbers may be derived from a phase function $\theta(\mathbf{x}, t)$,

$$\omega = -\varepsilon^{-1} \frac{\partial \theta}{\partial t}, \quad m = \varepsilon^{-1} \frac{\partial \theta}{\partial x}, \quad n = \varepsilon^{-1} \frac{\partial \theta}{\partial y}, \quad (4.9)$$

which are all $O(1)$. Approximations to solutions of Eq. (4.8) are found by assuming that the wave packets have a WKB form (after Wentzel-Kramers-Brillouin). The WKB expansion may be written (Müller 1978; Zeng 1983a, 1983b)

$$\psi'(\mathbf{x}, t) = \sum_{n=0}^{\infty} \varepsilon^n a_n(\mathbf{x}, t) e^{i\theta(\mathbf{x}, t)/\varepsilon}. \quad (4.10)$$

In our derivation it is assumed that a wave packet can be approximated to lowest order in the expansion, which implies that the plane waves are approximately sinusoidal

$$\psi' = a_0(\mathbf{x}, t) e^{i\theta(\mathbf{x}, t)/\varepsilon}. \quad (4.11)$$

Substitution of Eq. (4.11) into Eq. (4.8) yields, to zero'th order in ε , the dispersion relation for Rossby waves, which relates the local frequency at a fixed point to the local wavenumber,

$$\omega = \mathbf{k} \cdot \bar{\mathbf{v}} + \hat{\omega} = m\bar{u} + n\bar{v} + \frac{n \frac{\partial \bar{\zeta}}{\partial x} - m \left(\frac{\partial \bar{\zeta} + f}{\partial y} \right)}{k^2}, \quad (4.12)$$

with

$$k^2 = m^2 + n^2 \quad (4.13)$$

the total horizontal wavenumber squared. Here $\hat{\omega}$ is the intrinsic frequency, i.e. the local frequency measured by an observer moving with the mean flow, and $\mathbf{k} \cdot \bar{\mathbf{v}}$ is the Doppler shifted frequency. Using the phase function relations (Eq. (4.9)), so-called propagation equations can be written in a canonical form

$$\frac{d\alpha}{dt} = c_{g\alpha} = \frac{\partial \omega}{\partial k_\alpha}, \quad (4.14)$$

$$\frac{dk_\alpha}{dt} = -\frac{\partial \omega}{\partial \alpha}, \quad (4.15)$$

where $\alpha = x, y$ and $k_x = m, k_y = n$ and

$$\frac{d}{dt} = \frac{\partial}{\partial t} + c_{gx} \frac{\partial}{\partial x} + c_{gy} \frac{\partial}{\partial y} \quad (4.16)$$

stands for differentiation moving with the group velocity $\mathbf{c}_g = (c_{gx}, c_{gy})$. The integration of Eq. (4.14) yields the path described by an observer always moving with the local value of the group velocity. This path is known as a ray (see e.g. Lighthill 1978). The change of wavenumber along the ray is described by Eq. (4.15). Using the dispersion relation, we find a set of four coupled equations that can be solved when initial conditions are specified.

$$c_{gx} = \bar{u} + \frac{(m^2 - n^2) \left(\frac{\partial \bar{\zeta} + f}{\partial y} \right) - 2mn \frac{\partial \bar{\zeta}}{\partial x}}{k^4}, \quad (4.17)$$

$$c_{gy} = \bar{v} + \frac{2mn(\frac{\partial \bar{\zeta} + f}{\partial y}) + (m^2 - n^2)\frac{\partial \bar{\zeta}}{\partial x}}{k^4}, \quad (4.18)$$

$$\frac{dm}{dt} = -m\frac{\partial \bar{u}}{\partial x} - n\frac{\partial \bar{v}}{\partial x} + \frac{m}{k^2}\frac{\partial^2 \bar{\zeta}}{\partial x \partial y} - \frac{n}{k^2}\frac{\partial^2 \bar{\zeta}}{\partial x^2}, \quad (4.19)$$

$$\frac{dn}{dt} = -m\frac{\partial \bar{u}}{\partial y} - n\frac{\partial \bar{v}}{\partial y} + \frac{m}{k^2}\left(\frac{\partial^2 \bar{\zeta} + f}{\partial y^2}\right) - \frac{n}{k^2}\frac{\partial^2 \bar{\zeta}}{\partial y \partial x}. \quad (4.20)$$

The number of degrees of freedom is reduced to only four by using the WKB approximation. The price that is paid for this reduction is the loss of linearity. The original *linear* equation ((Eq. 4.8)) is replaced by a set of *nonlinear* equations.

Substitution of Eq. (4.11) into Eq. (4.8) yields to first order in ϵ the so-called transport equation

$$\begin{aligned} & \left(\frac{\partial}{\partial t} + \bar{u}\frac{\partial}{\partial x} + \bar{v}\frac{\partial}{\partial y} \right) k^2 a_0 - (\omega - m\bar{u} - n\bar{v}) \left[2 \left(m\frac{\partial}{\partial x} + n\frac{\partial}{\partial y} \right) a_0 \right. \\ & \left. + a_0 \left(\frac{\partial m}{\partial x} + \frac{\partial n}{\partial y} \right) \right] + \frac{\partial \bar{\zeta}}{\partial x} \frac{\partial a_0}{\partial y} - \left(\frac{\partial \bar{\zeta}}{\partial y} + f \right) \frac{\partial a_0}{\partial x} = 0. \end{aligned} \quad (4.21)$$

The transport equation can be rewritten in a more convenient expression (see appendix B, Eq. (4.75)). We now take the envelope of the packet to be of sinusoidal form

$$a_0 = |a_0| e^{i\alpha(\mathbf{x}, t)}. \quad (4.22)$$

Thus, we can derive equations for the phase α and the amplitude $|a_0|$ of the packet,

$$\frac{d\alpha}{dt} = 0 \quad (4.23)$$

$$\begin{aligned} & k^2 \frac{d|a_0|}{dt} + \frac{|a_0|}{2} \frac{dk^2}{dt} + \frac{k^2 |a_0|}{2} \nabla \cdot \mathbf{c}_g \\ & - |a_0| [\mathbf{k} \cdot (\mathbf{k} \cdot \nabla) \bar{\mathbf{v}} + \mathbf{k} \cdot (\bar{\mathbf{v}} \cdot \nabla) \mathbf{k} - \bar{\mathbf{v}} \cdot (\mathbf{k} \cdot \nabla) \mathbf{k}] = 0. \end{aligned} \quad (4.24)$$

Equation (4.23) states that the envelope of the wave packet propagates with the group velocity and thus moves along a ray. When we assume that the mean flow is time independent, the frequency ω of a packet is constant along a ray (see Bretherton and Garrett 1969),

$$\frac{d\omega}{dt} = 0. \quad (4.25)$$

This implies that the frequency of the packet is conserved.

Multiplying Eq. (4.24) by $|a_0|$ yields an equation for the kinetic energy density $E = \frac{1}{2} k^2 |a_0|^2$ of the packet. On a plane, the energy density equation can be written as

$$\frac{dE}{dt} + E \nabla \cdot \mathbf{c}_g = s_{\alpha\beta} e_{\alpha\beta}, \quad (4.26)$$

where \mathbf{s} is the stress tensor, also called the radiation stress tensor (Longuet-Higgins and Stewart 1964), and \mathbf{e} is the strain rate tensor,

$$\mathbf{s} = \frac{2E}{k^2} \begin{pmatrix} m^2 & mn \\ mn & n^2 \end{pmatrix}, \text{ and } \mathbf{e} = \begin{pmatrix} \frac{\partial \bar{u}}{\partial x} & \frac{1}{2} \left(\frac{\partial \bar{u}}{\partial y} + \frac{\partial \bar{v}}{\partial x} \right) \\ \frac{1}{2} \left(\frac{\partial \bar{u}}{\partial y} + \frac{\partial \bar{v}}{\partial x} \right) & \frac{\partial \bar{v}}{\partial y} \end{pmatrix}. \quad (4.27)$$

When the wave packet propagates on a non-uniform mean flow the strain rate tensor has non-vanishing components. The mean flow does work on the wave packet at a rate $s_{\alpha\beta}e_{\alpha\beta}$ (first shown in the paper of Longuet-Higgins and Stewart (1961) on gravity waves). This implies that the total energy of the wave packet is not conserved and changes as

$$\frac{\partial}{\partial t} \iint_W E \, dx \, dy = \iint_W |a_0|^2 \left[(m^2 - n^2) \frac{\partial \bar{u}}{\partial x} + mn \left(\frac{\partial \bar{u}}{\partial y} + \frac{\partial \bar{v}}{\partial x} \right) \right] dx \, dy. \quad (4.28)$$

Here the packet is denoted by W and is used that the flow is non-divergent. The integral equation states that the packet gains energy through conversion of mean kinetic energy by the Reynolds stress mechanism.

On a sphere (with radius $R = 6.37 \times 10^6 \text{m}$), additional metric terms must be added in order to get the correct expression (see appendix B, Eqs (4.77)-(4.80))

$$\begin{aligned} \frac{\partial}{\partial t} \iint_W E \, dx \, dy = & \iint_W |a_0|^2 \left[(m^2 - n^2) \frac{\partial \bar{u}}{\partial x} + mn \left(\frac{\partial \bar{u}}{\partial y} + \frac{\partial \bar{v}}{\partial x} \right) \right. \\ & \left. + (n^2 - m^2) \frac{\bar{v} \tan \phi}{R} + mn \frac{\bar{u} \tan \phi}{R} \right] dx \, dy. \end{aligned} \quad (4.29)$$

Simmons et al. (1983) have derived an equation for the growth of perturbation kinetic energy (their equation (4)), which can be written in a similar form

$$\begin{aligned} \frac{\partial}{\partial t} \iint E' \, dx \, dy = & \iint \left[(v'^2 - u'^2) \frac{\partial \bar{u}}{\partial x} - u'v' \left(\frac{\partial \bar{u}}{\partial y} + \frac{\partial \bar{v}}{\partial x} \right) \right. \\ & \left. + (u'^2 - v'^2) \frac{\bar{v} \tan \phi}{R} - u'v' \frac{\bar{u} \tan \phi}{R} \right] dx \, dy. \end{aligned} \quad (4.30)$$

Here, the integration is performed globally. The equivalence becomes clear by observing that in the WKB approximation the zonal velocity is proportional to the meridional wavenumber $u' \sim n$ and the meridional velocity to the zonal wavenumber $v' \sim -m$, as can be derived from Eq. (4.11).

In the particular case of a zonal mean flow it can be easily derived from Eq. (4.19) that the zonal wavenumber of the packet remains constant. Zonal asymmetry results in a packet with varying zonal wavenumber. (In a similar way: the meridional asymmetry leads to variations of the meridional wavenumber.) In a zonal background flow total wave action is also conserved. Wave action density A is defined as the relation between the energy density and the intrinsic frequency, thus $A = \frac{E}{\omega}$. One can derive, combining the equation for the kinetic energy density Eq. (4.26) and the evolution equation of the intrinsic frequency

$$\frac{d\omega}{dt} = \frac{\partial \omega}{\partial x} \bar{u} + \frac{\partial \omega}{\partial y} \bar{v} - k_\beta \hat{c}_{g\alpha} \frac{\partial u_\beta}{\partial x_\alpha} \quad (4.31)$$

that for a zonal mean flow the wave action density equation equals

$$\frac{\partial A}{\partial t} + A \nabla \cdot \mathbf{c}_g = 0. \quad (4.32)$$

This implies that total wave action, i.e. the wave action density integrated over the wave packet, is conserved. This result was first obtained by Bretherton and Garrett (1969). It means that total wave energy is proportional to the instantaneous intrinsic frequency; that is, an energetically growing disturbance propagates in a direction such that its intrinsic frequency increases accordingly. In the absence of zonal symmetry this conservation equation is not valid anymore. This is in contrast to a system of propagating surface gravity waves on a slowly varying non-zonal mean flow. Here total wave action is still conserved (see Bretherton and Garrett 1969). However, when variations of the mean vorticity terms are of $O(\varepsilon)$ so that the terms $\frac{\partial \bar{\zeta}}{\partial x} u'$ and $\frac{\partial \bar{\zeta}}{\partial y} v'$ can be neglected in Eq. (4.8) conservation of total wave enstrophy can be derived (Young and Rhines 1980),

$$\frac{\partial P}{\partial t} + P \nabla \cdot \mathbf{c}_g = 0. \quad (4.33)$$

The wave enstrophy density P is equal to the product of the square of total wavenumber and energy density, $P = k^2 E$. The conservation law states that an energetically growing disturbance modifies its shape such that total wavenumber decreases. Conservation of total wave enstrophy breaks down when the terms $\frac{\partial \bar{\zeta}}{\partial x}$ and $\frac{\partial \bar{\zeta}}{\partial y}$ become of $O(1)$. In the simulations these conservation laws are used to calculate the energy of the wave packet.

4.4 Numerical simulations

In this section we will describe some examples of integrations of wave packets using the WKB approximation. For this purpose, a WKB model is written. Computations are performed with a barotropic model and its tangent linear and adjoint model for comparison. A short description of these models will be given first.

4.4.1 Formulation of the WKB model and the barotropic model

The WKB model integrates the set of four coupled nonlinear propagation equations (Eqs (4.17)-(4.20)). A wave packet is characterized by its zonal and meridional wavenumbers (m, n) and its position on the sphere (x, y) . The spatial size of the wave packet is thus reduced to a single point, which represents its center of wave energy. The wavenumbers and position of the packet are prescribed at initial time. Making use of recurrence relations for Legendre functions, see for example Machenhauer (1979), we have derived analytical expressions for the mean velocity and mean relative vorticity and their first and second derivatives to x and y for every position on the sphere. The computations of these variables are exact for mean flows with spectral components up to T21. The group velocity components and the tendency of the wavenumbers can be calculated by substitution of their values into Eqs (4.17)-(4.20). The energy of the packet propagates at the group velocity. Its new position and wavenumbers

after one time step are calculated by integrating the propagation equations. The integration is performed using a fourth-order Runge-Kutta method with a time step of 1 hour. In this way, the path of the packet, i.e. the ray, is found. In order to satisfy the assumptions of the WKB approach we have to consider smooth mean flows. The mean flows we will study are time invariant or have a sufficiently small tendency to be considered time invariant. The computations can then be validated by testing whether the dominant frequency of the packet along the ray is conserved, see Eq. (4.25).

Because we assume that the packet has no spatial size, the integrand of Eq. (4.29) is considered as a measure of the tendency of the total energy of the packet. When total wave action or total wave enstrophy is conserved total wave energy can also be calculated using these conservation properties.

The barotropic model integrates the barotropic vorticity equation (Eq. (4.7)). The model is triangularly truncated at T42, using a time step of 30 minutes, or at T85, using a time step of 15 minutes. Orography and forcing are not included. The tangent linear model describes the linear evolution of small perturbations along a reference forecast, which will be time invariant in the following experiments. The adjoint of the tangent linear model is computed with respect to the squared norm inner product.

4.4.2 Zonal mean background flow

To introduce the WKB method we will first describe the development of a wave packet superimposed on a zonal mean flow. Zonal mean flows are stationary solutions of Eq. (4.7). A zonal mean flow with a reasonably realistic velocity profile is obtained by zonally averaging the low-sensitive cluster mean LS3, which is shown in Fig. 5c of chapter 3 (see also Fig. 4.9). This LS3 pattern is nearly zonal. In Fig. 4.6 the zonally averaged mean flow is shown. The velocity field has a jet-like structure, with a maximum velocity equal to 18.5 m s^{-1} at 35°N . We have located the wave packet in the shear stream centered at 22°N and 62°W , its envelope having the sinusoidal form of Eq. (4.22). The dominant zonal and meridional wavelengths are both equal to about 4500 km ($\lambda_m = \lambda_n = \lambda$), so that $m = n = \frac{2\pi}{\lambda} \simeq 1.4 \times 10^{-6} \text{ m}^{-1}$. In Fig. 4.6 snapshots of the evolution of the packet are shown at 24 hourly intervals. The integration is performed with the T85 version of the tangent linear model. The packet propagates northeastward in the direction of the jet maximum. The group velocity, i.e. the velocity at which the energy or envelope propagates, is larger than the phase velocity, i.e. the velocity of the individual crests. The crests appearing in front of the packet are a clear sign of this. The tilt of the packet, initially from northwest to southeast, vanishes the first few days due to an increase of the meridional wavelength. After 72 hours the packet splits into two parts. The southern part of the packet bends southward and acquires a tilt from southwest to northeast. The northern part passes through the jet maximum. The tilt changes again from northwest to southeast.

In Fig. 4.7 integrations are repeated with the WKB model. In the integrations two packets are superimposed on the same zonal mean flow. The packets have slightly different initial positions and are integrated for 10 days. We will show that both packets describe parts of the evolution of the perturbation. Every 24 hours the position of the packet is marked by a vector and circle. The slope of the vector reflects the ratio of the meridional

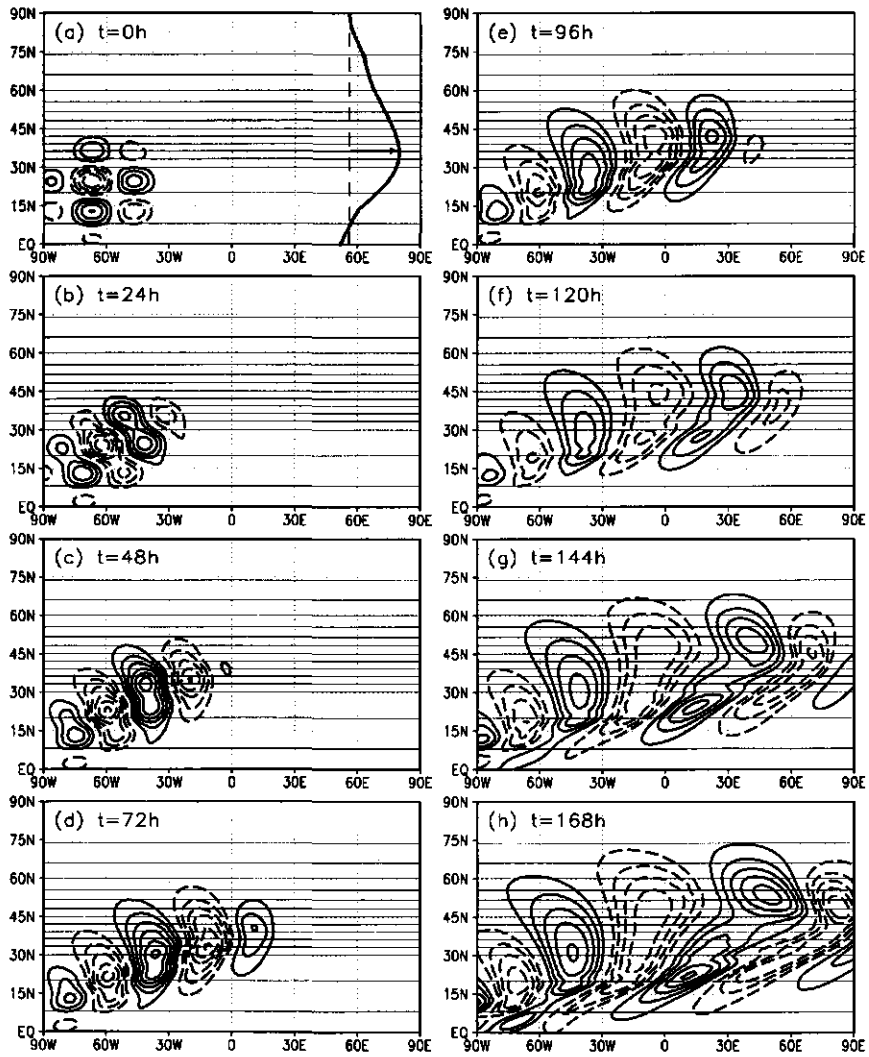


Figure 4.6: Development of a wave packet superimposed on a zonal mean flow. The velocity field has a jet-like structure shown in panel (a), with a maximum of 18.5 m s^{-1} at 35°N . The initial dominant zonal and meridional wavelengths are equal to about 4500 km. The evolution is shown every 24 hours. The fields are given in streamfunction. The contour interval is equal to $5.9 \times 10^6 \text{ m}^2 \text{ s}^{-1}$ for the zonal mean flow and arbitrary for the perturbation.

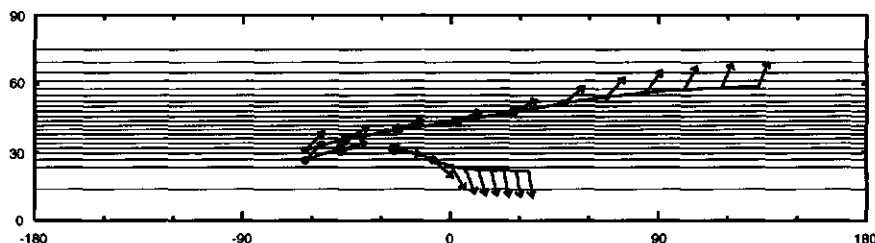


Figure 4.7: Development of two wave packets with slightly different initial positions superimposed on a zonal mean flow. The velocity field has a jet-like structure shown in Fig. 4.6. The initial zonal and meridional wavelengths are equal to about 4500 km. Every 24 hours the location of the packets is marked by a vector and circle. The slope of the vector reflects the proportion of the meridional to the zonal wavenumber. The size of the circle reflects the energy of the packet.

wavenumber to the zonal wavenumber of the packet. A vertical (horizontal) slope indicates that the zonal (meridional) wavenumber is zero. When the wavenumbers are equal the slope is 45° . The size of the circle marks the energy of the packet, i.e. the radius increases a factor of 2 when the energy doubles. The horizontal wavenumber and total wave action are conserved, because of the zonal symmetry. The southern packet bends to the south before it reaches the jet maximum. Near the jet maximum the intrinsic frequency $\hat{\omega}$ is largest and therefore, due to the conservation of waveaction, the energy of the packet is maximal. The turning point is located at the position where $n = 0$ ($\lambda_n \rightarrow \infty$), see Eq. (4.18). Here the meridional velocity $c_{gy} \simeq \frac{2mn\partial f}{k^4}$ reverses sign. In the neighborhood of the turning point the WKB method is obviously invalid. However the terms in Eqs (4.17)-(4.20) in which the meridional wavenumber is involved become very small near the turning point and can be neglected. The integration can thus be continued. The packet loses energy and approaches a critical line at which all energy will be lost (both E as $\hat{\omega}$ become zero). The northern packet passes through the jet maximum before $n = 0$. At the jet maximum the term $\frac{\partial \bar{u}}{\partial y}$ reverses sign, so that n increases again before it reaches $n = 0$, see Eq. (4.20). The ray has therefore no turning point and approaches a critical line further to the north where all energy has been lost to the mean flow.

We are not interested in the asymptotic behaviour of the packet, as the WKB approach fails in this limit (due to viscosity and nonlinearity). Moreover, by this time the packet is spread over the whole northern hemisphere. Its description by just a single wavenumber pair has become unrealistic. As can be observed by comparing Figs. 4.6 and 4.7, both rays in Fig. 4.7 describe parts of the development of the wave packet in Fig. 4.6. We will also encounter this in the following examples because the spatial size of the perturbations is not negligible entirely. The main conclusions that can be drawn from the WKB approach are that it predicts an amplification of the wave energy when wave crests lean into the mean shear ($\frac{m}{n} > 0$ where $\frac{\partial \bar{u}}{\partial y} > 0$, or $\frac{m}{n} < 0$ where $\frac{\partial \bar{u}}{\partial y} < 0$) and a decrease when they lean along the mean shear ($\frac{m}{n} < 0$ where $\frac{\partial \bar{u}}{\partial y} > 0$, or $\frac{m}{n} > 0$ where $\frac{\partial \bar{u}}{\partial y} < 0$). In the first case, the meridional

wavelength increases and the packet propagates toward the jet maximum. In the second case, the meridional wavelength decreases and the packet propagates into weaker westerly flow. We will concentrate on such mechanisms in the description of the development of wave packets superimposed on non-zonal mean flows. The description can only be qualitative because of the assumptions made in the WKB method. The goal is not to develop an accurate forecast model but to interpret the main processes using simple relations.

4.4.3 Non-zonal background flows

Rossby-Haurwitz waves

We will proceed with background flows that are non-zonal. First, we will consider packets superimposed on stationary Rossby-Haurwitz waves. Rossby-Haurwitz waves are solutions of the barotropic vorticity equation of the form

$$\psi(\lambda, \mu) = A \operatorname{Re}\{Y_{M,N}(\lambda, \mu)\} - c\mu, \quad (4.34)$$

where $Y_{M,N}$ is the spherical harmonic function of order M and degree N . $\operatorname{Re}\{z\}$ denotes the real part of the complex number z , λ is longitude and $\mu = \sin \phi$, where ϕ is latitude. In our application we choose $M = 3$, $N = 6$, $c = 0.05$ and vary the amplitude A . For these values, the flow possesses characteristics of sensitive periods like a strong jet stream upstream of a diffuent flow (see chapter 3) on the same spatial scale. By enlarging A the strength of the jet becomes stronger and the diffuence of the flow larger. In chapter 3 it was also shown that the initially fastest growing perturbations (regional singular vectors) are formed by eddies that have a zonally elongated structure, see Fig. 3.13. These perturbations can be represented reasonably well by wave packets that have a large dominant zonal wavelength and a small dominant meridional wavelength.

In Fig. 4.8 we show several evolutions of wave packets, which may be compared to the evolution of the regional singular vectors (RSVs) shown in Fig. 3.13. In Figs. 4.8a to 4.8c we have set the initial zonal wavenumber equal to 10000 km and the meridional wavenumber to 1000 km. The packets are located at three different positions, so that we have one packet located at the jet maximum and two located just to the north and south of it. In Fig. 4.8a $A = 0.1$ so that the mean flow is nearly zonal with a very weak jet. The zonal wavenumber m is not conserved anymore but changes slowly. The energy of the packet grows by maximally 5%. In Fig. 4.8b $A = 1.0$ and in Fig. 4.8c $A = 2.5$. In these diffuent background flows $\frac{\partial u}{\partial x} < 0$ is the dominant term. The energy of the packet grows the first 48 to 72 hours, since $m < n$ (see Eq. (4.28)). In the second part, $\frac{\partial u}{\partial x} > 0$ so that the packet loses energy where $m < n$ (angle of vector with meridional direction is smaller than 45°) and gains energy where $m > n$ (angle of vector with zonal direction is smaller than 45°). At initial time m is very small and even decreases and goes through zero during the first 24 hours of the most northern path. The prediction of the value of m is incorrect (one can compare this to the turning point in a zonal flow where $n \rightarrow 0$), but there all terms in which m appears can be neglected, the (temporary) incorrect value of m does not affect the evolution of the group velocity, meridional wavenumber, the energy and even its own evolution. Since zonal variations in the mean flow are small at the initial position of the packet, the small value of m does

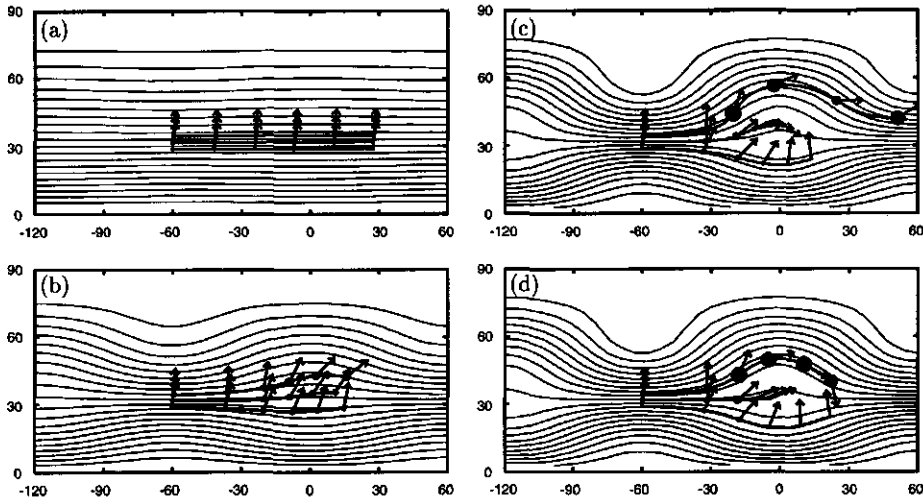


Figure 4.8: Wave packets superimposed on Rossby-Haurwitz waves with amplitudes equal to 0.1 (a), 1.25 (b), 2.5 (c) and 2.5 (d). The initial zonal and meridional wavelengths are equal to about 10000 km and 1000 km, respectively (panels (a), (b) and (c)). In panel (d) the initial zonal and meridional wavelengths are 10 times smaller. Every 24 hours the location of the packets is marked.

not directly imply that the WKB approximations fail. After this initial period, the WKB equations predict an increase of m and decrease of n . The zonal size of the packet decreases and the meridional size increases. This goes together with an increase of the energy. It appears that the wave enstrophy is nearly conserved: k^2 decreases as a result of a decreasing 'aspect ratio' (m increases and n decreases), which leads to an increase of E . The energy of the northern packet increases more than that of the southern packet. The main difference in the mean flow is that the term $\frac{\partial \bar{v}}{\partial x}$ is positive in the former case and negative in the second one. Therefore, the zonal wavenumber m of the northern packet increases more slowly (Eq. 4.19), by which the packet gains more energy because the negative contribution of $m^2 \frac{\partial \bar{v}}{\partial x}$ in Eq. (4.28) is smaller during a longer time interval.

In Fig. 4.8c it can be observed that the path of the northern packet deviates from the mean flow. This occurs when m becomes of the same order of n so that the second term of the meridional group velocity component $\frac{2mn \frac{\partial \bar{v}}{\partial y} + f}{k^4}$ can not be disregarded in comparison to \bar{v} . However, at that time the size of the packet has become so large compared to deviations in the mean flow, that the WKB approximation is not valid anymore. In Fig. 4.8d the evolutions of wave packets superimposed on the same mean flow ($A = 2.5$) are shown but with a meridional and zonal wavelength each 10 times smaller. The evolutions of the packets are comparable to the ones in Fig. 4.8c. The main difference is that the packets follow the mean flow, because of the smaller size of the packets. As can be deduced by looking at the group velocity components, only when the total wavenumber k is small the group velocity will differ substantially from the background velocity field. However, in this range

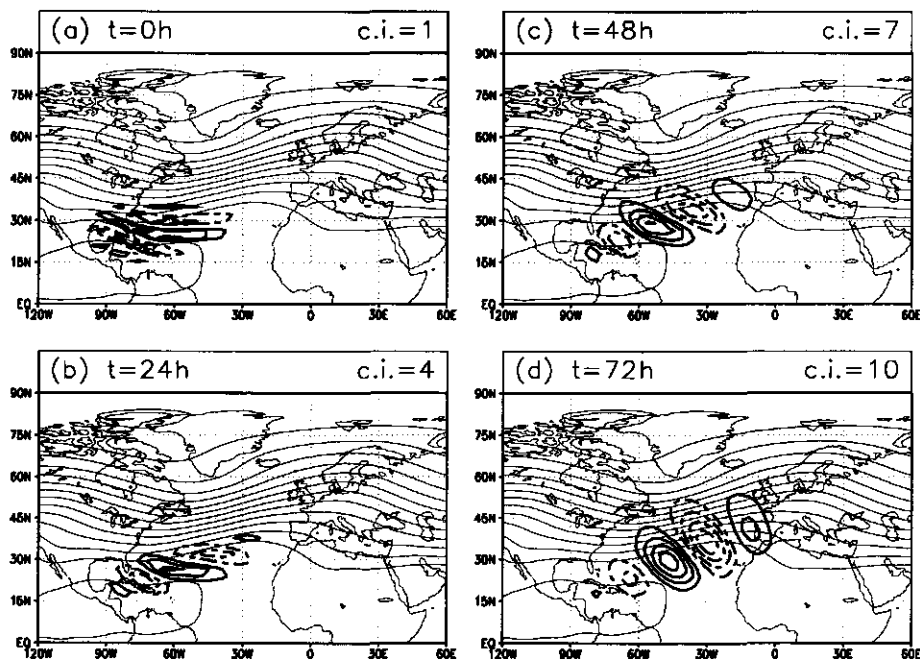


Figure 4.9: Evolution of the fastest growing RSV of the low-sensitive cluster mean LS3. Fields are in streamfunction. The contour interval of the background flow is $5.9 \times 10^6 \text{ m}^2 \text{ s}^{-1}$. The contour interval of the RSV is given in each panel.

the validity of the WKB approximation becomes questionable.

Low-sensitive background flow

We will now proceed with more realistic large-scale mean flows. We will first consider the low-sensitive cluster mean LS3 (Fig. 5c of chapter 3) and then the high-sensitive cluster mean HS1 (Fig. 5e of chapter 3). The cluster means were found by applying a cluster algorithm (see Michelangeli 1995 and chapter 3) to a high- and low-sensitivity set, each containing 1000 flow patterns. The sensitivity is defined again with respect to the initial conditions for onset of BL and SZF regimes. The flow of LS3 is fairly zonal (Fig. 4.9), the flow of HS1 is characterized by a stronger jet stream upstream of a diffluent flow (Fig. 4.11). For both cluster means we have computed regional singular vectors (RSVs) that maximize rms error over the Atlantic-European area ($10^\circ\text{N} - 85^\circ\text{N}$ and $90^\circ\text{W} - 60^\circ\text{E}$, see also Eq. (3.10)). The computations are performed with the barotropic model at T42. The cluster mean LS3 and its fastest growing RSV are shown in Fig. 4.9a. The RSV is located to the south of the jet maximum where the shear is largest. The RSV can be represented quite well by a wave packet with a small zonal wavenumber and a large meridional wavenumber. During the evolution the tilt decreases. The energy of the packet increases by the Reynolds stress

mechanism already described above. The sensitivity for BL/SZF onset is low because the sensitivity is located here more to the south. Furthermore the perturbation exists of zonally instead of meridionally oriented wave trains.

Results of ray tracing of wave packets superimposed on the flow pattern LS3 are shown in Fig. 4.10. Four different initial positions are chosen. The packets have a structure similar

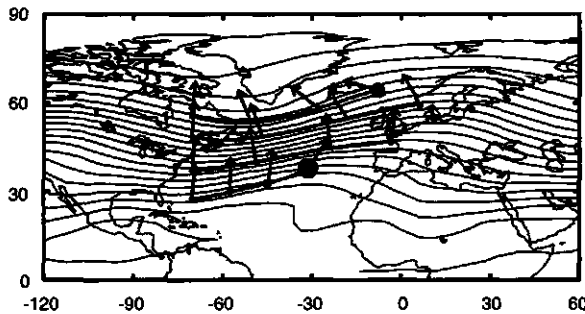


Figure 4.10: Development of four wave packets with only different initial position superimposed on the low-sensitive cluster mean LS3. The initial zonal wavelength is 10000 km, the initial meridional wavelength is 1000 km. Every 24 hours the location of the packets is marked by a vector and circle.

to the RSV, a large initial zonal wavelength (10000 km) and a small initial meridional wavelength (1000 km). One directly observes that the sensitivity of the flow is not in the neighborhood of the jet maximum but southward and northward of it. Near the jet maximum wave packets are being advected by the mean flow. The orientation of the packet follows the weakly meandering mean flow, so that m and n are not conserved. However, wave enstrophy is (nearly) conserved. During the first 48 hours the packet following the most southern ray conserves its wave enstrophy. Its tilt changes by an increasing meridional wavelength, which results in an increasing energy. The zonal wavelength decreases only after 48 hours. Because n continues to decrease, k becomes so small that the group velocity differs from the velocity field of the mean flow. By this time, the WKB approximation again loses its validity. Still, it predicts rather well that the energy will grow and the ray will bend northward. The most northern packet gains energy by tilting in the opposite direction, because the shear is opposite.

High-sensitive background flow

In Fig. 4.11 the evolution of the fastest growing RSV superimposed on the high-sensitivity cluster mean HS1 is shown. The RSV can again be represented quite well by a wave packet with a very small zonal wavenumber and a large meridional wavenumber. (Notice that the barotropic part of the optimal perturbation (Fig. 4.1d) has a similar elongated structure.) The northern part of the RSV has initially much less amplitude but after 72 hours it has about the same strength as the southern part. This indicates that wave energy moves northward.

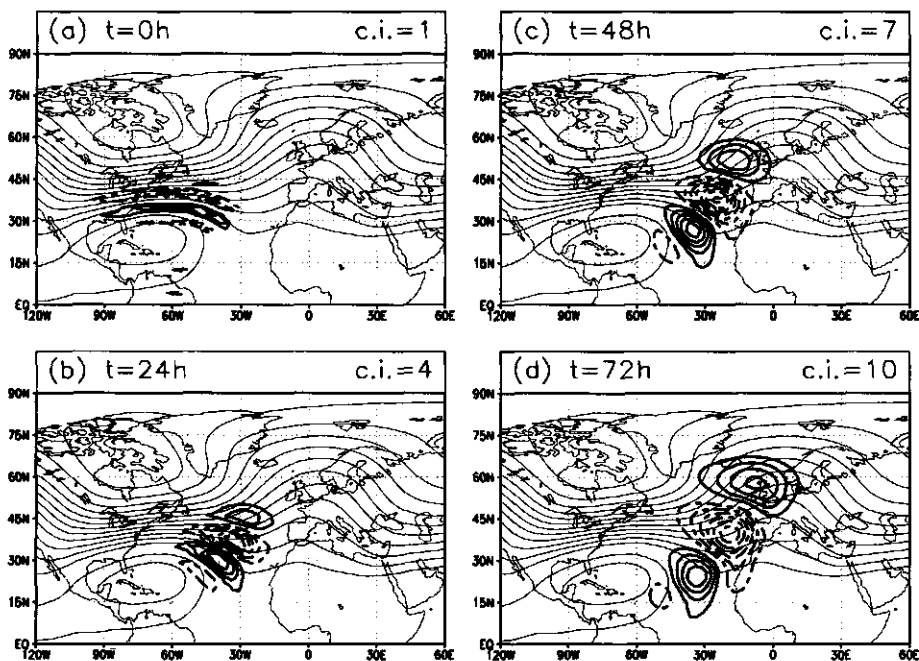


Figure 4.11: Evolution of the fastest growing regional SV of the high-sensitive cluster mean HS1. Fields are in streamfunction. The contour interval of the background flow is $5.9 \times 10^6 \text{ m}^2 \text{ s}^{-1}$. The contour interval of the RSV is given in each panel.

This could be partly due to the mechanism described at the end of the discussion of Fig. 4.8.

The WKB results are shown in Fig. 4.12. It appears that the ratio of m and n must be very small, otherwise the tilt of the packet will change so fast that when it enters the diffuent area it is not able to gain energy anymore. Therefore, the initial zonal wavelength is chosen to be 25000 km and the meridional wavelength 1000 km. The packets gain energy initially through the shear in the jet stream and after that by the diffuence of the flow. In fact, the same mechanisms that are responsible for the energy growth of the packets in Figs. 4.7 and 4.8 apply here. Due to the shear and the diffuence n decreases, whereas m increases in the diffuent area. The WKB approach predicts that the wave energy propagates along a ray that bends to the north with respect to the mean flow. This has also been observed in Fig. 4.11. However, the WKB approximations exceed their validity range at the end of the integration period.

4.5 Summary and discussion

In this chapter, we have focused on the evolution of optimal transition perturbations. These perturbations optimally trigger a blocking (BL) or strong zonal flow (SZF) regime within

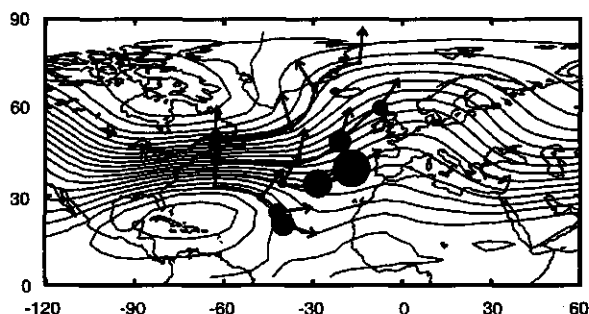


Figure 4.12: Development of four wave packets with only different initial position superimposed on the high-sensitive cluster mean HS1. The initial zonal wavelength is 25000 km, the initial meridional wavelength is 1000 km. Every 24 hours the location of the packets is marked by a vector and circle.

3 days. The computation of the perturbations has been described extensively in chapter 2 (Oortwijn and Barkmeijer 1995). In chapter 3 (Oortwijn 1998a) we have mainly investigated characteristics of sensitive and insensitive reference flows, with sensitivity defined as sensitivity to the initial conditions for onset of BL and SZF regimes. Here, we have investigated the development of perturbations superimposed on background flows that possess these characteristics.

First, we have considered the full evolution of a fast growing optimal transition perturbation during a sensitive three-day period. This example is illustrative for many evolutions of optimal transition perturbations. The evolution is studied using a vertical mode decomposition and a normal mode decomposition of the perturbation. We can identify two phases in the evolution of the perturbation. At initial time, the perturbation can be described by a small-scale wave train with a strong baroclinic structure. However, its barotropic part has a zonally elongated large-scale structure. After 24 hours the barotropic part almost completely dominates the pattern. During this first phase, available potential energy is converted to kinetic energy. The growth is strongly non-modal: the perturbation is a combination of many (amplifying as well as decaying) NMs that point into directions that partially cancel each other. Due to different oscillation and amplification properties of the NMs, the growth is faster than exponential. After 24 hours, the development of the perturbation is mainly barotropic, extracting kinetic energy from the reference flow. During this second phase, the growth is still non-modal but not as strong as before. The growth is mainly due to amplification of the NMs.

In the second part of this chapter we have described the barotropic evolution using a WKB approximation. This approach is based on the following assumptions: the perturbation can be represented in the form of a Rossby wave packet, and a scale separation between the perturbation and the background flow can be made. These assumptions are rather strong and will not be valid for realistic basic states. In particular, the initial baroclinic development of an optimal perturbation depends strongly on the small-scale structure of the background

flow. This initial inverse energy cascade can therefore not be described within the context of the WKB method. However, the barotropic part of the evolution and the structure of the perturbation seem to depend more on the large-scale structure of the background flow. As is shown in chapter 3 (Oortwijn 1998a), the sensitivity is higher when the background flow has a strong jet stream to the west of a diffluent flow. We have therefore focused on the barotropic part of the development. Because of its assumptions, the WKB approach is meant as a diagnostic tool to interpret the evolution of perturbations qualitatively. Thus, we are able to determine which terms in the equations are important for the development of the packets.

We have considered idealized basic states in order to satisfy the WKB assumptions. It appears that the WKB method describes the development of the perturbations reasonably well, although the scale separation between the perturbations and the background flow is fairly small. First the evolution of a perturbation on a zonal mean flow is described. In this case, total wave action of a packet is conserved. The growth of an optimal perturbation embedded in a zonal background flow may be interpreted by this conservation law: a wave packet that propagates into a stronger zonal flow acquires a larger intrinsic frequency and therefore gains energy. However, sensitive periods with respect to BL onset are essentially non-zonal and cannot be interpreted therefore by conservation of total wave action. Under certain conditions, conservation of total wave enstrophy may be derived for wave packets embedded in a non-zonal background flow: a growing disturbance modifies its shape such that its total wavenumber decreases. The evolution of the perturbations may be understood partly by this concept. However, when the scale of a packet becomes in the order of variations in the mean flow, total wave enstrophy will not be conserved anymore. We have focused on the evolution of zonally elongated wave packets located initially in or near the jet stream that propagate into a diffluent area. For these packets wave enstrophy is not conserved anymore in the latter part of the integration when the meridional wavenumber becomes of the order of the meridional variations in the background flow. In this limit, the WKB approach is not valid anymore, although it still predicts rather well that the energy will grow and that the packet propagates along a ray that bends northward with respect to the background flow. When variations in the mean flow appear on smaller scales, like for daily flow patterns, conservation of total wave enstrophy will not apply anymore. However, the evolution of the barotropic mode of the considered fast growing perturbation on an unsmoothed background flow, indicates that the mechanisms we have studied with the WKB approach still play an important role in these more realistic situations.

A Vertical modes of the T21QG model

Vertical modes can be found when horizontal and vertical coordinates are separated. Formally, the streamfunction ψ and planetary vorticity f can be expanded by summing over vertical modes (Flierl 1978; Neven 1993; Ambaum 1997),

$$\psi(\lambda, \phi, p, t) = \sum_{m=0}^{\infty} G_m(p) \tilde{\psi}_m(\lambda, \phi, t), \quad (4.35)$$

$$f(\phi) = \sum_{m=0}^{\infty} \tilde{G}_m(p) \tilde{f}_m(\phi). \quad (4.36)$$

When the vertical modes $\tilde{G}_m(p)$ satisfy the Sturm-Liouville equation

$$f_0^2 \frac{\partial}{\partial p} \left(\frac{1}{\sigma} \frac{\partial}{\partial p} \right) \tilde{G}_m - \chi_m \tilde{G}_m = 0 \quad (4.37)$$

with boundary conditions

$$\left(\frac{\partial \tilde{G}_m}{\partial p} \right)_{p=p_h} = \left(\frac{\partial \tilde{G}_m}{\partial p} \right)_{p=p_l} = 0, \quad (4.38)$$

then the quasigeostrophic potential vorticity

$$q = \nabla^2 \psi + f + f_0^2 \frac{\partial}{\partial p} \left(\frac{1}{\sigma} \frac{\partial}{\partial p} \right) \psi \quad (4.39)$$

can be written as

$$q(\lambda, \phi, p, t) = \sum_{m=0}^{\infty} \tilde{G}_m(p) \tilde{q}_m(\lambda, \phi, t), \quad (4.40)$$

with

$$\tilde{q}_m(\lambda, \phi, t) = \nabla^2 \tilde{\psi}_m(\lambda, \phi, t) + \tilde{f}_m(\phi) + \chi_m \tilde{\psi}_m(\lambda, \phi, t). \quad (4.41)$$

The eigenfunctions $\tilde{G}_m(p)$ can be chosen such that they form an orthonormal system with respect to the inner product

$$\frac{1}{p_h - p_l} \int_{p_l}^{p_h} \tilde{G}_m(p) \tilde{G}_n(p) dp = \delta_{mn}. \quad (4.42)$$

The imposed boundary conditions (Eq. (4.38)) imply that the vertical velocity on the bottom of the atmosphere ($p = p_l$) and at the top of the atmosphere ($p = p_h$) is equal to zero, i.e. the stratosphere acts as a rigid lid over the troposphere. The expression for the coefficient \tilde{q}_m (Eq. (4.41)), which is associated with the m 'th vertical mode, is such that it depends only on the m 'th streamfunction coefficient ψ_m . We will use this property in the derivation of the vertical modes of the T21QG model.

The T21QG model integrates prognostic equations for quasigeostrophic potential vorticity q at 200, 500, and 800 hPa (levels 1,2 and 3 respectively). The quasigeostrophic potential vorticity is related to the streamfunction as

$$\begin{aligned} q_1 &= \nabla^2 \psi_1 + f - R_1^{-2}(\psi_1 - \psi_2), \\ q_2 &= \nabla^2 \psi_2 + f + R_1^{-2}(\psi_1 - \psi_2) - R_2^{-2}(\psi_2 - \psi_3), \\ q_3 &= \nabla^2 \psi_3 + f \left(1 + \frac{h}{H_0} \right) + R_2^{-2}(\psi_2 - \psi_3). \end{aligned} \quad (4.43)$$

Here, f is the planetary vorticity, R_1 ($= 700$ km) and R_2 ($= 450$ km) are Rossby radii of deformation, h is the orographic height and H_0 ($= 9$ km) is a scale height. (In the model,

lengths are expressed in units of the radius of the earth $R = 6.37 \times 10^6 \text{m.}$) The discretization into three levels implies in fact that the streamfunction (and similarly the quasigeostrophic potential vorticity) is written as a sum of contributions of the three levels,

$$\psi = G_1\psi_1 + G_2\psi_2 + G_3\psi_3, \quad (4.44)$$

where the vertical modes have the structure as given in Fig. 4.13a. We can denote these modes by

$$G_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad G_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad G_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \quad (4.45)$$

In order to find the vertical coefficients \tilde{q}_m , we need to find a linear operator A_{mn} that transforms

$$\tilde{\psi}_{m-1} = \sum_{n=1}^3 A_{mn}\psi_n, \quad m = 1, 2, 3 \quad (4.46)$$

$$\tilde{q}_{m-1} = \sum_{n=1}^3 A_{mn}q_n, \quad m = 1, 2, 3 \quad (4.47)$$

such that the expression for \tilde{q}_{m-1} can be written as in Eq. (4.41). First, we rewrite Eq. (4.43) as

$$q_m = \nabla^2\psi_m + f_m + B_{mn}\psi_n, \quad m = 1, 2, 3 \quad (4.48)$$

where

$$f_m = f \begin{pmatrix} 1 \\ 1 \\ 1 + \frac{h}{H_0} \end{pmatrix}, \text{ and } B_{mn} = \begin{pmatrix} -R_1^{-2} & R_1^{-2} & 0 \\ R_1^{-2} & -(R_1^{-2} + R_2^{-2}) & R_2^{-2} \\ 0 & R_2^{-2} & -R_2^{-2} \end{pmatrix} \quad (4.49)$$

Substitution of Eq. (4.41) in Eq. (4.47) and using Eq. (4.46) yields an eigenvalue equation

$$V_m B^T = \chi_m V_m, \quad m = 1, 2, 3. \quad (4.50)$$

where the left eigenvector V_m is equal to the m 'th row of A_{mn} . The solution of the eigenvalue problem leads to the following relations between the streamfunction at model levels and the streamfunction associated with the vertical modes,

$$\psi_{BAR} = \tilde{\psi}_0 = 0.577 (\psi_1 + \psi_2 + \psi_3), \quad (4.51)$$

$$\psi_{BC1} = \tilde{\psi}_1 = -0.799 \psi_1 + 0.254 \psi_2 + 0.545 \psi_3, \quad (4.52)$$

$$\psi_{BC2} = \tilde{\psi}_2 = 0.168 \psi_1 - 0.776 \psi_2 + 0.608 \psi_3. \quad (4.53)$$

The corresponding vertical modes with eigenvalues $\chi_0 = 0$, $\chi_1 = -109$ and $\chi_2 = -465$ are equal to

$$\tilde{G}_0 = \begin{pmatrix} 0.577 \\ 0.577 \\ 0.577 \end{pmatrix}, \quad \tilde{G}_1 = \begin{pmatrix} -0.799 \\ 0.254 \\ 0.545 \end{pmatrix}, \quad \tilde{G}_2 = \begin{pmatrix} 0.168 \\ -0.776 \\ 0.608 \end{pmatrix}, \quad (4.54)$$

such that

$$\psi = \tilde{G}_0 \tilde{\psi}_0 + \tilde{G}_1 \tilde{\psi}_1 + \tilde{G}_2 \tilde{\psi}_2. \quad (4.55)$$

The vertical structure of the modes is shown in Fig. 4.13b. We will call the streamfunction associated with the vertical mode with $m = 0$ the barotropic mode ψ_{BAR} because the vertical structure is constant at all model levels. The vertical structures associated with the so-called first and second baroclinic modes ψ_{BC1} and ψ_{BC2} change their sign once and twice, respectively.

The vertical modes derived above correspond to a discretized version of the Sturm-Liouville equation (Eq. (4.37)). The Sturm-Liouville equation reduces to an eigenvalue equation in which the vertical modes \tilde{G}_m are eigenfunctions of B . The orthonormalization property (4.42) reduces to a vector product of vertical modes

$$\tilde{G}_m \cdot \tilde{G}_n = \delta_{mn}. \quad (4.56)$$

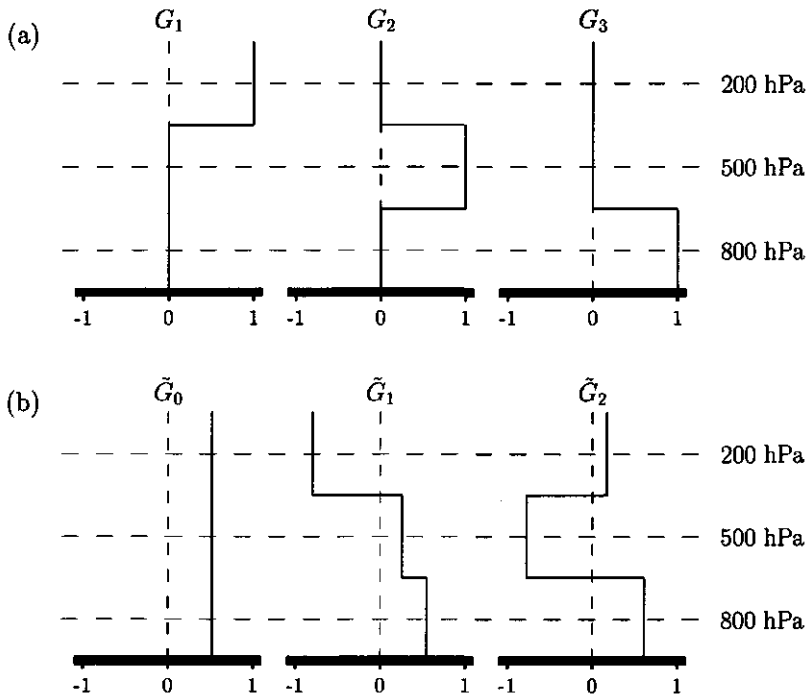


Figure 4.13: (a) Vertical modes associated with the three-level model and (b) the vertical modes that are associated with a discretized version of the Sturm-Liouville equation (Eq. (4.37)).

B Derivation of the KE-equation

We will derive the KE-equation (Eq. (4.26)) in general orthogonal coordinates. The linearised barotropic vorticity equation can be written as

$$\frac{\partial \zeta'}{\partial t} + \bar{\mathbf{v}} \cdot \nabla \zeta' + \mathbf{v}' \cdot \nabla (\bar{\zeta} + f) = 0. \quad (4.57)$$

Furthermore, we use the following expressions

$$\mathbf{k} = \frac{1}{\varepsilon} \nabla \theta, \text{ and thus } \nabla \times \mathbf{k} = 0, \quad (4.58)$$

$$\psi' = a_0 e^{\frac{i\theta}{\varepsilon}}, \quad (4.59)$$

$$\mathbf{v}' = \mathbf{e}_3 \times \nabla \psi'. \quad (4.60)$$

The gradient of ψ' is written

$$\nabla \psi' = \nabla (a_0 e^{\frac{i\theta}{\varepsilon}}) = (\nabla a_0) e^{\frac{i\theta}{\varepsilon}} + i \frac{\nabla \theta}{\varepsilon} a_0 e^{\frac{i\theta}{\varepsilon}}. \quad (4.61)$$

It follows that the perturbation relative vorticity is given by

$$\begin{aligned} \zeta' &= \nabla \cdot \nabla \psi' = \\ &= \nabla^2 a_0 e^{\frac{i\theta}{\varepsilon}} + \frac{2i}{\varepsilon} (\nabla a_0 \cdot \nabla \theta) e^{\frac{i\theta}{\varepsilon}} - a_0 \frac{(\nabla \theta)^2}{\varepsilon^2} e^{\frac{i\theta}{\varepsilon}} + \frac{i}{\varepsilon} a_0 \nabla^2 \theta e^{\frac{i\theta}{\varepsilon}} \\ &\simeq -k^2 a_0 e^{\frac{i\theta}{\varepsilon}} + i \left\{ 2(\mathbf{k} \cdot \nabla a_0) e^{\frac{i\theta}{\varepsilon}} + a_0 \nabla \cdot \mathbf{k} e^{\frac{i\theta}{\varepsilon}} \right\}, \end{aligned} \quad (4.62)$$

when terms of $O(\varepsilon^2)$ are omitted. To $O(1)$ we now find for Eq. (4.57):

$$i\omega k^2 a_0 e^{\frac{i\theta}{\varepsilon}} - ik^2 a_0 \bar{\mathbf{v}} \cdot \mathbf{k} e^{\frac{i\theta}{\varepsilon}} + ia_0 (\mathbf{e}_3 \times \mathbf{k}) \cdot \nabla (\bar{\zeta} + f) e^{\frac{i\theta}{\varepsilon}} = 0. \quad (4.63)$$

So that

$$\omega = \mathbf{k} \cdot \bar{\mathbf{v}} + \frac{(\mathbf{e}_3 \times \nabla (\bar{\zeta} + f)) \cdot \mathbf{k}}{k^2}, \quad (4.64)$$

$$\mathbf{c}_g = \frac{\partial \omega}{\partial \mathbf{k}} = \bar{\mathbf{v}} + \frac{\mathbf{e}_3 \times \nabla (\bar{\zeta} + f)}{k^2} - \frac{2\mathbf{k}(\mathbf{e}_3 \times \nabla (\bar{\zeta} + f)) \cdot \mathbf{k}}{k^4}, \quad (4.65)$$

where \mathbf{e}_3 is the local vertical unit vector.

To $O(\varepsilon)$ Eq. (4.57) yields the transport equation (see Eq. (4.21)):

$$\begin{aligned} -\left(\frac{\partial}{\partial t} + \bar{\mathbf{v}} \cdot \nabla\right) k^2 a_0 + (\omega - \mathbf{k} \cdot \bar{\mathbf{v}}) [2(\mathbf{k} \cdot \nabla) a_0 + a_0 \nabla \cdot \mathbf{k}] \\ + (\mathbf{e}_3 \times \nabla a_0) \cdot \nabla (\bar{\zeta} + f) = 0. \end{aligned} \quad (4.66)$$

We will now have to carry out some manipulations with the equations. Combining Eqs (4.64) and (4.65) yields

$$k^2 \mathbf{c}_g = k^2 \bar{\mathbf{v}} + \mathbf{e}_3 \times \nabla (\bar{\zeta} + f) - 2(\omega - \mathbf{k} \cdot \bar{\mathbf{v}}) \mathbf{k} \quad (4.67)$$

Using the rules for vector operations, we get

$$\begin{aligned}\nabla \cdot (k^2 \mathbf{c}_g) &= \bar{\mathbf{v}} \cdot \nabla k^2 + k^2 \nabla \cdot \bar{\mathbf{v}} + \nabla(\bar{\zeta} + f) \cdot \nabla \times \mathbf{e}_3 - \mathbf{e}_3 \cdot \nabla \times \nabla(\bar{\zeta} + f) \\ &\quad - 2(\omega - \mathbf{k} \cdot \bar{\mathbf{v}}) \nabla \cdot \mathbf{k} - 2\mathbf{k} \cdot \nabla(\omega - \mathbf{k} \cdot \bar{\mathbf{v}}) \\ &= \bar{\mathbf{v}} \cdot \nabla k^2 - 2(\omega - \mathbf{k} \cdot \bar{\mathbf{v}}) \nabla \cdot \mathbf{k} - 2\mathbf{k} \cdot \nabla(\omega - \mathbf{k} \cdot \bar{\mathbf{v}}),\end{aligned}\quad (4.68)$$

because the mean flow is non-divergent and the rotation of the local vertical unit vector and the rotation of a gradient are zero. Next, we use that

$$-2\mathbf{k} \cdot \nabla \omega = -2\mathbf{k} \cdot \nabla \left(-\frac{1}{\varepsilon} \frac{\partial \theta}{\partial t} \right) = 2\mathbf{k} \cdot \frac{\partial \mathbf{k}}{\partial t} = \frac{\partial k^2}{\partial t}, \quad (4.69)$$

$$\begin{aligned}2\mathbf{k} \cdot \nabla(\mathbf{k} \cdot \bar{\mathbf{v}}) &= 2\mathbf{k} \cdot [(\mathbf{k} \cdot \nabla) \bar{\mathbf{v}} + (\bar{\mathbf{v}} \cdot \nabla) \mathbf{k} + \mathbf{k} \times (\nabla \times \bar{\mathbf{v}}) + \bar{\mathbf{v}} \times (\nabla \times \mathbf{k})] \\ &= 2\mathbf{k} \cdot (\mathbf{k} \cdot \nabla) \bar{\mathbf{v}} + 2\mathbf{k} \cdot (\bar{\mathbf{v}} \cdot \nabla) \mathbf{k}.\end{aligned}\quad (4.70)$$

With

$$\bar{\mathbf{v}} \cdot \nabla k^2 = 2\bar{\mathbf{v}} \cdot (\mathbf{k} \cdot \nabla) \mathbf{k} - 2\bar{\mathbf{v}} \cdot (\nabla \times \mathbf{k}) \times \mathbf{k} = 2\bar{\mathbf{v}} \cdot (\mathbf{k} \cdot \nabla) \mathbf{k} \quad (4.71)$$

Eq. (4.70) can be rewritten

$$2\mathbf{k} \cdot \nabla(\mathbf{k} \cdot \bar{\mathbf{v}}) = 2\mathbf{k} \cdot (\mathbf{k} \cdot \nabla) \bar{\mathbf{v}} + 2\mathbf{k} \cdot (\bar{\mathbf{v}} \cdot \nabla) \mathbf{k} - 2\bar{\mathbf{v}} \cdot (\mathbf{k} \cdot \nabla) \mathbf{k} + \bar{\mathbf{v}} \cdot \nabla k^2 \quad (4.72)$$

Now, combining Eqs (4.68), (4.69) and (4.72) we find

$$\begin{aligned}-(\omega - \mathbf{k} \cdot \bar{\mathbf{v}}) \nabla \cdot \mathbf{k} &= \frac{1}{2} \nabla \cdot (k^2 \mathbf{c}_g) - \left(\frac{\partial}{\partial t} + \bar{\mathbf{v}} \cdot \nabla \right) k^2 + \frac{1}{2} \frac{\partial k^2}{\partial t} \\ &\quad - \mathbf{k} \cdot (\mathbf{k} \cdot \nabla) \bar{\mathbf{v}} - \mathbf{k} \cdot (\bar{\mathbf{v}} \cdot \nabla) \mathbf{k} + \bar{\mathbf{v}} \cdot (\mathbf{k} \cdot \nabla) \mathbf{k}.\end{aligned}\quad (4.73)$$

Rewriting of the transport equation (Eq. (4.66)) leads to

$$\begin{aligned}&\left(\frac{\partial}{\partial t} + \bar{\mathbf{v}} \cdot \nabla \right) k^2 a_0 - (\omega - \mathbf{k} \cdot \bar{\mathbf{v}}) [2(\mathbf{k} \cdot \nabla) a_0 + a_0 \nabla \cdot \mathbf{k}] \\ &\quad + [\mathbf{e}_3 \times \nabla(\bar{\zeta} + f)] \cdot \nabla a_0 \\ &= a_0 \left(\frac{\partial}{\partial t} + \bar{\mathbf{v}} \cdot \nabla \right) k^2 + k^2 \frac{\partial a_0}{\partial t} \\ &\quad + [k^2 \bar{\mathbf{v}} + \mathbf{e}_3 \times \nabla(\bar{\zeta} + f) - 2(\omega - \mathbf{k} \cdot \bar{\mathbf{v}}) \mathbf{k}] \cdot \nabla a_0 - a_0 (\omega - \mathbf{k} \cdot \bar{\mathbf{v}}) \nabla \cdot \mathbf{k} \\ &= k^2 \frac{da_0}{dt} + a_0 \left(\frac{\partial}{\partial t} + \bar{\mathbf{v}} \cdot \nabla \right) k^2 - a_0 (\omega - \mathbf{k} \cdot \bar{\mathbf{v}}) \nabla \cdot \mathbf{k} = 0,\end{aligned}\quad (4.74)$$

where is used Eq. (4.67). Finally, substitution of Eq. (4.73) in Eq. (4.74) yields

$$\begin{aligned}&k^2 \frac{da_0}{dt} + \frac{a_0}{2} \frac{dk^2}{dt} + \frac{k^2 a_0}{2} \nabla \cdot \mathbf{c}_g \\ &\quad - a_0 [\mathbf{k} \cdot (\mathbf{k} \cdot \nabla) \bar{\mathbf{v}} + \mathbf{k} \cdot (\bar{\mathbf{v}} \cdot \nabla) \mathbf{k} - \bar{\mathbf{v}} \cdot (\mathbf{k} \cdot \nabla) \mathbf{k}] = 0.\end{aligned}\quad (4.75)$$

On a plane, this results in the following kinetic energy equation (after using the wave packet structure (Eq. (4.22)) and multiplying by $|a_0|$)

$$\frac{dE}{dt} + E \nabla \cdot \mathbf{c}_g = |a_0|^2 \left[(m^2 - n^2) \frac{\partial \bar{u}}{\partial x} + mn \left(\frac{\partial \bar{u}}{\partial y} + \frac{\partial \bar{v}}{\partial x} \right) \right]. \quad (4.76)$$

In order to transform the kinetic energy equation to spherical coordinates we use generalized expressions for the vector operators as given in Appendix A-3 of Haltiner and Williams (1980). Local Cartesian coordinates will be defined by

$$x = (R \cos \phi) \lambda, \quad y = R \phi, \\ \frac{\partial}{\partial x} = \frac{1}{R \cos \phi} \frac{\partial}{\partial \lambda}, \quad \frac{\partial}{\partial y} = \frac{1}{R} \frac{\partial}{\partial \phi}, \quad (4.77)$$

where R denotes the radius of the earth, $R = 6.37 \times 10^6 \text{m}$. Hence

$$\frac{\partial m}{\partial y} = \frac{\partial n}{\partial x} + m \frac{\tan \phi}{R}. \quad (4.78)$$

The derivatives of the unit vectors are given by

$$\begin{aligned} \frac{\partial \mathbf{e}_1}{\partial \lambda} &= \mathbf{e}_2 \sin \phi, \\ \frac{\partial \mathbf{e}_2}{\partial \lambda} &= -\mathbf{e}_1 \sin \phi, \\ \frac{\partial \mathbf{e}_1}{\partial \phi} &= 0, \\ \frac{\partial \mathbf{e}_2}{\partial \phi} &= 0. \end{aligned} \quad (4.79)$$

Using these equations we find extra metric terms in the kinetic energy density equation,

$$\begin{aligned} \frac{dE}{dt} + E \nabla \cdot \mathbf{c}_g &= |a_0|^2 \left[(m^2 - n^2) \frac{\partial \bar{u}}{\partial x} + mn \left(\frac{\partial \bar{u}}{\partial y} + \frac{\partial \bar{v}}{\partial x} \right) \right. \\ &\quad \left. + (n^2 - m^2) \frac{\bar{v} \tan \phi}{R} + mn \frac{\bar{u} \tan \phi}{R} \right]. \end{aligned} \quad (4.80)$$

References

- Ambaum, M. P. H., 1997: Large-scale dynamics of the tropopause, Ph. D. thesis, Eindhoven University of Technology, 113 pp.
- Anderson, J. L., 1994: Selection of initial conditions for ensemble forecasts in a simple perfect model framework. *J. Atmos. Sci.*, **53**, 22-36.
- Anderson, J. L., 1997: The impact of dynamical constraints on the selection of initial conditions for ensemble prediction: Low-order perfect model results. *Mon. Wea. Rev.*, **125**, 2969-2983.
- Anderson, J. R., and R. D. Rosen, 1983: The latitude-height structure of the 40-50 day variations in atmospheric angular momentum. *J. Atmos. Sci.*, **40**, 1584-1591.
- Barkmeijer, J., 1993: Local skill prediction for the ECMWF model using adjoint techniques. *Mon. Wea. Rev.*, **121**, 1262-1268.
- Barkmeijer, J., 1996: Constructing fast-growing perturbations for the nonlinear regime. *J. Atmos. Sci.*, **53**, 2838-2851.
- Barkmeijer, J., R. Mureau, J. Oortwijn, and J.D. Opsteegh, 1995: Research on regional predictability at KNMI. ECMWF seminar proceedings on predictability, Vol. 1, Reading, England, 247-274.
- Branstator, G., 1987: A striking example of the atmosphere's leading travelling pattern. *J. Atmos. Sci.*, **44**, 2310-2323.
- Bretherton, F. P., and C. J. R. Garrett, 1969: Wavetrains in inhomogeneous moving media. *Proc. Roy. Soc. A.*, **302**, 529-554.
- Buizza, R., 1994: Localization of optimal perturbations using a projection operator. *Quart. J. Roy. Meteor. Soc.*, **120**, 1647-1681.
- Buizza, R., and F. Molteni, 1996: On the role of finite-time barotropic instability during transition to blocking. *J. Atmos. Sci.*, **53**, 1675-1697.
- Buizza, R., and T. N. Palmer, 1995: The singular-vector structure of the atmospheric global circulation. *J. Atmos. Sci.*, **52**, 1434-1456.
- Buizza, R., J. Tribbia, F. Molteni, and T. Palmer, 1993: Computation of optimal unstable structures for a numerical weather prediction model. *Tellus*, **45A**, 388-407.

- Charney, J. G., and J. G. DeVore, 1979: Multiple flow equilibria in the atmosphere and blocking. *J. Atmos. Sci.*, **36**, 1206-1216.
- Courtier, P., J. Derber, R. Errico, J.-F. Louis, and T. Vukićević, 1993: Important literature on the use of adjoint, variational methods and the Kalman filter in meteorology. *Tellus*, **45A**, 342-357.
- Dole, R. M., 1978: The objective representation of blocking patterns. In *The general circulation: theory, modeling and observations*. Notes from a Colloquium, Summer, 1978. NCAR/CQ - 6 + 1978 - ASP, 406 - 426.
- Dole, R. M., 1986: The life cycles of persistent anomalies and blocking over the North Pacific. *Advances in Geophysics*, Vol. 29, 31-69.
- Dole, R. M., and N. D. Gordon, 1983: Persistent anomalies of the extratropical northern hemisphere wintertime circulation: Geographical distribution and regional persistence characteristics. *Mon. Wea. Rev.*, **111**, 1567-1586.
- ECMWF, 1993: *Expert meeting on ensemble prediction system 6-7 July 1993*. Report of expert meeting. Reading, England, ECMWF, 133 pp.
- Epstein, E. S., 1969: Stochastic dynamic prediction. *Tellus*, **21**, 739-759.
- Errico, R. M., T. Vukićević, and K. Reader, 1993: Examination of the accuracy of a tangent linear model. *Tellus*, **45A**, 462-477.
- Farrell, B. F., 1982: The growth of disturbances in a baroclinic flow. *J. Atmos. Sci.*, **39**, 1636-1686.
- Farrell, B. F., 1988: Optimal excitation of neutral Rossby modes. *J. Atmos. Sci.*, **45**, 163-172.
- Farrell, B. F., 1989: Optimal excitation of baroclinic waves. *J. Atmos. Sci.*, **46**, 1193-1206.
- Farrell, B. F., and P. J. Ioannou, 1996: Generalized stability theory. Part I: Autonomous operators. *J. Atmos. Sci.*, **53**, 2025-2040.
- Flierl, G. R., 1978: Models of vertical structure and the calibration of two-layer models. *Dyn. Atmos. Oceans*, **2**, 341-381.
- Haines, K., and J. C. Marshall, 1987: Eddy-forced coherent structures as a prototype of atmospheric blocking. *Quart. J. Roy. Meteor. Soc.*, **113**, 681-704.
- Haltiner, G. J., and R. T. Williams, 1980: *Numerical prediction and dynamic meteorology*, Wiley, 477 pp.
- Hansen, A. R., and A. Sutera, 1986: On the probability density distribution of large-scale atmospheric wave amplitude. *J. Atmos. Sci.*, **43**, 3250-3265.
- Hoskins, B. J., and D. Karoly, 1981: The steady linear response of a spherical atmosphere to thermal and orographic forcing. *J. Atmos. Sci.*, **38**, 1179-1196.

- Hoskins, B. J., I. N. James, and G. H. White, 1983: The shape, propagation and mean-flow interaction of large-scale weather systems. *J. Atmos. Sci.*, **40**, 1597-1612.
- Houtekamer, P. L., and J. Derome, 1995: Methods for ensemble prediction. *Mon. Wea. Rev.*, **123**, 2181-2196.
- Hsu, H.-H., and J. M. Wallace, 1985: Vertical structure of wintertime teleconnection patterns. *J. Atmos. Sci.*, **42**, 1693-1710.
- Kimoto, M., H. Mukougawa, and S. Yoden, 1992: Medium-range forecast skill variation and blocking transition: A case study. *Mon. Wea. Rev.*, **120**, 1616-1627.
- Kushnir, Y., 1987: Retrograding wintertime low-frequency disturbances over the North Pacific ocean. *J. Atmos. Sci.*, **44**, 2727-2742.
- 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. *Tellus*, **38A**, 97-110.
- Legras, B., and M. Ghil, 1985: Persistent anomalies, blocking and variations in atmospheric predictability. *J. Atmos. Sci.*, **42**, 433-471.
- Leith, C. E., 1974: Theoretical skill of Monte Carlo forecasts. *Mon. Wea. Rev.*, **102**, 409-418.
- Lighthill, J., 1978: *Waves in Fluids*. Cambridge University Press, 504 pp.
- Liu, Q., 1994: On the definition and persistence of blocking. *Tellus*, **46A**, 286-298.
- Liu, Q., and J. D. Opsteegh, 1995: Inter annual and decadal variations of blocking activity in a quasi-geostrophic model. *Tellus*, **47A**, 941-954.
- Longuet-Higgins, M. S., and R. W. Stewart, 1961: The changes in amplitude of short gravity waves on steady non-uniform currents. *J. Fl. Mech.*, **10**, 529-549.
- Longuet-Higgins, M. S., and R. W. Stewart, 1964: Radiation stresses in water waves; a physical discussion, with applications. *Deep-Sea Res.*, **11**, 529-562.
- 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.
- Machenhauer, B., 1979: The spectral method. *Numerical methods used in atmospheric models*, Vol. II, WMO/GARP Publ. Ser. 17, 121-275.
- Marshall, J., and F. Molteni, 1993: Toward a dynamical understanding of planetary-scale flow regimes. *J. Atmos. Sci.*, **50**, 1792-1818.
- Michelangeli, P.-A., and R. Vautard, 1998: The dynamics of Euro-Atlantic blocking onsets. *subm. to Quart. J. Roy. Meteor. Soc.*

- Michelangeli, P.-A., R. Vautard, and B. Legras, 1995: Weather regimes: Recurrence and quasi stationarity. *J. Atmos. Sci.*, **52**, 1237-1256.
- Mo, K., and M. Ghil, 1987: Statistics and dynamics of persistent anomalies. *J. Atmos. Sci.*, **44**, 877-901.
- Mo, K., and M. Ghil, 1988: Cluster analysis of multiple planetary flow regimes. *J. Geophys. Res.*, **93**, 927-952.
- Molteni, F., and T. N. Palmer, 1993: Predictability and finite-time instability of the northern winter circulation. *Quart. J. Roy. Meteor. Soc.*, **119**, 269-298.
- Molteni, F., R. Buizza, T. N. Palmer, and T. Petroliaigis, 1996: The ECMWF ensemble prediction system: Methodology and validation. *Quart. J. Roy. Meteor. Soc.*, **122**, 73-119.
- Müller, P., 1978: On the parametrization of eddy-mean flow interaction in the ocean. *Dyn. Atmos. Oceans*, **2**, 383-408.
- Mureau, R., F. Molteni, and T. N. Palmer, 1993: Ensemble predictions using dynamically-conditioned perturbations. *Quart. J. Roy. Meteor. Soc.*, **119**, 299-323.
- Nakamura, H., and J. M. Wallace, 1990: Observed changes in baroclinic wave activity during the life cycles of low-frequency circulation anomalies. *J. Atmos. Sci.*, **47**, 1100-1116.
- Nakamura, H., and J. M. Wallace, 1993: Synoptic behavior of baroclinic eddies during blocking onset. *Mon. Wea. Rev.*, **121**, 1892-1903.
- Namias, J., 1964: Seasonal interactions between the North Pacific Ocean and the atmosphere during the 1960s. *Mon. Wea. Rev.*, **97**, 173-192.
- Neven, E. C., 1993: Modons on a sphere. Ph. D. thesis, University of Utrecht, 175 pp.
- Oortwijn J., 1998a: Predictability of the onset of blocking and strong zonal flow regimes. *J. Atmos. Sci.*, **55**, 973-994.
- Oortwijn J., 1998b: Growth properties of optimal transition perturbations. *subm. to J. Atmos. Sci.*
- Oortwijn, J., and J. Barkmeijer, 1995: Perturbations that optimally trigger weather regimes. *J. Atmos. Sci.*, **52**, 3932-3944.
- Orr, W. M. F., 1907: The stability or instability of the steady motions of a perfect liquid and of a viscous liquid. *Proc. Roy. Irish Acad.*, **A27**, 9-138.
- Palmer, T. N., 1993: Extended-range atmospheric prediction and the Lorenz model. *Bull. Amer. Meteor. Soc.*, **74**, 49-65.
- Palmer, T. N., Č. Branković, F. Molteni, and S. Tibaldi, 1990: Extended-range predictions with ECMWF models: Interannual variability in operational model integrations. *Quart. J. Roy. Meteor. Soc.*, **116**, 799-834.
- Parlett, P., 1980: *The symmetric eigenvalue problem*. Series in computational mathematics.

Prentice Hall, 348 pp.

Plaut, G., and R. Vautard, 1994: Spells of low-frequency oscillations and weather regimes in the northern hemisphere. *J. Atmos. Sci.*, **51**, 210-236.

Rabier, F., P. Courtier, M. Hervéou, B. Strauss, and A. Persson, 1993: Sensitivity of forecast error to initial conditions using the adjoint model. ECMWF Technical Memorandum No. 197.

Rabier, F., E. Klinker, P. Courtier and A. Hollingsworth, 1996: Sensitivity of forecast errors to initial conditions. *Quart. J. Roy. Meteor. Soc.*, **122**, 121-150.

Reinhold, B. B., 1987: Weather regimes: The challenge in extended-range forecasting *Science*, **235**, 437-441.

Reinhold, B. B., and R. T. Pierrehumbert, 1982: Dynamics of weather regimes: Quasi-stationary waves and blocking. *Mon. Wea. Rev.*, **110**, 1105-1145.

Rex, D. F., 1950a: Blocking action in the middle troposphere and its effect upon regional climate. I. An aerological study of blocking. *Tellus*, **2**, 196-211.

Roads, J. O., 1987: Predictability in the extended range. *J. Atmos. Sci.*, **44**, 3495-3527.

Rex, D. F., 1950b: Blocking action in the middle troposphere and its effect upon regional climate. II. An aerological study of blocking. *Tellus*, **2**, 275-301.

Schuster, H. G., 1989: *Deterministic chaos. An introduction. 2th. revised edition.* VCH Publishers, 270 pp.

Simmons, A. J., J. M. Wallace, and G. W. Branstator, 1983: Barotropic wave propagation and instability, and atmospheric teleconnection patterns. *J. Atmos. Sci.*, **40**, 1363-1392.

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.

Thépaut, J.-N., and P. Courtier, 1991: Four-dimensional variational data assimilation using the adjoint of a multilevel primitive equation model. *Quart. J. Roy. Meteor. Soc.*, **117**, 1225-1254.

Tibaldi, S., and F. Molteni, 1990: On the operational predictability of blocking. *Tellus*, **42A**, 343-365.

Tibaldi, S., E. Tosi, A. Navarra, and L. Pedulli, 1994: Northern and Southern hemisphere seasonal variability of blocking frequency and predictability. *Mon. Wea. Rev.*, **122**, 1971-2003.

Toth, Z., and E. Kalnay, 1993: Ensemble forecasting at NMC: The generation of perturbations. *Bull. Amer. Meteor. Soc.*, **74**, 2317-2330.

Toth, Z., and E. Kalnay, 1997: Ensemble forecasting at NMC and the breeding method. *subm. to Mon. Wea. Rev.*

- Trefethen, L. N., A. E. Trefethen, S. C. Reddy, and T. A. Driscoll, 1993: Hydrodynamic stability without eigenvalues. *Science*, **261**, 578-584.
- Vautard, R., 1990: Multiple weather regimes over the North Atlantic: Analysis of precursors and successors. *Mon. Wea. Rev.*, **118**, 2056-2081.
- Vautard, R., and B. Legras, 1988: On the source of low-frequency variability. Part II: Nonlinear equilibration of weather regimes. *J. Atmos. Sci.*, **45**, 2845-2867.
- Vautard, R., B. Legras, and M. Deque, 1988: On the source of low-frequency variability. Part I: A statistical approach to persistence. *J. Atmos. Sci.*, **45**, 2811-2843.
- Vukićević, T., 1991: Nonlinear and linear evolution of initial forecast errors. *Mon. Wea. Rev.*, **119**, 1602-1611.
- Yang, S., and B. Reinhold, 1991: How does low-frequency variance vary? *Mon. Wea. Rev.*, **119**, 119-127.
- Young, W. R., and P. B. Rhines, 1980: Rossby wave action, enstrophy and energy in forced mean flows. *Geophys. Astrophys. Fluid Dynamics*, **15**, 39-52.
- Zeng, Q.-C., 1983a: The development of quasi-geostrophic baroclinic disturbances. *Tellus*, **35A**, 337-349.
- Zeng, Q.-C., 1983b: The evolution of a Rossby-wave packet in a three-dimensional baroclinic atmosphere. *J. Atmos. Sci.*, **40**, 73-84.
- Zupanski, M., 1993: Regional four-dimensional variational data assimilation in a quasi-operational forecasting environment. *Mon. Wea. Rev.*, **121**, 2396-2408.

Samenvatting

Voorspelbaarheid van weerregime transities

Weersverwachtingen worden al sinds 1950 gemaakt met behulp van computers. Hiervoor worden steeds snellere en duurdere computers gebruikt. De verbetering van de weersverwachting houdt echter geen gelijke tred met de investeringen in computers en computermodellen die het weer simuleren. De modellen geven nog al eens een verkeerde verwachting. Waarom is het weer eigenlijk zo moeilijk te voorspellen? Er zijn perioden dat weersverwachtingen slechts voor één of enkele dagen betrouwbaar zijn, terwijl in andere perioden het weer meer dan een week goed voorspelbaar is. Het ene weertype blijkt beter voorspelbaar te zijn dan het andere weertype. Soms houdt een weertype erg lang stand. In zo'n geval is de zomer bijvoorbeeld erg mooi of is de winter zo langdurig koud dat er een elfstedentocht geschaatst kan worden. Echter in soortgelijke situaties kan het weer ook heel snel omslaan naar een ander weertype. Waarom is het weer zo variabel en hoe komt het dat een omslag zo snel kan optreden?

De voorafgaande vragen hebben erg veel te maken met het onderwerp dat in dit proefschrift bestudeerd wordt. Ik zal eerst op deze vragen ingaan en proberen een beeld te schetsen van het weer, of beter, van de atmosferische circulatie, en van de problemen bij de voorspelling ervan. Tot slot zal het onderzoek dat beschreven is in dit proefschrift samengevat worden.

De tijdsevolutie van de atmosfeer is erg complex. Daarvoor is een aantal redenen aan te wijzen. Twee daarvan zal ik hier bespreken. Als eerste is de atmosfeer een zeer groot systeem. Veel processen spelen zich gelijktijdig af op verschillende plaatsen in de atmosfeer en met verschillende intensiteit en omvang. Het ene proces, zoals bijvoorbeeld de ontwikkeling van een krachtig hogedrukgebied, bepaalt het weer voor een lange periode en een groot gebied, terwijl een ander proces, zoals bijvoorbeeld een onweersbui, slechts een korte duur heeft en erg lokaal is. Een tweede belangrijke reden is dat de atmosfeer een niet-lineair systeem is. Dit heeft belangrijke consequenties voor de voorspelbaarheid. Hier zal ik later op terug komen. Eerst zal ik proberen duidelijk te maken wat lineair en niet-lineair precies betekenen.

Een verband tussen twee variabelen is lineair indien de verandering van de ene variabele een evenredige verandering geeft in de andere variabele. Zo zal een auto die twee keer zo hard rijdt als een andere auto, in een gegeven tijdsinterval, een twee keer zo grote afstand afleggen. Als in een elektrisch circuit de spanning verdubbeld wordt, zal de elektrische stroom ook in sterkte verdubbelen. De stijging van de luchtdruk leidt tot een evenredige stijging van de

kwikkolom in een kwikbarometer (iedere stijging van 13.6 mbar geeft een stijging van 1 cm kwik).

De gevolgen van een verandering in een variabele (zoals de snelheid, spanning en druk in bovenstaande voorbeelden) zijn in een lineair systeem meestal duidelijk. Niet-lineaire verbanden, waarin een verandering in het algemeen geen evenredige verandering tot gevolg heeft, zijn vaak veel complexer. Dat de natuur of andere systemen eigenlijk zelden (zuiver) lineair reageren mag blijken uit de volgende voorbeelden.

Als een auto harder gaat rijden, zal deze door een onevenredige toename in de luchtweerstand minder zuinig rijden. Als iemand twee keer zo lang in de zon blijft liggen, wordt hij of zij niet twee keer zo bruin (wellicht wel twee maal zo rood). In de supermarkt is een produkt in een grote verpakking meestal voordeliger dan dezelfde hoeveelheid bestaande uit kleinere verpakkingen. Verdubbeling van de dosis medicijnen leidt zelden tot een verdubbeling in de snelheid van het genezingsproces. Integendeel, dit kan de gezondheid ernstige schade toebrengen.

Bij niet-lineaire processen hangt de tijdsevolutie af van niet-lineaire verbanden. Een dergelijk systeem wordt een niet-lineair dynamisch systeem genoemd. Een goed voorbeeld hiervan is de beweging van de beurskoersen. Wanneer de koersen een stijgende lijn vertonen leidt dit vaak tot een extra stijging van de koersen. Het gevoel dat het goed gaat overheerst, meer mensen stappen in (want ze willen de boot niet missen), met als gevolg dat de beurskoersen verder stijgen. Andersom geldt dit ook. Als het minder goed gaat besluiten mensen te verkopen, de koers daalt, waardoor nog meer mensen gaan verkopen. In beide gevallen is er sprake van een zogenaamde positieve terugkoppeling (alhoewel het effect in het laatste geval negatief is). Een dergelijk proces zal door een positieve terugkoppeling zichzelf versterken.

Mathematisch gezien betekent het niet-lineair zijn dat er termen in de vergelijkingen voorkomen die bijvoorbeeld kwadratisch zijn. Dit niet-lineair zijn betekent dat variabelen met elkaar vermenigvuldigd worden of wel gekoppeld zijn. Een verdubbeling van de amplitude van een bepaalde variabele hoeft dan niet te leiden tot een verdubbeling in de intensiteit van de respons. Deze kan veel sterker of zwakker zijn. Dit maakt het systeem complexer: de invloed van een verandering in een variabele op andere variabelen is niet direct in te zien. Hier komt de voorspelbaarheid om de hoek kijken.

De atmosferische circulatie is ook een niet-lineair dynamisch systeem. In feite is de beweging van lucht niets anders dan een grote verzameling botsende moleculen die bewegen in het zwaartekrachtsveld van de aarde. De beweging van ieder molecuul wordt beschreven door de wetten van Newton. Door de krachten te beschouwen die op een luchtpakketje (een klein volume lucht met nog steeds enorm veel moleculen) werken, kunnen de zogenaamde Navier-Stokes-vergelijkingen afgeleid worden. Deze beschrijven de beweging van lucht (en in het algemeen van vloeistoffen), waarbij geen rekening gehouden hoeft te worden met de beweging van individuele moleculen. De Navier-Stokes-vergelijkingen zijn als het ware de bewegingsvergelijkingen van Newton voor een continuum. Ze bevatten niet-lineaire termen.

De wetten van Newton en de Navier-Stokes-vergelijkingen zijn deterministisch. Dit houdt in dat, indien de snelheid en positie van ieder molecuul in de lucht op een begintijdstip bekend zouden zijn, in principe alle volgende botsingen uitgerekend kunnen worden. Met andere woorden, als de begintoestand van een deterministisch systeem bekend is, is de verdere evolutie volledig bepaald. Zelfs tot het einde van de vorige eeuw dacht men daarom dat

men slechts nauwkeuriger hoefde te rekenen om een betere voorspelling te maken. Echter, het gedrag van niet-lineaire systemen is in veel gevallen chaotisch. Deze systemen worden gekenmerkt door een grote gevoeligheid voor kleine veranderingen in de begintoestand. Dit betekent dat een kleine fout in de schatting van de begintoestand steeds sterker doorwerkt in de tijd. Dit heeft zelfs als gevolg dat na een eindige tijd het gedrag volledig onvoorspelbaar wordt. Deze eindige tijd wordt de voorspelbaarheidshorizon genoemd: voorbij deze tijd kunnen we niet voorspellen. De voorspelbaarheidshorizon van de atmosferische circulatie wordt geschat op 1 à 2 weken. We kunnen dus nog zoveel investeren in computers en computermodellen, we zullen nooit deterministische voorspellingen kunnen maken voor langer dan 2 weken vooruit. *

Chaos betekent niet dat het systeem zich willekeurig ontwikkelt. De atmosfeer zal zich namelijk ontwikkelen volgens de deterministische Navier-Stokes-vergelijkingen. We spreken dan ook van deterministische chaos. Naast een grillig karakter vertoont het weer ook duidelijk een soort orde. Niet alle weersituaties zullen zich voordoen en de situaties die optreden gebeuren in het algemeen in dezelfde vorm. Zo zal in Nederland nooit een temperatuur van 60°C bereikt worden, zullen we nooit kunnen schaatsen in de zomer, waait de wind rond een lagedrukgebied altijd tegen de klok in, en gaat mooi en standvastig weer gepaard met een krachtig hogedrukgebied. Het weer heeft dus bepaalde voorkeurspatronen die steeds weer opnieuw ontstaan.

De belangrijkste bronnen die het weer in beweging brengen zijn de temperatuurverschillen tussen de polen en evenaar, door een ongelijke instraling van de zon, en de draaiing van de aarde. Hierdoor bevindt er zich op het noordelijk halfrond, op breedten tussen 40° en 70°, hoog in de atmosfeer een krachtige westenwind (het sterkst op zo'n 10 kilometer hoogte) die geheel rond de aarde waait. Deze stroming wordt de straalstroom genoemd. De straalstroom waait niet precies langs een breedtegraad maar vertoont slingeringen. Door de aanwezigheid van bergen als de Himalayas en de Rocky Mountains en verschillen in de warmtecapaciteit tussen land en oceanen, is de stroming ten oosten van deze bergketens sterker dan op andere plaatsen. Op deze plaatsen ontstaan depressies (lagedrukgebieden) vanwege de grote snelheidsverschillen die zich in de straalstroom voordoen (de warme Golfstroom is echter ook een belangrijke factor). Dit proces is te vergelijken met de waterstraal uit een kraan. Wanneer men de kraan niet te ver opendraait, is de straal erg glad en kan men er doorheen kijken. Wordt de kraan vervolgens verder opengedraaid dan wordt de straal veel chaotischer, turbulenter en ondoorzichtig. Dit komt door de snelheidsverschillen in de stroming waardoor allerlei storingen ontstaan. Storingen in de straalstroom kunnen zich onder gunstige omstandigheden ontwikkelen tot depressies. Een depressie mengt warme lucht uit het zuiden met koude lucht uit het noorden zodat temperatuurverschillen tussen de pool en evenaar genivelleerd worden. De grootte van een depressie, hetgeen de synoptische schaal genoemd wordt, is in de orde van zo'n duizend kilometer. De levensduur van een depressie is ongeveer

*Dit betekent niet dat seizoens- of klimaatverwachtingen zinloos zijn. Het weer hangt namelijk ook af van de stroming en temperatuur van de oceanen. De evolutie van de oceanen, die veel langzamer variëren dan de atmosfeer, kan voor een veel langere termijn voorspeld worden. Een voorbeeld is het fenomeen El Niño: de hogere oppervlaktetemperaturen van het zeewater aan de oostkant van de tropische Pacific hebben een maandenlange invloed op bijna de gehele atmosferische circulatie. Ook de invloed van broeikasgassen op het klimaat heeft voorspelbare componenten.

5 dagen.

Depressies die het weer in Europa beïnvloeden ontstaan meestal ergens boven het westen van de Atlantische oceaan. Door de straalstroom worden ze meegevoerd in de richting van ons land. Soms bevindt er zich echter een sterk hogedrukgebied boven Europa en wordt de straalstroom afgebogen naar het noorden of zuiden of splitst hij zich zelfs in een noordelijke en zuidelijke tak. Depressies kunnen ons land dan niet bereiken. Zo'n situatie wordt een blokkade genoemd, omdat de oostwaartse beweging geblokkeerd is. Zo'n blokkade kan zich langere tijd handhaven, soms wel een aantal weken tot één of enkele maanden, en doet zich herhaaldelijk voor. Een blokkade is een voorkeurspatroon van het weer op een langere tijdschaal dan die van depressies. In het gebied van een blokkade is het weer doorgaans erg mooi en droog, in de zomer warm en in de winter juist koud. In gebieden ten noorden of ten zuiden van een blokkade is het weer veel natter met meer wind, omdat depressies nu hier overtrekken.

Een blokkade is een voorbeeld van een weerregime. De term weerregime is geïntroduceerd door Reinhold en Pierrehumbert (1982). Zij veronderstelden dat de atmosfeer een aantal voorkeurspatronen heeft die een soort evenwichtstoestanden zijn tussen processen op een zeer grote ruimtelijke schaal (met afmetingen van vele duizenden tot tienduizenden kilometer) en processen op een synoptische schaal. Doordat deze processen elkaar wederzijds positief beïnvloeden kan een patroon ontstaan dat lange tijd stand houdt. Zo'n voorkeurspatroon wordt een weerregime genoemd. De atmosferische circulatie zal zich afwisselend in verschillende weerregimes bevinden, met tussenliggende perioden waarin een omslag (transitie) plaatsvindt. Het noordoostelijk deel van de Atlantische oceaan en het noordoostelijk deel van de Pacific zijn voorkeursgebieden waar deze regimes ontstaan. Dit heeft te maken met de ligging van de bergketens en de sterkte van de straalstroom ten westen van deze gebieden.

Een tweede regime dat het weer boven West-Europa in grote mate beïnvloedt, is een sterk zonale stroming regime. In dit geval vertoont de straalstroom bijna geen slingeren en loopt de sterke oostwaartse stroming door tot boven West-Europa. Depressies zullen dan over ons land trekken. Het weer is erg wisselvallig en nat. Dit kan een lange tijd aanhouden. De winters van 1994 en 1995 zijn voorbeelden waarin het regime van sterk zonale stroming zich lange tijd manifesteerde. Beide keren leidde de overvloedige hoeveelheid regen bijna tot een watersnoodramp.

Hoewel een weerregime een lange periode kan standhouden, is de duur niet erg goed te voorspellen. Dit komt doordat de instandhouding van de weerregimes afhankelijk is van depressies die slechts zo'n 5 dagen leven. Het tijdstip waarop de atmosferische circulatie een transitie maakt naar een ander regime blijkt erg moeilijk voorspelbaar te zijn. Deze omslagpunten zijn juist erg interessant omdat het weer daarna een totaal ander karakter kan hebben. Dit is te vergelijken met de omslagpunten in de beurskoersen. Het omslagpunt van een positieve-stemming regime naar een negatieve-stemming regime is zowel erg interessant als lastig te voorspellen.

In dit proefschrift zijn de transities tussen een blokkade regime en een sterk zonale stroming regime bestudeerd. De studie is gedaan door de atmosferische circulatie te simuleren met behulp van een computermodel. Dit model heeft drie lagen op verschillende hoogten (op ongeveer 2, 5 en 12 km) in de atmosfeer. Iedere laag bestaat uit een rooster van punten

(64×32) dat horizontaal verdeeld is. De horizontale afstand tussen twee roosterpunten is zo'n 500 km. Dit betekent dat alleen de grootschalige stromingen bestudeerd kunnen worden. In ieder roosterpunt zijn variabelen als de luchtdruk, snelheid en temperatuur bekend. Door middel van vergelijkingen die afgeleid zijn uit de Navier-Stokes-vergelijkingen kan de tijdsevolutie van de atmosfeer bepaald worden. Door daarnaast gebruik te maken van geavanceerde wiskundige technieken, kan de gevoeligheid in de begintoestand voor het ontstaan van regimes berekend worden. Zo ben ik in staat om transities tussen een geblokkeerd regime en een sterk zonale stroming zeer effectief te bestuderen. Deze studie is beschreven in de hoofdstukken van het proefschrift. In de volgende alinea's zal de inhoud van deze hoofdstukken worden samengevat.

Hoofdstuk 1 is een inleidend hoofdstuk. Als eerste wordt een kort overzicht gegeven van de theorie van weerregimes. Daarna wordt met een eenvoudig voorbeeld aangetoond dat kleine verschillen in de begintoestand tussen twee stromingen (bijvoorbeeld een andere windsnelheid of temperatuur in een aantal roosterpunten van het model) zeer snel kunnen toenemen. Ook wordt uitgelegd hoe een berekening gemaakt kan worden van de maximale verschillen na een bepaalde (niet al te lange) periode. Aan de hand van een relatief zeer eenvoudig niet-lineair dynamisch systeem, het zogenaamde Lorenz model, laat ik zien dat de gevoeligheid in de begintoestand voor regime transities erg groot kan zijn. Dit model ligt ten grondslag aan de beroemde metafoer van de vlinder van Lorenz: het fladderen van een vlinder in Noord-Amerika kan een aantal dagen later een storm in Europa veroorzaken of te niet doen. Als laatste wordt een beschrijving gegeven van het computermodel dat gebruikt is om experimenten uit te voeren.

In hoofdstuk 2 wordt de methode beschreven waarmee optimale transitie-perturbaties berekend kunnen worden. Deze storingen zullen, indien ze bij de begintoestand van een atmosferische stroming opgeteld worden, na een aantal dagen de maximale verandering (ten opzichte van de geëvolueerde toestand zonder deze perturbaties) in het weerregime teweegbrengen. De maximale verandering die in een periode bereikt kan worden is een maat voor de gevoeligheid in de begintoestand in deze periode. Het blijkt dat de gevoeligheid sterk van dag tot dag fluctueert. De wiskundige methoden die gebruikt worden zijn slechts geldig voor niet al te grote perturbaties. In dat geval zijn de veranderingen in de processen (bij benadering) lineair. Als we een periode die langer duurt dan 3 dagen bestuderen, worden de perturbaties echter te groot. De groei wordt dan ook bepaald door niet-lineaire termen in de vergelijkingen. Daardoor worden de berekeningen veel complexer. Door de lineaire methode herhaaldelijk toe te passen is het toch mogelijk om perioden van zo'n 6 dagen te bestuderen. Het blijkt dan dat er voor bijna ieder weertype hele kleine perturbaties in de begintoestand gevonden kunnen worden die ervoor zorgen dat het weer na 6 dagen compleet is omgeslagen. Met andere woorden, de atmosferische circulatie kan erg gevoelig zijn voor kleine storingen. Deze storingen kunnen binnen een korte termijn het weer sterk beïnvloeden.

In hoofdstuk 3 is een groot experiment beschreven. Uit een grote set van perioden zijn de situaties met de hoogste en laagste gevoeligheid in de begintoestand voor het ontstaan van blokkades en sterk zonale stroming regimes geanalyseerd. Het blijkt, onder andere, dat stromingen met een hoge gevoeligheid een sterke straalstroom ten westen van een diffuente stroming (dit is een uitwaaiende stroming) vertonen. Laag gevoelige stromingen hebben een zwakkere en meer zonale stroming. Ook wordt onderzocht waarom deze karakteristieken

de gevoeligheid beïnvloeden.

Als laatste wordt in hoofdstuk 4 de ontwikkeling bestudeerd van optimale transitie-perturbaties. Onderzocht wordt hoe het komt dat deze zo snel kunnen groeien. Het blijkt dat er twee fasen onderscheiden kunnen worden. Tijdens de eerste fase groeit de perturbatie door middel van zogenaamde barotrope en barokliene mechanismen. Tijdens de tweede fase domineren barotrope mechanismen. Op verschillende manieren worden deze mechanismen bestudeerd.

Dankwoord

Ik heb in de afgelopen jaren bij het KNMI een prima tijd gehad. Daar heeft een groot aantal mensen toe bijgedragen. Op deze plaats wil ik hen graag bedanken.

Als eerste wil ik mijn begeleider en copromotor Jan Barkmeijer bedanken. Jan, jij hebt ervoor gezorgd dat ik me snel kon inwerken en dat het onderzoek direct in goede banen werd geleid. Onze samenwerking was erg prettig. Ik vond het dan ook jammer dat we de laatste twee jaar voornamelijk via telefoon en email hebben gecommuniceerd. In het laatste jaar heeft Wim Verkleij gefungeerd als een stand-in begeleider. Wim, ik ben je erg dankbaar dat je altijd klaar stond om mee te denken of om commentaar op mijn werk te leveren. Johan Grasman ben ik erkentelijk dat hij als promotor wilde optreden.

Theo Opsteegh, je hebt het onderzoek op afstand gevolgd maar door de discussies die we gevoerd hebben, die telkens leidden tot nieuwe vragen, kreeg ik een steeds beter fysisch inzicht in de materie. Bedankt daarvoor. Robert Mureau wil ik bedanken voor het redigeren van mijn Engelse teksten, de meteorologische lessen en de informele gesprekken zomaar tussendoor. Ook Frank Selten en Ruben Pasmaanter wil ik bedanken voor hun commentaar op delen van het proefschrift.

I would like to thank Jeffrey Anderson for his hospitality during my visit to Princeton, his interest in my work and his efforts to arrange a position for me at the Geophysical Fluid Dynamics Laboratory in Princeton.

In het bijzonder wil ik Maarten Ambaum en Hans Hersbach bedanken. Maarten, ik heb erg genoten van onze (al dan niet luchtige) conversaties over het werk of (meestal) over andere onderwerpen, tijdens verloren uurtjes op het KNMI of in de kroeg. Tegen Hans zou ik eigenlijk hetzelfde willen zeggen. Ik wil daar nog aan toevoegen dat ik met jou als kamer- en reisgenoot een geweldig laatste jaar heb gehad. Ook Geert Lenderink en Hanna Brands hebben een belangrijke bijdrage geleverd aan een prettige promotietijd.

Mijn dank gaat verder uit naar de overige VO-ers, OO-ers en KNMI-ers die hebben gezorgd voor een zeer goede sfeer. André van Lammeren en al de andere voetballers hebben mij heel wat uurtjes voetbalplezier bezorgd. Het voetballen (en, echter in mindere mate, ook het hardlopen) in de lunchpauzes was voor mij een ideale combinatie tussen in- en ontspanning.

Mijn vrienden en in het bijzonder mijn huisgenoot Cas van Arendonk wil ik bedanken voor de perfecte tijd naast het werk. Mijn ouders, mijn zus en mijn zwager wil ik bedanken voor de belangstelling, gastvrijheid en steun. Ik ben erg gelukkig dat ik zo'n fijne thuisbasis heb en draag daarom dit proefschrift op aan mijn ouders. Als laatste wil ik Corina Bulk bedanken. Onze toevallige ontmoeting in New York toont aan dat kleine invloeden vanuit Amerika inderdaad grote gevolgen kunnen hebben (en dat niet alleen voor het weer in Europa).

Curriculum vitae

Ik ben geboren in Tilburg op 10 augustus 1970. In 1974 ben ik verhuisd naar Goirle, waar ik opgegroeid ben. Van 1982 tot 1988 heb ik de middelbare school R.S.G. Koning Willem II in Tilburg doorlopen. Na het behalen van het VWO-diploma ben ik Technische Natuurkunde gaan studeren aan de Technische Universiteit Eindhoven. Mijn afstudeeronderzoek heb ik verricht aan het Koninklijk Shell Laboratorium in Amsterdam onder begeleiding van dr. Paul Iske. Het onderzoek betrof het opstellen van een theoretisch model voor elastische golfvoortplanting in heterogene media. In juni 1993 heb ik mijn ingenieurstitel behaald. Vanaf september 1993 heb ik mijn promotie-onderzoek verricht aan het Koninklijk Nederlands Meteorologisch Instituut (KNMI) in De Bilt. Het onderzoek werd gefinancierd door de stichting Geologisch, Oceanografisch en Atmosferisch Onderzoek (GOA) van de Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO). Tijdens mijn promotie heb ik verschillende workshops en instituten bezocht in Europa, Canada en de Verenigde Staten. Sinds 1 januari 1998 ben ik werkzaam bij Moret Ernst & Young Management Consultants te Utrecht.