Data assimilation and covariance dynamics in atmospheric models

translated: Datenassimilation und Dynamik von Fehlerkovarianzen in atmosphärischen Modellen

Dissertation

zur Erlangung des akademischen Grades Doktor der Naturwissenschaften an der Fakultät für Naturwissenschaften und Mathematik der Universität Wien

Eingereicht von

Mag. Alexander Beck

Wien, Februar 2003
Meinem Vater
# Contents

Abstract vi

Zusammenfassung vii

1 Introduction 1

1.1 Uncertainty in atmospheric models 2

1.2 Variational data assimilation 3

1.3 Atmospheric predictability 5

1.4 Motivation and outline 7

2 Four-dimensional variational data assimilation (4D–Var) 9

2.1 4D–Var: Method 10

2.2 Estimating analysis quality: Hessian of the costfunction 12

3 Error covariances in data assimilation and ensemble prediction 14

3.1 The Extended Kalman Filter (EKF) 14

3.2 Hessian singular vectors 15

3.2.1 Singular vectors and analysis errors 16
4 Multi–Gaussian sampling in ensemble prediction 18
  4.1 Ensemble prediction: Generation of initial perturbations 18
  4.2 The singular vector based Monte Carlo technique 19

5 The quasigeostrophic (QG) assimilation system 21
  5.1 The quasigeostrophic model 21
  5.2 QG/4D–Var system 23
  5.3 Observational network 23

6 Background formulation in QG/4D–Var system 26
  6.1 4D–Var in a cycling environment 26
  6.2 The static system 28
    6.2.1 Tuning of the static system — Determining of the static B 28
  6.3 Dynamic background: The Kalman Filter 29
  6.4 The Reduced–Rank Kalman Filter (RRKF) 30
    6.4.1 The RRKF with model error 34
    6.4.2 Scaling of the background 35

7 4D–Var experiments and results 36
  7.1 Summary of experiments 36
    7.1.1 Static system and EKF 37
    7.1.2 Background formulation and RRKF 38
  7.2 Cycling experiments 38
  7.3 Static system vs. EKF 40
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.4</td>
<td>Role of the background formulation and RRKF</td>
<td>54</td>
</tr>
<tr>
<td>7.5</td>
<td>Further experiments</td>
<td>61</td>
</tr>
<tr>
<td>7.5.1</td>
<td>Nonlinear covariance prediction</td>
<td>61</td>
</tr>
<tr>
<td>7.5.2</td>
<td>Scaling issues in RRKF</td>
<td>64</td>
</tr>
<tr>
<td>8</td>
<td>Covariance prediction with an operational forecasting system</td>
<td>69</td>
</tr>
<tr>
<td>8.1</td>
<td>Experiments with ECMWF EPS</td>
<td>69</td>
</tr>
<tr>
<td>8.2</td>
<td>EPS: results</td>
<td>70</td>
</tr>
<tr>
<td>8.2.1</td>
<td>EPS: skill scores</td>
<td>77</td>
</tr>
<tr>
<td>9</td>
<td>Summary and conclusions</td>
<td>83</td>
</tr>
<tr>
<td>9.1</td>
<td>QG/4D–Var system</td>
<td>83</td>
</tr>
<tr>
<td>9.1.1</td>
<td>Comments on the role of the RRKF in operational 4D–Var</td>
<td>86</td>
</tr>
<tr>
<td>9.2</td>
<td>Sampling experiments</td>
<td>87</td>
</tr>
<tr>
<td>9.3</td>
<td>Outlook</td>
<td>87</td>
</tr>
<tr>
<td>A</td>
<td>List of acronyms and abbreviations</td>
<td>89</td>
</tr>
<tr>
<td></td>
<td>List of figures</td>
<td>90</td>
</tr>
<tr>
<td></td>
<td>List of tables</td>
<td>95</td>
</tr>
<tr>
<td></td>
<td>Bibliography</td>
<td>96</td>
</tr>
<tr>
<td></td>
<td>Acknowledgments</td>
<td>103</td>
</tr>
<tr>
<td></td>
<td>Curriculum vitae</td>
<td>104</td>
</tr>
</tbody>
</table>
Abstract

Numerical weather prediction (NWP) requires the accurate specification of the initial state of the atmosphere, denoted the analysis. In order to obtain this analysis, information from observations as well as from previous forecasts from an NWP model is combined through a statistical process denoted data assimilation. In modern data assimilation systems the analysis is computed using variational formulations of least–squares estimation constrained to the model dynamics. This approach is called four dimensional variational data assimilation (4D–Var) since it adjusts globally the model state thereby taking into account the temporal distribution of the observations. The formulation of 4D–Var is dependent on the error statistics required to match the information from the previous forecast, denoted the background field, and from the observations. Specifically, the specification of the background error statistics has been the subject of intense research. In current operational data assimilation systems the background error statistics is essentially static and thus independent of the uncertainties associated with present atmospheric flow. At least theoretically, it is possible to use flow–dependent background error statistics in 4D–Var through exploiting the duality between variational data assimilation and estimation theory.

This thesis makes use of this duality to study the impact of different flow–dependent background error specifications in 4D–Var compared to a static formulation. The main emphasis is placed on quantifying the potential improvement in analysis quality that is possible through flow–dependent background errors available from an Extended Kalman Filter. Additionally, the performance of a suboptimal but still flow–dependent background specification, known as Reduced–Rank Kalman Filter (RRKF), is investigated. These experiments are carried out within a quasigeostrophic (QG) assimilation system using artificial data.

A related problem motivates the second main question for this thesis. On the basis of the products available from the data assimilation process ensemble forecasts are carried out. These ensemble forecasts require the specification of the set of initial states in a way consistent with the error statistics estimated from the 4D–Var process. The method adopted here, denoted sampling method, uses so–called singular vectors (SVs) computed on the basis of the current estimate of the analysis error statistics to specify a set of initial perturbations that is multivariately normal with covariance structure completely consistent with the information contained in the SVs. The method has been implemented within an operational ensemble prediction system (EPS) to study properties related to the dynamics of forecast errors and quantify its impact in terms of skill scores.

The results obtained from 4D–Var within the QG/4D–Var system suggest that a flow–dependent background specification is beneficial especially in data sparse areas where the structure of analysis increments is dominated by the background error statistics. The RRKF is able to improve the analyses compared to the static system although a reasonable approximation to the EKF is computationally very expensive.

The performance of the sampling method within the operational EPS is similar to the present operational method for generating initial perturbations. Increasing the number of perturbations and/or singular vectors improves the estimated forecast error covariance structures and shows a small positive impact in terms of skill scores.
Zusammenfassung

Numerische Wettervorhersage (NWP) erfordert die genaue Spezifikation des Ausgangszu-
standes der Atmosphäre, die sogenannte Analyse. Um diese Analyse zu erhalten, wer-
den Informationen von den Beobachtungen sowie von den vorhergehenden Modellpro-
gnosen durch einen statistischen bezeichneten Prozeß kombiniert Datenaassimilation. In op-
erationellen Datenassimilationssystemen wird die Analyse durch eine variationelle For-
uml;mulierungen der Methode der kleinsten Quadrate (least–squares) berechnet, wobei die
Modelldynamik als starke Nebenbedingung fungiert. Dieses Verfahren wird als vier–
dimensionale variationelle Datenaassimilation (4D–Var) bezeichnet, da es den Modellzu-
stand global anpasst und dabei die zeitliche Verteilung der Beobachtungen miteinbezieht.

Die Formulierung der 4D–Var ist von der Spezifizierung der Fehlerstatistiken abhängig, die
notwendig sind, um die Information aus den vorhergehenden Modellprognosen, dem back-
ground und den Beobachtungen zu kombinieren. Dabei ist die Spezifization der “back-
ground Statistik” von zentraler Bedeutung. In den operationellen Datenassimilationssyste-
men ist diese background Statistik im Wesentlichen statisch (d.h. zeitlich konstant) und fol-
glich unabhängig von den Unsicherheiten des aktuellen Zustands der Atmosphäre. Durch
die formale Äquivalenz von Variationassimilation und der zugehörigen Schätztheorie ist
e jedoch möglich, dynamische (d.h. strömungsabhängige) background Statistiken in die
4D–Var zu integrieren.

Auf Grundlage dieser Äquivalenz wird in dieser Arbeit die Auswirkung von unter-
schiedlichen dynamischen background Formulierungen im Vergleich zu einer statischen
Formulierung in der 4D–Var studiert. Im Mittelpunkt steht dabei die Quantifizierung der möglichen Verbesserung in der Genauigkeit der Analyse durch die Verwendung einer dyna-
maschen background Formulierung. Diese dynamische Formulierung ist aus der Kalman
Filter Theorie vorhanden. Zusätzlich wird das Potential einer suboptimalen, aber dennoch
dynamischen background Spezifizierung, des sogenannten Reduced–rank Kalman Filters
(RRKF) untersucht. Diese Experimente werden in einem quasigeostrophischen (QG) Assi-
minationssystems auf Grundlage künstlicher Daten durchgeführt.

Ein formal ähnliches Problem motiviert die zweite Hauptfrage für diese Arbeit: Auf
Grundlage der Resultate aus der Datenaassimilation werden Ensembleprognosen

durchgeführt. Diese Ensembleprognosen erfordern die konsistente Spezifikation der unter-
schiedlichen Ausgangszustände in Übereinstimmung mit den Fehlerstatistiken aus dem 4D–

Var ProzeßDie Methode, die hier angenommen wird, bezeichnet Stichprobenverfahren,
verwendet also – die benannten Singulären Vektoren (SVs) berechnet auf der Grundlage
von die gegenwärtige Schätzung der Analyse Störung Statistiken, um einen Satz Aus-
gangsstörungen zu spezifizieren, der multivariat Normal mit der Kovarianzstruktur ist,
die mit den Informationen vollständig gleichbleibend ist, die in der SVS enthalten wer-
den. Die Methode ist innerhalb eines funktionsfähigen Ensemblevorhersagesystems (EPS)
um die Eigenschaften zu studieren eingeführt worden, die auf der Dynamik der progno-
tizierten Störungen und seine Auswirkung in Fähigkeit Kerben ausgedrückt quantitativ zu
bestimmen bezogen werden.
Chapter 1

Introduction

Although there is no doubt that the practise of weather forecasting has improved immensely over the last decades weather forecasts are inherently uncertain (see, e.g., Bengtsson 1999; Holligsworth et al. 2002). These forecasts are routinely produced with numerical weather prediction (NWP) models of the atmosphere. The uncertainty of forecasts using NWP models is caused through uncertainties in the specification of the initial state as well as errors in the model itself. Both sources of errors deteriorate the quality of the forecast systematically and thus limit the time interval over which accurate numerical weather forecasts are possible (see, e.g., Tribbia and Baumhefner 1988). Additionally, uncertainties associated with lateral boundary conditions, such as the sea–surface temperature, might deteriorate the quality of the forecast.

In view of the work by Lorenz (Lorenz 1963, 1969), the specification of the initial state of the model, denoted analysis, has been the focus of numerous research efforts. This analysis is obtained from data assimilation (DA) systems based on statistical and physical assumptions (see, e.g., Lorenc 1986; Daley 1991; Ghil and Melanotte-Rizzoli 1991; Talagrand 1997). Within these DA systems, statistical methods based on estimation theory have replaced conceptually simpler approaches like Cressman analysis (see, e.g., Courtier 1997; Bouttier and Courtier 1999). In modern DA systems variational formulations of estimation theory are used to produce the analysis using information contained in observations, previous forecasts of the NWP model (denoted background) as well as known physical properties of the system (e.g., geostrophic or hydrostatic balance). All these sources of information contribute to the analysis but at the same time are subject to errors. Thus one has to take into account the uncertainties associated with the observations (observation error), the previous NWP forecast (background error), and the model dynamics (model error).
1.1 Uncertainty in atmospheric models

It is well known that present NWP models and the atmosphere (Lorenz 1996, 1993) are fundamentally chaotic. Neglecting model deficiencies, this fact a priori imposes a limit upon useful deterministic weather forecasts. In view of the chaotic behavior of the NWP system the (most) important component of forecast error results from errors in the estimation of the initial state. Although these initial errors have been reduced substantially by taking more observations and also more accurate observations probably the most important issue is to extract the maximum amount of information from a given set of observations. Sophisticated algorithms have been developed to exploit this issue and thus give the most accurate specification of the state of the atmosphere at a given time. A better specification of the initial state should improve the quality of subsequent forecasts and also extend the range of useful forecasts.

Assuming that the problem of finding the best initial state could be solved there still remains the problem that all NWP models, although fairly high dimensional, still are only models of the real atmosphere and hence are subject to model errors.

It is in view of these sources of errors and the corresponding uncertainty in the estimation of the initial state of the flow that it seems natural to address the prediction problem in the context of a probabilistic prediction system (PPS). For problems related to the evaluation of PPS refer to Talagrand and Vautard (1998). Approaches based on PPS have led to the field of ensemble prediction where a number of forecasts is performed on the basis of initial conditions that are consistent with the current estimate of the initial uncertainties.

Fig. 1.1 illustrates conceptually modern NWP: On the basis of observations, an NWP model as well as information from previous forecasts (denoted, background field) the
analysis is computed using an algorithm that basically solves a constrained optimization problem (see, section 1.2). Having solved this optimization problem and thus computed the current analysis the actual weather forecast is carried out. In the past the main emphasise has been on producing a single (high-resolution) forecast starting from the best initial state (i.e., the analysis). This approach is denoted as "deterministic" forecast in Fig. 1.1. However, in view of the uncertainties associated with the estimation of the analysis and the mathematics involved an ensemble forecast using a PPS seems more appropriate (see, section 1.3).

### 1.2 Variational data assimilation

Variational assimilation methods effectively deal with the solution of (linear) estimation problems thereby incorporating model dynamics (see, e.g., Courtier 1997). The basic idea is to adjust globally a model solution to the observational information available over the so-called assimilation interval. The misfit between this solution, denoted analysis \( x^a \), and the observations \( y \) (as well as the background, see 2.1) is measured by a scalar costfunction denoted as \( J \). This scalar function is defined as the sum of squared differences between the model values \( x_l \) and the observations \( y_l \) at time \( t \):

\[
J \equiv \sum_{l} \alpha_l (y_l - x_l)^2
\]

here \( \alpha_l \) are weights (i.e., variances) reflecting the accuracy of the observation \( y_l \). The model solution that minimizes the cost function (i.e., the analysis) is uniquely defined by its initial condition and hence the minimization of \( J \) is performed with respect to the initial conditions: \( x_{t=0} \equiv x_0 \). In RMS sense, this solution \( x^a \) is optimal. Assuming that the probability density functions (pdf) for observation (and background) errors are Gaussian, \( x^a \) is also the maximum likelihood estimator of the (unknown) truth \( x^t \).

In a global sense this analysis problem is under-determined and one has to rely on an a priori estimate of the model state referred to as the background field \( x^b \). Moreover, by including this background information into the current analysis one is able to accumulate in time the (observational) information from the past into the present model state.

In recent years state of the art three dimensional (3D–Var) and four dimensional variational data assimilation (4D–Var) systems, that globally adjust a model solution to the available observations, have replaced former sequential methods at a number of operational forecasting centers (e.g., at the European Centre for Medium–Range Weather Forecasts (ECMWF; Rabier et al. 2000), at the National Center for Environmental Prediction (Par- rish and Derber 1992), or at the Met. Office (Lorenc et al. 2000)). In present operational 3D/4D–Var systems the statistics relating the background information \( x^b \) are, although highly sophisticated, essentially flow–independent and specified through so–called struc-
ture functions (see, e.g., Rabier et al. 1998; Derber and Bouttier 1999). In other words, these statistics do not reflect information and uncertainties of the current state of the atmosphere and hence will be regarded as static background formulation.

Theoretically, flow–dependent background error statistics are available from estimation theory (see, e.g., Cohn 1997). Specifically, the extended Kalman Filter (EKF) theory (see, e.g., Jazwinski 1970) provides a systematic way for computing flow–dependent error structures in weakly nonlinear systems (see, e.g., Ghil et al. 1981). The correspondence between the EKF and 4D–Var in view of (linear) estimation theory is exploited in this study to include flow–dependent background error statistics in 4D–Var.

To illustrate the relevance of the background specification results from a (chaotic) three–variable Lorenz model (Lorenz 1984) are presented (Beck 2000). Within these experiments both background and observation errors are crudely specified by (static) diagonal covariance matrices proportional to the identity matrix — static background error covariance matrix (B): $B = \sigma^2_b I$ and observation error covariance matrix (R): $R = \sigma^2_o I$). The 4D–Var algorithm as well as the role of $B$ and $R$ are described in chapter 2.

As an example, two assimilation experiments with different specifications for $B$ are given. The assimilation interval is six (dimensionless) time units and the number of observations is $L = 12$. For a given number of observations the specification of these statistics forces the analysis to be either closer to the (generally) smoother background field or to the (noisy) observations. In both experiments the first guess for starting the minimization algorithm has been taken different from the background (note that in current operational practise usually the background is taken as the first guess). Fig. 1.2 shows the $Z$ component of the corresponding trajectories for the first guess (fg), the analysis (ana), and the background (bg) as a function of time. The observations $y_t$ are shown as crosses (obs).

![Figure 1.2: 4D–Var in the Lorenz 1984 model: Assimilation experiment involving both observation and background term with (a) large ($\sigma^2_b = 1.0, \sigma^2_o = 1.0$) and (b) small ($\sigma^2_b = 0.1, \sigma^2_o = 1.0$) background error variances. Shown are the trajectories ($Z$ component) of the first guess (fg), the analysis (ana), and the background (bg) as a function of time. The observations $y_t$ are shown as crosses (obs).](image-url)
units. The observations are included as blue crosses. Comparing Figs. 1.2a and 1.2b clearly shows the impact of the specification of the background error variance, which is ten times larger in Fig. 1.2a. Hence, the analysis remains much closer to the background in Fig. 1.2b. On the other hand, a further increase of the background error variance (compared to Fig. 1.2a) results in qualitatively the same picture as in Fig. 1.2a (not shown). Clearly, this is a consequence of imposing the model dynamics as a strong constraint. These results demonstrate, in the framework of this conceptual model, that the specification of the background error statistics (i.e., $\sigma_b$ in this example) has a substantial impact on analysis obtained.

1.3 Atmospheric predictability

An accurate analysis of the current state of the atmosphere is the basis for forecasting with a NWP model. Vilhelm Bjerknes formulated the problem of weather forecasting in terms of mathematical equations as an initial–value problem (Bjerknes 1904). Bjerknes also referred to the prediction of future atmospheric states as the "ultimate problem in meteorology". However, this deterministic approach lacks the ability to quantify the potential errors of the forecast a priori. Since the pioneering work of Lorenz it is accepted that atmospheric predictability is limited by the inherent error growth as a consequence of the nonlinearity and instability of atmospheric dynamics (Lorenz 1963). It is well know (Leith 1978) that the sensitivity to the specification of the initial state due to this inherent error growth limits the predictability of the atmosphere. In recent years, probabilistic approaches, and thus PPS, to predicting the future state of the atmosphere on the basis of the associated pdfs are widely used. These probabilistic approaches are know as ensemble prediction because a (small) ensemble of forecast is performed instead of only one forecast. One the basis of the set of ensemble members properties of the pdf can be derived. The concept of ensemble prediction is basically a sort of Monte Carlo (MC) approach consisting of quasi–random sampling of the initial state pdf producing an ensemble of initial states and than evolving these states by the NWP model. This approach has been proposed by Leith (1974) in the field of meteorology as an alternative to stochastic dynamic prediction (see, Epstein 1969) and the solution of the Liouville equation (see, e.g., Ehrendorfer 1994).

Following the work by Edward Lorenz (Lorenz 1965), atmospheric predictability is often determined by the divergence of two predictions as a result from minor changes in the initial state. Obviously this dependence on the initial conditions implies that even a deterministic system is only predictable for a certain time period given small errors in the specification of the initial state. To illustrate the internal error growth of an NWP model Fig. 1.3 shows total temperature variance at 850hPa (i.e., the sum over roughly 5000 gridpoints) as a function of forecast time estimated from the ECMWF Ensemble Prediction System (EPS). It is clearly seen from Fig. 1.3 that small errors at initial time
grow rapidly for the first two days of the forecast. Extending the curves beyond forecast day 10 leads to a saturation of the variance due to the nonlinear processes of the model. Note that in a linear model the error growth would be unbounded. To further illustrate the rapid growth of small initial time errors Fig. 1.4 shows T850 variances at (a) initial time and (b) at forecast day 5. The contour interval is 10 times larger in Fig. 1.4b.

Referring back to Fig. 1.1, there is a strong connection between data assimilation and probabilistic forecasts. It is on the basis of the results from data assimilation process, namely the analysis and the corresponding estimate for the analysis uncertainty, that the
initial time ensemble is constructed. This connection will be further exploited in this work in the context of the so-called sampling technique (see, chapter 4).

1.4 Motivation and outline

The primary motivation for the present study has been the question of the potential improvement of 4D–Var by incorporating flow–dependent error statistics within the assimilation system. In other words, for a given set of observational information the behavior of the 4D–Var system with respect to changes of the background statistics is investigated. Basically, this question amounts to understanding how the provided data can be used as efficiently as possible.

The formulation of the background error statistics in an operational assimilation system has been the subject of major research interest in that past and still is the subject of present research efforts. It is however impossible in operational context to study the potential benefit of a dynamic background formulation without crude assumptions. One possibility is to use simplified dynamics for the propagation of the error statistics (see, Ehrendorfer and Bouttier 1998b). As an alternative one might try to approximate the equations governing the error dynamics as is done in various forms of ”approximated” Kalman filters. Among these suboptimal filters is the so–called Reduced Rank Kalman Filter (RRKF) introduced by Rabier et al. (1997).

Within this thesis, a simple yet reasonably realistic model of the atmospheric circulation based on quasigeostrophic (QG) dynamics will be employed. To study the potential benefit of a dynamic background formulation the error statistics are computed from a full EKF. This allows for studying the performance of the RRKF compared to the static system and the system using the dynamic background from the EKF. Theoretically, one would expect the static system and the EKF to form the lower and upper bounds for the performance of the system with different background specifications. In between these two extremes the RRKF is designed to estimate flow–dependent background error statistics on the basis of so–called Hessian singular vectors (HSVs; see, section 3.2). This work is carried out in a perfect model scenario using artificial data sampled from a truth–run of the model. This allows for computing the statistics as well as the errors with respect to this truth run.

The second main question is related to estimation of forecast error covariances in an operational EPS. On the basis of the products of the DA cycle an initial ensemble is generated that is consistent with the structure of analysis errors as implied through the analysis error covariance matrix. The method adopted here uses a Monte Carlo approach, denoted sampling technique, to sample initial perturbations from the singular vectors designed to approximate the analysis error covariance structure. This sampling technique is described in detail in section 4.2. Selected experiments using the sampling technique in comparison to the method applied operationally at ECMWF are discussed in section 8.1. Addition-
ally, the impact of increasing the number of ensemble member on the estimated covariance
structure is investigated.

The outline of this work is as follows: Chapter 2 introduces the concept of four di-
mensional variational data assimilation with emphasis on the application within this work. The following chapter 3 presents a brief review of the concept of Kalman filtering an its rel-
evance in both data assimilation as well as covariance prediction. Chapter 4 describes the
so–called sampling technique as a means for constructing initial time ensemble members
consistent with analysis error structure. The experiments concerning data assimilation are
carried out in the framework of a quasigeostrophic assimilation system that is introduced
in chapter 5. With regard to the primary question of this work, chapter 6 introduces the dif-
ferent background specifications investigated as well as the RRKF algorithm in detail. The
main results from the data assimilation experiments are summarized in chapter 7. Chapter
8 gives a brief overview of the application of the sampling technique in the framework of
the ECMWF EPS. Finally, chapter 9 summarizes the key features of this work and restates
the overall conclusions on the basis of the experiments that have been carried out.
Chapter 2

Four–dimensional variational data assimilation (4D–Var)

4D–Var aims at defining the model initial conditions of an NWP system using all available pieces of information over some period of time (e.g., 6 or 12 hours). These pieces of information include observations taken within the finite time interval as well as information from previous assimilations (i.e., background field). On the basis of this information and by taking into account the model dynamics, 4D–Var globally adjusts the model solution in an optimal way by solving a (global) constrained optimization problem.

![4D-Var Diagram](image)

**Figure 2.1: Illustration of 4D–Var (adapted from Bouttier and Courtier 1999).**

Let us now illustrate the concept of 4D–Var for a simple scalar problem: Fig. 2.1 shows conceptually the 4D–Var problem in terms of a scalar variable X to be estimated on the
2.1 4D–Var: Method

Implementation of 4D–Var requires the specification of a cost function that quantifies the misfit between the observational information and the present estimate of the analysis trajectory. More precisely, this scalar cost function measures the misfit between the observations $y_l$ and the model solution $x_l$ at times $l$ as well as the departure from the background field $x^b$ at initial time $l = 0$. The misfit between the model state vector and the observations is computed in the observation space through the use of the (nonlinear) observation operator $\mathcal{H}_l$. The subscript $l$ denotes the quantities at any given observation time, distributed among $L + 1$ dates in the assimilation interval. The covariance matrices relating the background ($P^f$) and the observations ($R$) to the model solution have to be specified externally or, in case of $P^f$, can be modeled by a suitable variant of the Kalman filter (see, section 6.3). The analysis problem is defined as finding the $n$–dimensional model state $x_0$ that minimizes the following nonlinear cost function:

$$J(x_0) = \frac{1}{2} (x_0 - x^b_0)^T (P^f)^{-1} (x_0 - x^b_0) + \frac{1}{2} \sum_{l=0}^{L} \{y_l - \mathcal{H}(x_l)\}^T R_l^{-1} \{y_l - \mathcal{H}(x_l)\}$$

subject to the strong constraint that the sequence of model states $x_l$ must be a solution of the nonlinear forecast model $\mathcal{M}$:

$$x_l = \mathcal{M}_l(x_0).$$

From eq. (2.1.1) it is evident that the costfunction is actually the sum of two terms, namely, the background term $J^b$ and the observation term $J^o$. Note that it is possible to include other terms for controlling gravity–waves (i.e., $J^c$, see, Gauthier and Thepaut 2001) or an appropriate formulation of model uncertainties (Talagrand 2002, personal communication). Within the framework of conceptual models, weak constraint variational formulations have been proposed especially for strongly nonlinear models as well as for including errors within the dynamical model (see, e.g., Evensen and Fario 1997; Ahrens 1999).
2.1 4D–Var: Method

The solution of the cost function eq. (2.1.1) is a global optimization problem subject to a nonlinear constraint and thus in general very difficult to solve. To simplify the solution of eq. (2.1.1) the following assumptions are made: Assuming that the action of the model operator $M_l$ can be expressed as a sequence of operators reflecting causality of nature, $M_l$ is written as a product of operators:

$$M_l = M'_l M'_{l-1} \ldots M'_1$$  \hspace{1cm} (2.1.3)

where $M'_l$ is the operator that propagates the state vector one time step ahead (from $t_{l-1}$ to $t_l$) and $M_0 \equiv I$. The most important assumption in terms of simplifying eq. (2.1.1) and one of the major drawbacks of 4D–Var in general is the assumption of the validity of tangent–linear hypothesis (see, Errico et al. 1993): It is assumed that the nonlinear model operator $M_l$ (see, eq. 2.1.2) can be linearized in the vicinity of the background trajectory $x^b$:

$$M_l(x_0) \approx M_l(x^b_0) + \left. \frac{\partial M_l}{\partial x_0} \right|_{x_0} (x_0 - x^b_0)$$  \hspace{1cm} (2.1.4)

with $M_l$ denoting the tangent–linear (TL) model operator obtained as the derivative of the nonlinear dynamics along the reference trajectory (i.e. the background state $x^b$). These assumptions (i.e., causality and validity of the TL model) are common to all implementations of 4D–Var.

However, for the purpose of this study further assumptions are made to simplify the numerical implementation. The next assumption concerns the specification of the observation operator $H$. For the purpose of this study this operator is linear and thus will be denoted by $H$. Note that for nonlinear $H$ again the TL hypothesis has to be applied and $H$ has to be linearized in the vicinity of $x^b$ analogous to eq. (2.1.4). The last simplification assumes that $R_l$ is time–independent and will be denoted by $R$. This assumption does not degrade the realism of the system for the purpose of this study.

Subsequently, these assumptions are applied to the 4D–Var cost function (eq. 2.1.1). To allow for taking into account observations at specified time–slots $l$ only, an indicator $r_l$ is introduced: This indicator is equal to one ($r_l = 1$) in the case that observations are available at time–slot $k$ and equal to zero otherwise. With these assumptions and including the indicator $r$, eq. (2.1.1) is rewritten as:

$$\mathcal{J}(x_0) = (x_0 - x^b_0)^T (P^f)^{-1} (x_0 - x^b_0) + \sum_{l=0}^{L} r_l \{ y_l - Hx_l \}^T R_l^{-1} \{ y_l - Hx_l \}$$  \hspace{1cm} (2.1.5)

The gradient of the (modified) cost function (eq. 2.1.5) is obtained as:
2.2 Estimating analysis quality: Hessian of the costfunction

The evaluation of the gradient of the costfunction is done by making use of the explicit adjoint model operators $M_l^T$ (see, e.g., Talagrand and Courtier 1987; Thépaut and Courtier 1991). Thus the computation of equations 2.1.5 and 2.1.6 is computationally feasible.

As discussed in section 1.2, in current operational practise the background error covariance matrix $P_f$ is essentially flow–independent and will be denoted as $B$. This $B$ matrix will be referred to as the static background error covariance matrix. The 4D–Var system that makes use of this static $B$ matrix will be referred to as the static system (see also the simple example in 1.2).

2.2 Estimating analysis quality: Hessian of the costfunction

One weakness of 4D–Var, or variational assimilation methods in general, is that the algorithm does not provide an estimate of the quality of the analysis. In other words, the analysis error covariance matrix $P_a$ is not available directly after the assimilation has been carried out. It is however possible to infer $P_a$ from the matrix of second derivatives, known as the Hessian, of the cost function (eq. 2.1.5). Rabier and Courtier (1992) show that the relationship between the Hessian of the costfunction and $P_a$ is

$$P_a = (\mathcal{J}^\prime\prime)^{-1} = (\nabla \nabla \mathcal{J})^{-1} \quad (2.2.7)$$

If the costfunction $\mathcal{J}$ is exactly quadratic the Hessian does not depend on the current analysis $x_a$ and could in principle be determined from $\mathcal{J}$. However, the costfunction (eq. 2.1.5) is nonlinear because of the nonlinear model operator $M$ and therefore $\mathcal{J}$ (eq. 2.1.5) is not exactly quadratic. Hence the Hessian $\mathcal{J}^\prime\prime$ does depend on the (current) value of $x_a$.

Following Ehrendorfer and Bouttier (1998a) the Hessian $\mathcal{J}^\prime\prime$ of the costfunction (eq. 2.1.5) is estimated by taking differences of gradients computed from eq. (2.1.6). Specifically, the difference between two gradients gives an approximation to one column of the Hessian. To estimate the $i$–th column of the Hessian $\mathcal{J}^\prime\prime$ at the point $x_a$ (i.e., the current analysis) we compute:

$$(\mathcal{J}^\prime\prime)_i \approx \frac{(\nabla \mathcal{J})_{x^a+h_i} - (\nabla \mathcal{J})_{x^a}}{|h_i|} \quad (2.2.8)$$

where $h_i$ is a vector of length $n$ (i.e., the size of the model state vector) consisting of zeros except for the entry $i$ where $h_i \sim O(10^{-10})$. Two remarks apply to eq. (2.2.8):
For a quadratic function $J$ eq. (2.2.8) is fulfilled exactly. For the cost function defined in eq. (2.1.5) the result of eq. (2.2.8) has a weak dependence on the specification of $h_i$. Experiments using different choices of $h_i$ have shown a negligible impact in terms of the analyses obtained (not shown).
Chapter 3

Error covariances in data assimilation and ensemble prediction

As a starting point for representing uncertainty within an NWP model only the first two moments of the pdf are considered. In other words, this amounts to assuming that the errors encountered are distributed normally. The field of estimation theory provides a rigorous framework for dealing with problems encountered in data assimilation, and, in particular, for linear systems and Gaussian pdfs. In view of atmospheric data assimilation, the most prominent algorithm in this context is the well known Kalman Filter and its extension to nonlinear analysis updates, the Extended Kalman Filter. Cohn (1997) provides an outstanding introduction to the application of estimation theory in the field of data assimilation.

An important result is that under certain assumptions (linearity, Gaussian errors, and perfect model assumption) the 4D–Var analysis is equal to the EKF analysis at the end of the assimilation interval (see, e.g., Jazwinski 1970). On the basis of this equivalence we apply the EKF to the propagation of the error covariances only and compute the analysis using 4D–Var.

3.1 The Extended Kalman Filter (EKF)

The Extended Kalman Filter is a sequential assimilation method based on least–squares theory (Kalman 1960; Kalman and Bucy 1961; Maybeck 1979). For the purpose of this work we will consider only the error covariance forecast step of the EKF:

\[ P^f = M_t P^a M_t^T + Q \]  \hspace{1cm} (3.1.1)

where \( P^f \), \( P^a \) and \( Q \) are the \( n \)-by–\( n \) covariance matrices of forecast, analysis and model error and \( M_t \) is the tangent linear model resolvent for the time–interval \([0 : t]\) (i.e., \( M_t \) is a mapping between the initial state at time \( t = 0 \) and the final state at time \( t \)). Numerical evaluation of eq. (3.1.1) is very expensive since it requires \( n \) forecasts of the TL and
3.2 Hessian singular vectors

adjoint model and thus is not feasible even for "small" dimensional NWP models. To circumvent this problem simplifications have to be introduced (see, e.g., Todling and Cohn 1994; Todling et al. 1998). The line of attack for these simplifications is twofold:

- approximate the dynamics governing the error propagation
- reduce the dimension of the error space

A number of algorithms have been introduced for simplification of eq. (3.1.1). Generally speaking, these algorithms belong to one of the groups mentioned above. The most prominent approaches are either based on reduced–rank approximations of the full error covariance matrices or on Monte Carlo approaches for estimating the predicted covariance structure on the basis of a (small) finite–size ensemble of model states. For filter algorithms based on reduced–rank approximations reference is made to (e.g., Verlaan and Heemink 2001, Farrell and Ioannou 2001a). The class of MC approaches is known as Ensemble Kalman Filter (enKF) since it represents the error statistics by an ensemble of randomly generated model states (see, e.g., Evensen (1994) as well as Houtekamer and Mitchell (1998) and references therein). The implementation of an exact EKF based on QG dynamics is described by Ehrendorfer and Bouttier (1998b).

The method adopted here is known as Reduced–Rank Kalman Filter (RRKF; Rabier et al. 1997). Note that the RRKF has been denoted Simplified Kalman Filter in Rabier et al. (1997). The basic idea of the RRKF is to combine a reduced–rank estimate of the forecast error covariance structure with a static, flow–independent, background error model. The RRKF has been proposed by Courtier (1993) for use within an operational data assimilation system. A detailed description of the RRKF as well as the implementation used within this study is given in section 6.4.

3.2 Hessian singular vectors

In a perfect model scenario the covariance forecast equation of the Kalman filter (eq. 3.1.1) becomes

$$\hat{P}_f = M_t P^a M_t^T$$

(3.2.2)

Let us define the following generalized eigenproblem:

$$M^T C^T C X_0 = (P^a)^{-1} X_0 \Lambda \quad \text{with} \quad X_0^T (P^a)^{-1} X_0 = I$$

(3.2.3)

where $C^T C$ is a norm defining metric and $M$ denotes the TL resolvent at the optimization time $t_{opt}$: $M \equiv M_{t=t_{opt}}$. The solution to this eigenproblem are the so–called Hessian singular vectors (Barkmeijer et al. 1998) denoted as $X_0$ with corresponding eigenvalues
written in matrix form as $\Lambda$. Eq. 3.2.3 implies that the set of vectors $X_0$ is constraint at initial time by the $P^a$ matrix. Multiplying eq. (3.2.3) from the left with $CMP^a$ we obtain

$$ CMP^aM^T C^T CMX_0 = CMX_0\Lambda. $$

(3.2.4)

Defining the evolved eigenvectors as $X = CMX_0$ and using eq. (3.2.2) gives:

$$ C\hat{P}^f C^T X = X\Lambda $$

(3.2.5)

Eq. 3.2.5 implies that the set of evolved eigenvectors $X$ are the eigenvectors of the forecast error covariance matrix $\hat{P}^f$ in the norm $C^T C$ at the optimization time $t_{opt}$. It is this property of the HSVs that makes them a primary candidate for estimating forecast error covariances (Ehrendorfer and Tribbia 1997).

HSV's are of fundamental importance for the research described in this work as they appear in both the context of ensemble prediction (see, sections 3.2.1 and 4.2) as well as the formulation of the RRKF (see, section 6.4).

### 3.2.1 Singular vectors and analysis errors

This section briefly reviews the correspondence between (Hessian) singular vectors (SVs) and analysis errors as described through the analysis error covariance matrix $P^a$. In view of eq. (3.2.3) we obtain a square–root decomposition of $P^a$ as:

$$ P^a = X_0X_0^T. $$

(3.2.6)

The significance of eq. (3.2.6) is that this SV–decomposition of $P^a$ (Ehrendorfer 1999) evolves (under tangent–linear dynamics) into the eigendecomposition of the forecast error covariance matrix. This property is immediately evident when (3.2.6) is multiplied from the left with $CM$ and from the right with $(CM)^T$, yielding:

$$ CMP^aM^T C^T = XX^T $$

(3.2.7)

Eq. (3.2.7) is equivalent to the eigendecomposition of $P^f$ (in the norm $C$). This fact is seen when (3.2.7) is compared to eq. (3.2.5) showing that the evolved SVs $X$ are the eigenvectors of $P^f$ with eigenvalues $\Lambda$. Equations (3.2.5) and (3.2.7) are equivalent because of the orthogonality of the final–time SVs:

$$ X^T X = (CMX_0)^T (CMX_0) = X_0^T M^T C^T CMX_0 = X_0^T (P^a)^{-1}X_0\Lambda = \Lambda. $$

(3.2.8)
In other words, if $P^a$ is, as in (3.2.6), represented in terms of the initial SVs, and if this representation is truncated to a representation in terms of $N$ SVs, this truncated representation will lead, when time–evolved, into the eigenstructure (represented by $N$ eigenvectors) of the forecast error covariance matrix (Ehrendorfer 1999).
Chapter 4

Multi–Gaussian sampling in ensemble prediction

Producing an accurate forecast is the most challenging task of modern NWP centers. Ideally, one would like to have not only an accurate forecast but also an estimate for its potential accuracy on a day–to–day basis. On the basis of these estimates valuable parameters such as probabilities for certain events ideally, extreme events, can be derived. Ensemble prediction tries to accomplish this task by running a series of forecasts each starting from slightly different initial states. Mathematically, the sample of initial states has to be consistent with the current estimate of the initial time pdf namely the analysis and the corresponding analysis error covariance matrix. The most important question in an operational context is how to generate these initial states on the basis of the information provided from the data assimilation process.

4.1 Ensemble prediction: Generation of initial perturbations

Finite–time optimal perturbations, described by the singular vectors of tangent–linear model resolvents, are a primary candidate for generating initial–time members in ensemble prediction. One of the rationales for using SVs to generate perturbations in ensemble prediction is that they evolve, at optimization time, into the eigenvectors of the forecast error covariance matrix (Ehrendorfer and Tribbia 1997). Consequently, they describe important information about the time–evolving pdf of the model state vector.

For example, ECMWF operationally uses a method based on SVs for the generation of initial–time perturbations (see, Molteni et al. 1996). In the methodology considered here for selecting ensemble members from an initial pdf the initial–time SVs are used to represent the covariance structure of the pdf. The procedure described is referred to as SV–based Monte Carlo method (Ehrendorfer 1999), since it combines random sampling
4.2 The singular vector based Monte Carlo technique

(i.e., the Monte Carlo approach) with the analysis error covariance information provided through the SVs. Because of this random sampling we will refer to the method as sampling method. The sampling method adopted here uses a set of (say) N SVs to describe a square–root decomposition (restricted to N vectors) of the covariance matrix that is implied by the initial–time metric used in the SV computation. Given such an SV–decomposition of the analysis error covariance matrix a set of initial perturbations (size M) can be obtained that is multivariately normal with covariance structure completely consistent with the information contained in the N SVs.

4.2 The singular vector based Monte Carlo technique

In view of the limited information available about the initial pdf in an ensemble–prediction context, it is reasonable to assume that the analysis errors are normally distributed. Information about their covariances is contained in the analysis error covariance matrix $P^a$. From eq. (3.2.6) a square–root decomposition of $P^a$ is available through the initial–time HVSs: $P^a = X_0^T X_0$. On the basis of this SV–decomposition a set of initial perturbations is constructed as follows. Let us consider a multinormal random variable $x$ with mean $\bar{x}$ and covariance $V$ defined as:

$$x = \bar{x} + V^{1/2} q \quad \Rightarrow \quad x \sim \mathcal{N}(\bar{x}, V) \quad \text{if} \quad q \sim \mathcal{N}(0, I) \quad (4.2.1)$$

In view of eq. (3.2.6), the sampling technique consists of specifying the set of M initial perturbed states $x_i$ from N HVSs as:

$$x_i = x^a + X_0^{(N)} q_i \quad i = 1, 2, ..., M \quad (4.2.2)$$

where $X_0^{(N)}$ indicates the matrix containing as columns the N HVSs available, and $q_i$ is a realization from $\mathcal{N}(0, I)$. The mean value $\bar{x}$ corresponds to the analysis $x^a$. The sample covariance of an ensemble of size M generated by (4.2.2) converges (with increasing M) – by construction – towards:

$$(P^a)^{(N)} \equiv (X_0^{(N)}) (X_0^{(N)})^T \quad (4.2.3)$$

where $(P^a)^{(N)}$ is the version of $P^a$ (see eq. 3.2.6) that is truncated in terms of the N leading HVSs. Consequently, the set of perturbed states $x_i$ generated through (4.2.2) is a set of M realizations from $\mathcal{N}(x^a, (P^a)^{(N)})$. Therefore, the sampling technique produces a random sample of M perturbations from the initial pdf that is multivariately normal with covariance structure consistent with the $P^a$ information (restricted to the leading N SVs).

The method will be illustrated in ensemble experiments within the framework of the ECMWF Integrated Forecasting System (IFS). By construction, the method possesses
4.2 The singular vector based Monte Carlo technique

basic characteristics similar to the operationally applied rotation method (Molteni et al. 1996). However, the sampling method has the appealing property that the number of ensemble members $M$ can be chosen independently of the number of SVs available. Within the rotation method the number of ensemble members basically is twice the number of SVs. Additionally, the sampling method directly makes use of the properties of the initial–time SVs and thus generates a statistically consistent ensemble of initial states. It will be shown that increased $M$ is beneficial when properties (e.g., correlation structures) of the time–evolving ensemble are estimated, or in terms of reducing the number of outliers. The dependencies on the ensemble size $M$ and the number of SVs $N$ used for constructing the initial covariance structure is investigated (see, section 8.1).
Chapter 5

The quasigeostrophic (QG) assimilation system

The investigation of the questions outlined in section 1.4 requires a simple yet reasonably realistic model of the atmosphere. Specifically, the dimension of the model has to be fairly small (say, O(1000)) to solve explicitly the equations involved in 4D–Var and the EKF. To that aim the assimilation system used within this study is based on quasigeostrophic (QG) dynamics. A brief description of this QG model is summarized in the following section 5.1.

5.1 The quasigeostrophic model

The extratropical circulation can be described approximately by the QG equations (see, e.g., Pedlosky 1987; Holton 1992). The 4D–Var system developed is based on the quasigeostrophic potential–vorticity (PV) model developed by Marshall and Molteni (1993). The model is formulated spectrally on the sphere at resolution T21 with three vertical layers of 300hPa each. The midpoints of these model layers are considered as model levels corresponding to roughly 200, 500, and 800hPa (see, Figure 5.1). With this choice of the (spectral) horizontal and vertical discretization the system possesses \( n = 1449 \) degrees of freedom. This QG model has been used extensively for a broad range of studies in the field of predictability (e.g., Vannitsem and Nicolis 1997, Oortwijn 1998, Reynolds and Errico 1999, Gelaro et al. 2002), ensemble prediction (e.g., Molteni and Palmer 1993, Ehrendorfer 1999) and data assimilation (e.g., Ehrendorfer and Bouitter 1998b, Swanson et al. 1998, Swanson et al. 2000).

The quasigeostrophic model is based on the standard filtered prognostic equations for vorticity and temperature and includes dissipation \( D \) and an artificial forcing \( S \). It integrates prognostic equations for QG potential–vorticity \( q_i \) at model level \( i \) written here as:
5.1 The quasigeostrophic model

\[
\left( \frac{\partial}{\partial t} + \mathbf{v}_{\Psi_i} \cdot \nabla \right) (q'_i + f) + D_i - S_i = 0
\]  

(5.1.1)

with

\[
q'_i = \nabla^2 \psi_i + f_0^2 \frac{\partial}{\partial p} \sigma^{-1} \frac{\partial \psi_i}{\partial p}
\]

(5.1.2)

where \( v_{\Psi_i} \) denotes the geostrophic velocity and \( f = 2\Omega \sin \varphi \) is the Coriolis parameter (on the sphere). To obtain energetic consistency a constant value \( f_0 \) is required in the expression of \( q'_i \) (see, eq. 5.1.2). Additionally, the static stability \( \sigma \) is only dependent on height: \( \sigma = \sigma(p) \).

\[\begin{array}{c|c}
\text{p} & \text{q'}_1, \psi_1 \hline
200 \text{ hPa} & T_{\text{top}} \\
\text{500 hPa} & T_I \\
800 \text{ hPa} & T_{\text{II}} \\
\end{array}\]

\[\begin{array}{c|c}
\text{} & \text{q'}_2, \psi_2 \hline
\text{200 hPa} & T_{\text{top}} \\
\text{500 hPa} & T_I \\
\text{800 hPa} & T_{\text{II}} \\
\end{array}\]

\[\begin{array}{c|c}
\text{} & \text{q'}_3, \psi_3 \hline
\text{200 hPa} & T_{\text{top}} \\
\text{500 hPa} & T_I \\
\text{800 hPa} & T_{\text{II}} \\
\end{array}\]

\[\begin{array}{c|c}
\text{} & \text{T_{bot}} \hline
\text{200 hPa} & Q_{_1} \hline
\text{500 hPa} & Q_{_2} \\
\text{800 hPa} & Q_{_3} \\
\end{array}\]

Fig. 5.1: Illustration of the vertical discretization of the QG model.

Fig. 5.1 illustrates the vertical discretization of the QG model. The variables \( q'_i \) and \( \Psi_i \) are defined in the middle of the three model layers whereas the expressions for the static stability are defined with respect to the "temperature gradients" \( (T_{\text{top}} - T_I) \), \( (T_I - T_{\text{II}}) \), and \( (T_{\text{II}} - T_{\text{bot}}) \), respectively. For details about the QG model as well as the discrete model formulation and its numerical implementation reference is made to Haas et al. (1999). In addition, Ehrendorfer (2000) discusses aspects of the energy expressions relevant in this QG model.

The dissipation terms \( D_i \) are linear terms including contributions from temperature relaxation, Ekman dissipation, and horizontal diffusion. Orographic effects are included through spatially varying drag coefficients within the formulation of the Ekman dissipation at the lowest model layer. The forcing terms \( S_i \) represent constant but spatially varying sources of PV that are necessary to obtain a realistic climatology within the model. For details reference is made to Marshall and Molteni (1993).
5.2 QG/4D–Var system

The 4D–Var system (see, Haas and Ehrendorfer 2001) is built on the basis of the T21L3 QG model described briefly in section 5.1 and additional code to compute HSVs (see, section 3.2) as well as code to estimate the background error covariance matrices by means of either the RRKF or the full KF equation (see, eq. 3.1.1). Note that due to the relatively small dimension of the model state vector (i.e., 1449 degrees of freedom) all matrices involved in 4D–Var and specifically the computation of the HSVs are available explicitly. To study the impact of different background specifications (see, chapter 6) series of cycling experiments (see, section 7.2) are carried out using (artificial) observations at the beginning, in the middle and close to the end of each assimilation window (see, section 5.3). The distribution of observations is assumed constant for subsequent assimilation cycles.

Fig. 5.2 illustrates this setup for a cycling experiment with 12–hour assimilation window. The individual cycles are numbered as 1, 2, 3, ... and the assimilation windows are shown as grey boxes. The observations are denoted as colored circles. The color of the circle indicates use within the corresponding assimilation window (colored frame). Note that this configuration ensures that a particular set of observations (denoted as colored circles in Fig. 5.2) enters the sequence of assimilation cycles only once. The cycling of the background error covariances is described in detail in chapter 6.

5.3 Observational network

We consider assimilation of artificial observations sampled from the truth–run $x^t$ with pre–specified variance structure $R$. The matrix $R$ is assumed to be diagonal with observation errors depending only on height. Specifically, these errors are: $\sigma_{200} = 60m$, $\sigma_{500} = 40m$, $\sigma_{800} = 20m$ for the three model levels. At each time slot (i.e., 00/06/11 in Fig. 5.2) a fixed number of observations on an irregular grid is used. The spatial distribution of the observations varies in time within the assimilation cycle but the error statistics $R$ (i.e., the size of the observation errors) is assumed to be constant.
Figure 5.3: Observation locations (×) at three model levels (200, 500, and 800 hPa) for one particular time-slot. The total number of observations is p=326 in this case.
Fig. 5.3 shows the location of the observations (×) for the three model levels (i.e., 200, 500, and 800hPa from top to bottom) for a total number of p=326 observations. The (random) distribution is such that the probability for an observation at a randomly selected gridpoint is much higher over land than over the ocean (the ratio is roughly 4:1).
Chapter 6

Background formulation in QG/4D–Var system

Specification of the background errors is a crucial part of setting up a variational assimilation system (see, e.g., Fisher and Courtier 1995; Derber and Bouttier 1999). In current operational practise these errors are essentially taken to be flow independent for reasons of computational restrictions. Specifically, the background error statistics do not reflect the day–by–day variations of the atmospheric flow. However, assuming that a reasonable guess of the background error covariance is available one might proceed one step further and investigate the impact of a flow–dependent background specification on the performance of the assimilation system. The quantification of the potential benefit of a flow–dependent background is the primary question of this study as outlined in section 1.4. Within the QG system used here, the quantifications of this benefit is possible since the computation of the fully dynamic system (i.e., the EKF) is feasible within this model and thus an upper bound for the performance of the system is available.

6.1 4D–Var in a cycling environment

To introduce the cycling environment that forms the framework for this study we consider again the background term of the costfunction $J(x_0)$ that is minimized in variational data assimilation (i.e., $J_b$ term in eq. 2.1.1):

$$J(x_0) = \left( x_0 - x^b \right)^T (P^f)^{-1} \left( x_0 - x^b \right) + J^o$$

(6.1.1)

with $P^f$ denoting the background error covariance matrix that weights the departures of $x_0$ with respect to the background field $x^b$. To illustrate the experimental setup that has been chosen Fig. 6.1 depicts the primary steps for running a 4D–Var system in a cycling environment. Assuming that we have access to a 4D–Var system, and that batches of observations
at certain predefined time–slots are available as well as a guess for the initial background field $\mathbf{x}^b$ and errors $\mathbf{B}$ we may consider the 4D–Var system as part of a cycling environment. Note that the matrix $\mathbf{B}$ is referred to as the static background error covariance matrix. For the cycling of the background errors $\mathbf{P}^f$ one might envisage the following three configurations as illustrated in Fig. 6.1. Within this figure the cycling process means following the vectors shown in Fig. 6.1 counterclock wise: Starting from a 4D–Var analysis $\mathbf{x}^a$ and the associated analysis error covariance matrix $\mathbf{P}^a$ the cycling of the error variance is possible through any of the three branches $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$. These branches allow for investigating the impact of the background specification ranging from a purely static ($\mathbf{A}$) to a fully dynamic formulation ($\mathbf{C}$) within the RRKF ($\mathbf{B}$) bridging the gap between these two extremes. On the basis of $\mathbf{P}^a$ it is possible to incorporate dynamic (i.e., flow–dependent) background error statistics into the next assimilation cycle.

![Figure 6.1: Illustration of background formulation for cycling in 4D–Var.](image)

The time–dimension is hidden in the cycling process: One analysis cycle covers a 12h period and thus $\mathbf{x}^b$ and $\mathbf{P}^f$ estimated on the basis of the current analysis are used in the background term of 4D–Var 12 hours later (i.e., within the next cycle). This setup is closely related to the current operational practise at NWP centers (e.g., at ECMWF).

For practical reasons the most straightforward choice ($\mathbf{A}$) is to use the static background $\mathbf{B}$ for the subsequent cycles. Obviously in this case no cycling of the background errors takes place at all. In view of the EKF ($\mathbf{C}$) the cycling of the covariances is obtained from eq. (3.1.1). From a theoretical point of view we would expect this configuration to be superior to the static formulation thereby improving the quality of the analyses. In between these to extremes the RRKF ($\mathbf{B}$) is regarded as a sub–optimal means for cycling the error covariances involved. In fact it is suboptimal for $k < n$ (i.e., the number of HSVs smaller than the number of degrees of freedom).
6.2 The static system

It is well known that the specification of the (static) background error statistics has a substantial impact on the performance of operational data assimilation systems (e.g., Anderson et al. 1998). To address the impact of different B specifications within the QG model two choices for the B matrix have been considered:

- \textbf{ec} B: ECMWF 3D–Var vorticity background specification as used in Ehrendorfer and Bouttier (1998b)
- \textbf{NMC}\textsuperscript{1} B: estimated \( B \) from statistics of differences between 24h and 12h forecasts valid for the same time (NMC method)

The first specification is essentially the vorticity part of the ECMWF background used in 3D–Var and has also been used by Ehrendorfer and Bouttier (1998b). The second specification is obtained from statistics of differences between 24h and 12h forecasts of the QG model valid at the same time (i.e., NMC method, see, Parrish and Derber 1992).

In both specifications, the estimated static B is assumed to correspond to a ”flat background” with homogenous, isotropic correlations. In view of this assumption the B matrix becomes block–diagonal in spectral space. This resembles to some degree the structure of the background error statistics for vorticity in present operational assimilation systems. Series of cycling experiments have shown that these two specifications show similar characteristics in terms of the analyses quality obtained (not shown). Hence all experiments described in the following make use of the NMC B.

6.2.1 Tuning of the static system — Determining of the static B

An adjustment of the static background error covariance matrix B is necessary to stabilize the system in the sense that analysis errors (i.e., the differences \( x^a - x^t \)) only vary between reasonable upper and lower bounds. In other words the system must not divert systematically from the truth–run \( x^t \) as a consequence of the internal error growth of the nonlinear model. This adjustment of the B matrix is implemented as a scaling of the overall level of variance implied through \( B = w^b \ast B^* \) where \( B^* \) is the matrix obtained from the NMC approach described above. By construction, the B matrix depends on the specification of the observational network and associated uncertainties since it reflects the short–term forecast errors computed from the previous analysis. This previous analysis however is the result of a 4D–Var based on the specification of B. Following suggestion by Errico (2002, personal communication) an appropriate B matrix for a given 4D–Var setup (i.e., set of observations) is obtained by evaluating directly the statistics \( x^b - x^t \) for a large number of assimilation

\textsuperscript{1}National Meteorological Center, now, National Centers for Environmental Prediction
cycles. On the basis of these realizations an update to the present \( B \) matrix is computed. Thus an iterative procedure is available for estimating a static \( B \) matrix appropriate for the particular system under consideration. A simplified version of this approach has been used here to compute an estimate for \( w^b \):

Statistics of the differences \( x^b - x^t \) have been computed from a limited set of cycling experiments (<50 cases) using a maximum of 3 iterations for the same set of cycling experiments. On the basis of the overall variance extracted from these statistics the scalar \( w^b \) is estimated. To illustrate the impact of the tuning procedure Fig. 6.2 shows analysis errors in terms of total energy (TE) [J/kg] as a function of assimilation cycle. In an reasonably balanced system one would expect the individual curves to reflect the (daily) changes in analysis quality but not a systematic trend compared to the truth–run \( x^t \). It is evident from Fig. 6.2 that for every configuration the tuning procedure has been able to stabilize the system reasonably well over 48 consecutive assimilation cycles. The four curves in Fig. 6.2 correspond to different choices for the observational network and thus are based on distinctive estimates for \( w^b \).

### 6.3 Dynamic background: The Kalman Filter

On the basis of the static system described in section 6.2 the impact of different dynamic background formulations is studied within the QG/4D–Var system. For this study, these dynamic formulations are RRKF (B) and EKF (C) as illustrated in Fig. 6.1. To address im more detail the cycling of the background error covariance \( P^f \) Fig. 6.3 focuses on the calculation of \( P^f \) for the next assimilation cycle on the basis of the current covariance matrix of analysis errors \( P^a \).

The RRKF algorithm does not make use of \( P^a \) directly. Instead information about analysis errors is contained in the set of HSVs precomputed along eq. (3.2.3). Note that the computation of the HSVs requires specification of the final–time norm \( C \) (see, eq. 3.2.3) as illustrated in Fig. 6.3. On the basis of the HSVs the RRKF algorithm gives an approximation to \( P^f \) constrained on the first leading \( k \) HSVs. The reason for this choice of vectors is related to the optimality property of the HSVs as given in eq. (3.2.5). To avoid rank–deficiency of the resulting (approximated) \( P^f \) the static background error covariance matrix \( B \) is used to fill–up \( P^f \) in the space orthogonal to the dynamic subspace spanned by
6.4 The Reduced–Rank Kalman Filter (RRKF)


The basic idea is that the dynamical propagation of errors is only applied in a subspace of relatively small dimension. This subspace is defined by Hessian singular vectors. This choice seems optimal in view of the properties of HSVs (eq. 3.2.5). Recently, another promising approach for constructing the (most) relevant subspace has been reported by (Farrell and Ioannou 2001a; Farrell and Ioannou 2001b). There the authors make use of the so–called Hankel SVs. Preliminary results using Hankel SVs in an operational RRKF seem encouraging (Fisher 2002, personal communication). In any case the RRKF is designed to give an approximation to the fully evolved forecast error covariance matrix through combining information contained in the set of SVs (HSV in this study) and the static background error covariance matrix B. Note that for the RRKF to be computationally feasible in an operational implementation it is required that the size of this subspace (i.e., the number of HSVs used) is small compared
to the dimension of the model space as illustrated in Fig. 6.4. For details about the implementation of the RRKF reference is made to Haas and Beck (2000).

Following Fisher (1998) we seek an approximation to eq. (3.2.2) and not eq. (3.1.1) thereby neglecting model errors (i.e., the matrix $Q$). We will however come back to the issue of including model error within RRKF algorithm in section 6.4.1. Note that the RRKF algorithm presented here is derived in terms of an approximation to $P^f$ itself and not the inverse of $P^f$ as is in Fisher (1998). For details about the relationship between the "non–inverse" formulation of the RRKF as described here and the derivation of Fisher (1998) reference is made to Haas (2003). The motivation for this "non–inverse" formulation is twofold: First, in view of the propagation of the covariance matrices within the (full) EKF, an approximation to $\hat{P}^f$ itself is obtained from the RRKF as is from the EKF according to eq. (3.2.2). The second motivation is that on the basis of this formulation it is possible to compute an estimate to eq. (3.1.1) and not eq. (3.2.2) thus allowing for representing model errors within the RRKF (see, section 6.4.1).

We start the derivation of the "non–inverse" formulation of the RRKF by defining

$$S_k \equiv C^T X_k \quad \text{and} \quad Z_k \equiv C^{-\frac{1}{2}} X_k \Lambda,$$

(6.4.2)

where $S_k$ and $Z_k$ are $n$-by-$k$ matrices consisting of the first $k$ evolved HSVs (as obtained from eq. 3.2.4) subject to the norm–defining metric $C^T C$ and the scaled by the eigenvalues $\Lambda$, respectively. From eq. (3.2.5) it is obvious that these two sets of vectors obey the following relationship

$$\hat{P}^f S_k = Z_k.$$

(6.4.3)

Note that this equation is of fundamental importance to the operational application of the RRKF (see also, Fisher 1998, eq. 16 therein) since it allows for expressing the action of the forecast error covariance matrix without explicit knowledge of $\hat{P}^f$ by means of the vectors $S_k$ and $Z_k$, respectively.

The static background error covariance matrix $B$ is decomposed as $B = LL^T$ through a Cholesky factorization. To combine the dynamic and the static parts we seek a coordinate rotation such that the $k$-by-$k$ submatrix $E$ reflects the $k$ dimensional subspace associated with the first $k$ HSVs. To construct this coordinate rotation a QR–factorization (see, e.g., Trefethen and Bau 1997) of $L^T S_k$ is computed as:

$$L^T S_k = Q\bar{R},$$

(6.4.4)

where $Q$ is a $n$-by-$n$ orthogonal matrix and $\bar{R}$ is a $n$-by-$k$ matrix. Rearranging the QR–factorization (eq. 6.4.4) we end up with
where $\bar{R}_k$ is a $k$-by-$k$ upper triangular matrix. Next we apply the following transformation to the forecast covariance matrix $\hat{P}^f$ (omitting model errors):

$$\hat{P} = Q^T L^{-1} \hat{P}^f L^{-T} Q,$$

and the corresponding backtransformation is obtained as:

$$\hat{P}^f = L \bar{Q} \hat{P} \bar{Q}^T L^T.$$

The fundamental idea of the RRKF algorithm is to partition $\hat{P}$ into four blocks

$$\hat{P} = \begin{bmatrix} E & F^T \\ F & G \end{bmatrix},$$

where $E$ has dimensions $k$-by-$k$ and corresponds to the (dynamic) subspace of interest. It remains to demonstrate how $E$ is computed on the basis of the HSVs. To relate $E$ to the HSVs, or more precisely, the matrices $S_k$ and $Z_k$ we multiply eq. (6.4.5) from the left with eq. (6.4.8) to obtain

$$\hat{P} Q^T L^T S_k = \begin{bmatrix} E & F^T \\ F & G \end{bmatrix} \begin{bmatrix} \bar{R}_k \\ 0 \end{bmatrix} = \begin{bmatrix} E \\ F \end{bmatrix} \bar{R}_k.$$

Replacing $\hat{P}^f$ in eq. (6.4.3) using eq. (6.4.7) we end up with:

$$L \bar{Q} \hat{P} Q^T L^T S_k = Z_k.$$

or after rearranging

$$\hat{P} Q^T L^T S_k = Q^T L^{-1} Z_k.$$

Combing this equation with eq. (6.4.9) we end up with

$$\begin{bmatrix} E \\ F \end{bmatrix} \bar{R}_k = Q^T L^{-1} Z_k.$$

Since $Z_k$ and $\bar{R}_k$ can be calculated it is possible to obtain $E$ and $F$, which are the first $k$ columns of $\hat{P}$. The computational costs can be controlled by the size of $k$. 

$$Q^T L^T S_k = \begin{bmatrix} \bar{R}_k \\ 0 \end{bmatrix}$$ (6.4.5)
Note that the partitioning in eq. (6.4.8) also requires specification of the block $G$. Since the transformed $B$ (denoted as $\hat{B}$) is the identity (substitute for $\hat{P}^f = B$ in eq. 6.4.6)

$$\hat{B} = Q^T L^{-1} B L^{-T} \hat{Q} = Q^T I Q = I,$$

(6.4.13)
where the first identity follows from the Cholesky decomposition of $B$ and the latter is a consequence of $\hat{Q}$ being an orthogonal matrix. In view of eq. (6.4.13) it makes sense to replace the block $G$ in eq. (6.4.8) by the identity matrix thus combining the dynamic information (given in $E$) and the static information (given by the identity). This substitution however might cause the resulting matrix $\hat{P}_k$ to be not positive definite any more. To overcome this problem, we apply a scaling factor $\alpha$ to the cross-covariances $F$ as

$$\hat{P}_k = \begin{bmatrix} E \sqrt{\alpha F^T} \\ \sqrt{\alpha} F \end{bmatrix},$$

(6.4.14)
and choosing the largest $\sqrt{\alpha}$ such that $E - \alpha F^T F$ is positive definite (see, Haas et al. 2000). Note that the most simple way to ensure $\hat{P}_k$ being positive-definite is to set $F$ to zero. To obtain the approximation $\hat{P}^f_k$ of $\hat{P}^f$ we transform $\hat{P}_k$ back using eq. (6.4.7):

$$\hat{P}^f_k = L \hat{Q} \hat{P}_k \hat{Q}^T L^T$$

(6.4.15)

$\hat{P}^f_k$ is the approximated forecast error covariance matrix (omitting model errors) estimated on the basis of the $k$ leading HSVs. It remains to evaluate eq. (6.4.15) for the two extreme cases $k = n$ and $k = 0$. Setting $k = n$, or in other words, using the full set of HSVs corresponds to a "perfect approximation". From eq. (6.4.14) it is obvious that $\hat{P}_n = \hat{P}$ and thus eq. (6.4.15) becomes for $k = n$

$$\hat{P}^f_n = L \hat{Q} \hat{P}_n \hat{Q}^T L^T.$$

(6.4.16)

From eq. (6.4.11) the underbraced term is written as

$$\hat{P}^f_n = L \hat{Q} \hat{Q}^T L^{-1} Z_k S_k^{-1} = Z_k S_k^{-1},$$

(6.4.17)
where the last identity follows from the orthogonality of $\hat{Q}$. Comparing eq. (6.4.17) to eq. (6.4.3) it is obvious that $\hat{P}^f_n = \hat{P}^f$ and thus the full forecast error covariance matrix is obtained.

In the case $k = 0$ the partitioning along eq. (6.4.8) gives $\hat{P} = I$ and thus from eq. (6.4.15) it is obvious that the static $B$ matrix is obtained:

$$\hat{P}^f_0 = L \hat{Q} I \hat{Q}^T L^T = B,$$

(6.4.18)

which follows from the Cholesky decomposition of the static $B$ matrix: $B = LL^T$. 

6.4 The Reduced–Rank Kalman Filter (RRKF)
6.4 The Reduced–Rank Kalman Filter (RRKF)

6.4.1 The RRKF with model error

In view of eq. (3.1.1) the derivation of the RRKF given in the previous section is suboptimal since it does not allow for including model error. It is however possible to include the impact of model error in the RRKF algorithm as shown below.

To that aim we seek an expression for the model error covariance matrix $Q$ applicable in the transformed (i.e., by $L$ and $\hat{Q}$) subspace of the RRKF. This transformed part of the model error covariance matrix $Q$ relevant within the subspace is then added to the dynamic part of $\hat{P}_k$ (i.e., the $E$ block). By including the effect of model error we are able to calculate an approximation to (the full) $P^f$ which we denote by $P^f_k$, with $k$ again denoting the size of the subspace (i.e., the number of HSVs). In view of the model error covariance $Q$, eq. (6.4.15) is now written as

$$P^f_k = L\hat{Q}\left(\hat{P}_k + U_k \otimes \left(\hat{Q}^T L^{-1} Q L^{-T} Q\right)\right) L^T,$$  \hspace{1cm} (6.4.19)

where $\otimes$ is the Hadamard product and

$$U_k \equiv \begin{bmatrix} U' & 0 \\ 0 & 0 \end{bmatrix}$$ \hspace{1cm} (6.4.20)

with $U'$ a $k$-by-$k$ matrix consisting of ones. The underbraced term in eq. (6.4.19) is the transformed $Q$ matrix. The matrix $U_k$ (acting on the transformed $Q$ via the Hadamard product) ensures that the model error effects the dynamic part of $\hat{P}_k$ (i.e., the $E$ block in eq. 6.4.14) only.

Let us now look at eq. (6.4.19) for the two extreme cases $k = n$ and $k = 0$. Recalling from eq. (6.4.17) that the case $k = n$ corresponds to the full EKF (omitting model errors) we now end up with

$$P^f_n = L\hat{Q}(\hat{P}_n + \hat{Q}^T L^{-1} Q L^{-T} Q) L^T = L\hat{Q} L^T + Q$$ \hspace{1cm} (6.4.21)

Comparing eq. (6.4.21) to the expression for the full EKF (eq. 3.1.1) we see that for $k = n$ the algorithm is equal to the full EKF. For the static case $k = 0$ (see, eq. 6.4.18) we end up with

$$P^f_0 = L\hat{Q}(I + 0)\hat{Q}^T L^T = LL^T = B$$ \hspace{1cm} (6.4.22)

From eq. (6.4.21) and eq. (6.4.21) it is obvious that the model error formulation eq. (6.4.19) is consistent with respect to both extreme cases (i.e., EKF and static system).
6.4.2 Scaling of the background

The blending of dynamic and static information along eq. (6.4.14) requires that the overall level of variance implied through HSVs and the static B are of comparable magnitude. Especially in the case of too large background variances the dynamic information as contained in the HSVs might become suppressed and thus not used efficiently. As a means for adjusting the possible misfit between background error variance and the variance implied through the HSVs a scaling factor \( \beta \) is introduced. Let us consider the scaling \( \beta \) acting on the static B matrix as follows:

\[
B' = \beta B \quad \rightarrow \quad L' = \beta^{1/2}L
\]

(6.4.23)

where \( \prime \) denotes the scaled quantities (i.e., B and corresponding L). To investigate the impact of this scaling factor we compare the equations involving \( L' \) (or \( L \), respectively) and express the \( \prime \) quantities in terms of original expressions. In view of eq. (6.4.4) it follows that:

\[
\bar{R}' = \beta^{1/2}R
\]

and thus using eq. (6.4.12) we find that:

\[
\begin{bmatrix}
E' \\
F'
\end{bmatrix} = \beta^{-1} \begin{bmatrix}
E \\
F
\end{bmatrix}
\]

(6.4.24)

and finally we get:

\[
\hat{P}_k = \begin{bmatrix}
\beta^{-1}E & \beta^{-1}\sqrt{\alpha F^T} \\
\beta^{-1}\sqrt{\alpha F} & I
\end{bmatrix} = \beta^{-1} \begin{bmatrix}
E & \sqrt{\alpha F^T} \\
\sqrt{\alpha F} & \beta I
\end{bmatrix}
\]

(6.4.25)

From eq. (6.4.25) it is obvious that the scaling has a nontrivial impact on the final result \( (\hat{P}_k) \) after backtransformation according to eq. (6.4.15). Note that a scaling of the HSVs and thus \( S_k \) and \( Z_k \) does not yield an impact of the final result. This property is evident from equations (6.4.12) and (6.4.9), respectively.
Chapter 7

4D–Var experiments and results

Referring back to the motivation given in sec. 1.4 and the illustration of the QG/4D–Var system (Fig. 6.1) we cast the primary objective of this work into to two main questions: The first being related to the potential benefit of a fully dynamic background (i.e., the EKF) compared to the static system and the second being how much of this benefit might be obtained from the RRKF. These questions are addressed through series of cycling experiments within the QG/4D-Var system described in chapter 5.

7.1 Summary of experiments

The experimental setup is such that the changes among the experiments described concern mainly the specification of the background errors. Specifically, all experiments share the same set of observations and use 12h assimilation windows.

<table>
<thead>
<tr>
<th>Specifications</th>
<th>section</th>
</tr>
</thead>
<tbody>
<tr>
<td>perfect model scenario</td>
<td>5.3, 7.1.1</td>
</tr>
<tr>
<td>12–hour assimilation window</td>
<td>5.2, 7.2</td>
</tr>
<tr>
<td>12–hour optimization time for HSVs</td>
<td>3.2, 7.1.2</td>
</tr>
<tr>
<td>48 consecutive assimilation cycles</td>
<td>7.2</td>
</tr>
<tr>
<td>pre–specified observational network</td>
<td>5.3</td>
</tr>
<tr>
<td>”numerical” model error term in EKF and RRKF</td>
<td>6.3, 6.4, 7.1.1</td>
</tr>
<tr>
<td>cross–covariances omitted in RRKF</td>
<td>6.4</td>
</tr>
</tbody>
</table>

Table 7.1: Common parameters and specifications in 4D–Var experiments.

Table 7.1 summarizes the parameters and specifications that are common to all experiments.
described in sections 7.3 to 7.5. This table collects individual specifications described in previous sections as well as the experimental setup that has been chosen. For easy reference, the corresponding section is given in the second column of table 7.1. The following sections summarize the main features of the experiments that are discussed in sections 7.3 to 7.5.

7.1.1 Static system and EKF

Table 7.2 lists the specifications for experiments with the static system and the EKF (configurations A and C in Fig. 6.1). These specifications include the experiment identifier, the choice for the static B matrix, the sample size for the nonlinear covariance propagation (NSVMC; see, 7.5.1), and a short description of the experiment. For details refer to the individual sections given in the last column.

<table>
<thead>
<tr>
<th>Exp ID</th>
<th>static B matrix</th>
<th>NSVMC</th>
<th>description</th>
<th>sec.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Xat</td>
<td>NMC B</td>
<td></td>
<td>static system</td>
<td>7.3</td>
</tr>
<tr>
<td>Xbx</td>
<td></td>
<td></td>
<td>fully dynamic system (EKF)</td>
<td></td>
</tr>
<tr>
<td>Xbc</td>
<td>”best” B (Xbx)</td>
<td></td>
<td>”best static” system</td>
<td></td>
</tr>
<tr>
<td>Xbv</td>
<td></td>
<td>5000</td>
<td>NL EKF</td>
<td>7.5.1</td>
</tr>
<tr>
<td>Xbs</td>
<td></td>
<td>20000</td>
<td>NL EKF</td>
<td></td>
</tr>
</tbody>
</table>

Table 7.2: Summary of specifications for static and EKF experiments.

Although this work is carried out in a perfect model scenario it is necessary to include a (very) small numerical model error term in both EKF and RRKF. This model error term is necessary for numerical stability. In view of eq. (3.1.1) the specification of the numerical model error covariance matrix is such that $Q$ is diagonal and $Q \ll MPaM^T$, i.e., the model error variance is at least two magnitudes smaller than the corresponding (dynamic) forecast error variance. The diagonal elements of $Q$ are estimated from differences between QG forecasts and corresponding ECMWF forecasts truncated at T21. These differences are taken as ”guesses” for the model error of the QG system. The statistics (i.e., variances) are estimated from $\sim 2000$ differences between the QG and the ECMWF system. The resulting variances are subsequently scaled to reflect the prechosen overall level of variance. We stress that this ”model error” term is necessary for numerical reasons only and does not contradict the perfect model assumption.
7.2 Cycling experiments

7.1.2 Background formulation and RRKF

Table 7.3 lists specifications for the experiments involving the RRKF (configuration B in Fig. 6.1). For these experiments the most important parameter is the number of HSVs used (denoted as $k$) and the choice of the scaling parameter $\beta$. Shown are the experiment identifier, the number of HSVs $k$, the choice of $\beta$, and a short description of the experiment. For details we refer again to the sections given in the last column.

<table>
<thead>
<tr>
<th>Exp ID</th>
<th>$k$</th>
<th>$\beta$</th>
<th>description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Xbd</td>
<td>$k=10$</td>
<td>1.0</td>
<td>RRKF using 10 HSVs</td>
</tr>
<tr>
<td>Xbj</td>
<td>$k=100$</td>
<td>1.0</td>
<td>RRKF using 100 HSVs</td>
</tr>
<tr>
<td>Xbo</td>
<td>$k=200$</td>
<td>1.0</td>
<td>RRKF using 200 HSVs</td>
</tr>
<tr>
<td>Xbm</td>
<td>$k=1000$</td>
<td>1.0</td>
<td>RRKF using 1000 HSVs</td>
</tr>
<tr>
<td>Xbk</td>
<td>$k=100$</td>
<td>0.5</td>
<td>scaling, $\beta = 0.5$</td>
</tr>
<tr>
<td>Xbl</td>
<td>$k=100$</td>
<td>2.0</td>
<td>scaling, $\beta = 2.0$</td>
</tr>
<tr>
<td>Xbn</td>
<td>$k=100$</td>
<td>5.0</td>
<td>scaling, $\beta = 5.0$</td>
</tr>
<tr>
<td>Xbf</td>
<td>$k=10$</td>
<td>2.0</td>
<td>scaling, $\beta = 2.0$</td>
</tr>
<tr>
<td>Xbg</td>
<td>$k=10$</td>
<td>5.0</td>
<td>scaling, $\beta = 5.0$</td>
</tr>
</tbody>
</table>

Table 7.3: Summary of specifications for RRKF experiments.

The optimization time $t_{opt}$ for the HSVs required in the definition of the RRKF algorithm is equal to the length of the assimilation window for all experiments: $t_{opt} = 12$ h. For consistency with the EKF the numerical model error term introduced in section 7.1.1 is also used within the RRKF formulation along the lines outlined in section 6.4.1.

7.2 Cycling experiments

The cycling process is started from a static background $B$ and a background field $x^b$ consistent with this $B$ specification: $x^b \sim \mathcal{N}(x^i; B)$, i.e., the initial background field is a realization of a normal distribution with mean $x^i$ and covariance $B$. After the first assimilation cycle the (new) background statistics is obtained from either the RRKF or the EKF. For the static system no update of the background error covariance matrix takes place (see, Fig. 6.1). Hence, the results shown below mainly address questions regarding the performance of the RRKF (in different configurations) in comparison to the static system and the full EKF. All results are taken from series of cycling experiments for 12h assimilation window using roughly 330 observations at times 00/06/11. This configuration corresponds
to the (spatial) observation distribution as shown in Fig. 5.3 (for one particular time–slot). Again we note that the spatial distribution varies in time (i.e., is different for the individual time–slots) but the observation error statistics is kept constant in time (see, sec. 5.3).

The results presented are taken from series of 48 consecutive assimilation cycles where the first 12 assimilation cycles are not taken into account for computing mean errors, skill scores, and RMS (root–mean–square) errors. After this initial period of 12 cycles it is assumed that the system has adjusted to the configuration selected (i.e., RRKF or EKF). Note that it is for technical reasons only that the cycling process is initiated with a static assimilation run independent of the configuration chosen.

### System performance: ana/fc errors (36 cases)

<table>
<thead>
<tr>
<th>ExpID</th>
<th>analysis error</th>
<th>forecast error: TE / RMS NH / RMS SH</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TE [J/kg]</td>
<td>RMS Z500 [m] NH&gt;30 SH&gt;30 t=24 t=48</td>
</tr>
<tr>
<td>Xat</td>
<td>64.0 21.2 53.2</td>
<td>88.1 / 37.1 / 61.7 / 148.0 / 82.6 / 77.3</td>
</tr>
<tr>
<td>Xbx</td>
<td>19.7  8.8  26.8</td>
<td>32.2 / 18.3 / 37.1 / 63.8 / 43.4 / 53.3</td>
</tr>
<tr>
<td>Xbc</td>
<td>31.4 12.8 33.6</td>
<td>43.0 / 23.1 / 41.4 / 88.1 / 55.8 / 55.0</td>
</tr>
<tr>
<td>Xbv</td>
<td>18.4  9.4  35.6</td>
<td>29.3 / 17.9 / 45.1 / 56.9 / 40.0 / 58.1</td>
</tr>
<tr>
<td>Xbs†</td>
<td>6.9   7.6  17.0</td>
<td>11.5 / 13.8 / 22.7 / 23.8 / 29.0 / 33.1</td>
</tr>
<tr>
<td>Xbd</td>
<td>68.8 18.7 57.0</td>
<td>91.4 / 34.5 / 66.5 / 149.5 / 78.2 / 82.0</td>
</tr>
<tr>
<td>Xbj</td>
<td>59.5 16.6 49.4</td>
<td>81.0 / 29.7 / 58.6 / 135.4 / 68.4 / 76.9</td>
</tr>
<tr>
<td>Xbo</td>
<td>44.6 15.3 38.7</td>
<td>62.7 / 26.3 / 49.2 / 108.0 / 58.7 / 68.3</td>
</tr>
<tr>
<td>Xbm</td>
<td>35.9 13.4 31.3</td>
<td>55.3 / 23.9 / 40.7 / 104.0 / 55.3 / 58.0</td>
</tr>
<tr>
<td>Xbk</td>
<td>71.4 18.4 53.2</td>
<td>96.7 / 32.9 / 64.0 / 158.4 / 77.5 / 79.6</td>
</tr>
<tr>
<td>Xbl</td>
<td>53.5 17.4 40.9</td>
<td>73.2 / 28.9 / 49.2 / 124.6 / 62.8 / 67.7</td>
</tr>
<tr>
<td>Xbn</td>
<td>62.8 19.6 46.7</td>
<td>84.0 / 31.0 / 52.9 / 139.0 / 68.7 / 70.6</td>
</tr>
<tr>
<td>Xbf</td>
<td>55.8 18.7 45.4</td>
<td>75.9 / 30.0 / 55.9 / 129.8 / 68.7 / 74.6</td>
</tr>
<tr>
<td>Xbg</td>
<td>64.1 21.1 47.0</td>
<td>86.7 / 33.7 / 55.3 / 147.6 / 75.9 / 74.6</td>
</tr>
</tbody>
</table>

Table 7.4: Summary of 4D–Var results: analysis and forecast errors.

Table 7.4 summarizes the performance in terms of analysis and forecast errors of all cycling experiments discussed. Every row in table 7.4 corresponds to one cycling experiment

---

†30 cases only for experiment Xbs
where the errors are obtained as the mean over 36 consecutive assimilation cycles (omitting cycles #1 – #12). For the characteristics of the individual experiments reference is made to tables 7.2 and 7.3, respectively. Table 7.4 lists (mean) analysis and forecast errors (day−1 (t=24) and day−2 (t=48)) in terms of total energy [J/kg] and RMS Z500 [m] for the northern and southern hemisphere. The (RMS) statistics is evaluated for gridpoints >30N (NH) and <30S (SH), respectively. The experiments are discussed in more detail in sections 7.3 to 7.5.

### 7.3 Static system vs. EKF

To study the potential benefit of a dynamic background formulation in the QG/4D–Var we compare the performance of the static system to the full EKF. The EKF configuration is regarded as fully dynamic since it explicitly propagates in time the current estimate of analysis errors \( P_a \) and thus provides the next assimilation cycle with the appropriate short–term forecast error statistics as contained in \( P_f \). In the static system however the background field \( x^b \) only is propagated in time by means of the nonlinear model. The background statistics \( B \) however is kept constant. Hence in every assimilation cycle the same (static) \( B \) matrix is used within the 4D–Var algorithm. This configuration to some degree mimics the current operational practise.

Fig. 7.1 shows analysis and forecast errors in terms of total energy [J/kg] as a function of lead time [h]. These errors are computed with respect to the truth run \( x^t \) and thus represent the true analysis and forecast errors, respectively. The three curves shown in Fig. 7.1 correspond to the static system (Xat; black), the EKF (Xbx; orange), and the so–called "best static" system (Xbc; violet). The curves are obtained as the mean over 36 consecutive assimilation intervals (skipping the first 12 intervals) as discussed in section 7.2. Part of the information shown in Fig. 7.1 can also be seen from table 7.4. It is obvious from Fig. 7.1 that the EKF is able to reduce the analysis error by a factor of three compared to the static system. Consequently, the forecast started from this improved analysis has the smallest error although the error growth rate is somewhat larger compared to the static system.

The violet curve in Fig. 7.1 corresponds to a static system where the \( B \) matrix has been obtained as the mean over the 48 realizations of \( P_f \) obtained from the EKF. In other words, this \( B \) matrix represents the mean forecast error structure as computed from the 48 consecutive EKF applications for the same period of time. This configuration is denoted as "best static" since it is optimal in the sense that it is derived from the \( P_f \) matrices obtained from the EKF for same sequence of analysis cycles. Note that this \( B \) matrix is indeed a full matrix and is not block–diagonal as the NMC \( B \) described in section 6.2. It is seen from Fig. 7.1 that the "best static" system gets reasonably close to the EKF in terms of

\[^1\text{30 cases only for experiment Xbs}\]
7.3 Static system vs. EKF

Figure 7.1: Analysis and forecast errors for static system (Xat; black), "best static" system (Xbc; violet), and EKF (Xbx; orange) in terms of total energy [J/kg] as a function of forecast time [h].

analysis and forecast errors. Nevertheless the EKF performs significantly better than the "best static" system which means that the day-to-day variability of the error structure does have a significant impact on the performance of the 4D-Var system.

To illustrate the positive impact of the EKF on the quality of the analyses and subsequent forecasts Fig. 7.2 shows difference in RMS analysis errors obtained from 4D-Var, between static system and the EKF, in terms of 500hPa geopotential height [m]. Note that as in Fig. 7.1 the RMS values are computed on the basis of 36 consecutive assimilation cycles (skipping cycles #1 to #12). The contour interval is given in the caption with shading indicating areas where the EKF has smaller RMS analysis errors than the corresponding static system. Referring back to the observational network shown in Fig. 5.3 it is evident that the EKF improves the analysis over data sparse areas. By construction these data sparse areas are the Pacific as well as the other southern hemisphere Oceans. There is also a clear indication of improvement over the Atlantic but given the spatial resolution of the QG model (T21) the Atlantic ocean is fairly limited in terms of the number of model grid-points. In data sparse areas the analysis increments are dominated by the structure implied through the background error covariance matrix $P^b$. It is evident from Fig. 7.2 that the dynamic covariance structure as obtained from the EKF is superior to that of the static B.

Fig. 7.3 shows differences in RMS errors of the background field between EKF and
Figure 7.2: Differences between 4D–Var RMS analysis errors [m] for fully dynamic background (EKF) and static at 500 hPa, averaged over 36 cases. The shaded areas indicate negative values, where EKF has smaller RMS errors than the static system. Contours: +/- 10/20/40/80/160 m (solid red, positive values; dashed blue, negative with areas shaded).

Figure 7.3: as Fig. 7.2, but RMS background errors.
7.3 Static system vs. EKF

Comparing differences in background errors (Fig. 7.3) to differences in analysis errors (Fig. 7.2) suggests that the improvement in analysis quality does not result from a consistently more accurate background field only but from the better use of the observational information and the more sophisticated background error covariance matrix available through the EKF.

To demonstrate the impact of improving the analysis as achieved by the EKF, Figs. 7.4 to 7.7 show differences in RMS forecast errors for 500hPa geopotential height between the static system and the EKF at forecast days 1, 2, 3, and 4. As in Fig. 7.2, the shading indicates regions where the EKF has smaller RMS (forecast) errors and thus more accurate forecasts compared to the static system. Focusing on the northern hemisphere it is evident that the improved analysis over the northern Pacific as well as the Atlantic ocean leads to a significant reduction of forecast errors throughout the forecast period. The regions where the forecast is improved have amplified rapidly during the first 3 or 4 days and the error pattern has spread across the whole northern hemisphere at days 3 and 4 as a result of the dynamics of the model. The picture for the southern hemisphere is different in the sense that although the EKF results in more accurate forecasts in general, there are areas where the forecast from the static system has smaller errors.

Fig. 7.8 shows differences of RMS analysis increments of 500hPa geopotential height [m] between the EKF configuration and the static system. Clearly, the most significant differences occur in data sparse regions (i.e., over the Oceans, especially in the southern hemisphere). However, the size of the increments is fairly similar in both con-
Figure 7.5: as Fig. 7.2, but RMS errors for day–2 forecasts.

Figure 7.6: as Fig. 7.2, but RMS errors for day–3 forecasts.
7.3 Static system vs. EKF

Figure 7.7: as Fig. 7.2, but RMS errors for day–4 forecasts.

Figure 7.8: Differences between 4D–Var RMS increments [m] for fully dynamic background (EKF) and static B at 500 hPa, averaged over 36 cases. The shaded areas indicate negative values, where EKF has smaller increments than the static system. Contours: +/- 2/4/8/16/32 m (solid red, positive values; dashed blue, negative with areas shaded).
7.3 Static system vs. EKF

figurations. Referring back to Fig. 7.3, showing improved background fields in case of the EKF, the overall size of the increments is again an indication that the improvement of the system using the EKF is not due to the fact that the background field is more accurate compared to the static system but is a direct consequence of the flow–dependent and thus more realistic covariance structure of $P_f$. In other words, although the EKF analyses make use of a more accurate background, the size of the increments is fairly similar compared to the static system thus highlighting the influence of the background formulation.

Comparing analysis increments for analysis cycle #2 for the static configuration (Xat) and the EKF (Xbx) reveals the systematic differences implied through the different background specifications. Cycle #2 is the first cycle where the system makes use of the flow–dependent background provided through the EKF. As a consequence of the experimental setup (see, section 7.2), this is the only difference between these two experiments since both share the same analysis $x^a$ and background field $x^b$ obtained from cycle #1 (i.e., the initial static run). Figs. 7.9 and 7.10 show analysis increments in terms of geopotential height [m] for the three model levels (200, 500, and 800hPa) for the static system (Fig. 7.9) and the EKF (Fig. 7.10). The contour interval is 25m in both figures. The precise location of the static increments is strongly dependent on the location of the observations showing substantially smaller increments over data void areas. As discussed previously the main differences occur over the Oceans and thus the data sparse areas. Focusing on the Pacific ocean the increments obtained from the EKF have a more pronounced structure and are more small scale. Although the differences between the two experiments might seem to be small we stress that they are solely due to the different background error covariance matrix applied.

The configuration denoted "best static" system (Xbc) allows for investigating the role of the day–to–day changes of the background error structures since it is computed as the mean over the 48 consecutive realizations of $P_f$ (obtained from the EKF). Thus the comparison of the performance of this system to the system based on the static B (Xat) gives an indication for the (theoretical) improvement that could be achieved on the basis of a static 4D–Var system. Fig. 7.11 shows RMS differences for 500hPa geopotential height between configurations Xat (static system) and Xbc ("best static" system). As in previous figures, the shading indicates areas where the latter (i.e., the "best static" system) has smaller analysis errors. Comparing Fig. 7.11 to the corresponding figure for the EKF (Fig. 7.2) shows a strong agreement among the two figures. The spatial location of the regions where the analysis is improved is very similar but the difference field looks somewhat smoother in Fig. 7.11. This is due to the fact that the background error covariance used for the "best static" system is computed as the mean of the EKF $P_f$ matrices and thus, by construction, possesses less variability. In terms of analysis increments the picture is qualitatively similar however the agreement between the "best static" system and the EKF is weaker. Fig. 7.12 shows the RMS difference for analysis increments between the "best static" (Xbc) and
Figure 7.9: Geopotential height analysis increments [m] for static system (Xat) at three model levels (200, 500, and 800 hPa). Contour interval is 25m, with negative values dashed blue.
7.3 Static system vs. EKF

Figure 7.10: as Fig. 7.9, but for EKF.
7.3 Static system vs. EKF

Figure 7.11: as Fig. 7.2, but for "best static" (Xbc) and static (Xat) B. The shaded areas indicate negative values, where configuration Xbc has smaller RMS errors than the static system. Contours: +/- 10/20/40/80/160 m (solid red, positive values; dashed blue, negative with areas shaded).

Figure 7.12: as Fig. 7.8, but for "best static" (Xbc) and static (Xat) B. The shaded areas indicate negative values, where "best static" system has smaller increments than the static system. Contours: +/- 2/4/8/16/32 m (solid red, positive values; dashed blue, negative with areas shaded).
7.3 Static system vs. EKF

the (ordinary) static system (Xat). This figure corresponds to Fig. 7.8 where differences between the EKF and the static system are shown.

From Fig. 7.1 (and also table 7.4) it is clearly seen that the one the basis of a purely static system the quality of the analyses could in principle be improved substantially (i.e., from 64 to 31 J/kg in terms of TE). However, this improvement would only be possible if a highly optimized B matrix, such as in the "best static" system, were available. Nevertheless the EKF and thus the daily variations of the atmospheric predictability have to be taken into account to achieve the largest benefit (i.e., reduction of analysis errors to below 20 J/kg in terms of TE). The (good) performance of the "best static" system is also related to the specification of the observational network which is assumed constant among all assimilation cycles (see, section 5.3). Thus, by construction, the overall structure of the Pf matrices used for computing the "best static" B matrix is similar. For a time–varying observational network, one would expect the "best static" system performing worse compared to the setup chosen here.

Looking at the covariance matrices directly allows for investigating further the differences between the static configurations and the EKF. Fig. 7.13 shows geopotential height forecast error variances at 500hPa for (a) the static system (Xat), (b) the "best static" system (Xbc), and (c) the EKF (Xbx) for one particular assimilation cycle (cycle #25) chosen randomly. For the static configurations (top and middle panel) the variance field shown is independent of the assimilation cycle by construction. It is clearly seen from Fig. 7.13a that the static B matrix is a very crude approximation to the actual variances obtained from the EKF (Fig. 7.13c). In particular, the static B matrix does not give any indication for (spatial) regions where large errors are commonly observed (i.e., storm tracks over the Atlantic and Pacific Ocean). From Fig. 7.13c it is evident that the largest forecast error variances are located in the southern hemisphere and in particular in the South Pacific. This is a direct consequence of the choice of the spatial observation distribution within this study. In view of this dependence of the variances on the particular observational network it is evident that a reasonable static B matrix has to reflect both the error dynamics of the underlying model and the observational distribution. Clearly, the latter dependency has not been taken into account by the specification of the (static) NMC B matrix used here (see, section 6.2).

The forecast error variance obtained from the EKF (panel c) reflects the impact of the previous assimilation and the error growth over the 12h interval. The B matrix used in the "best static" system (Fig. 7.13b) shows a much more realistic picture of the forecast error variance. In particular, the spatially varying error structure is taken care of and thus the main features of Fig. 7.13c are present in Fig. 7.13b. Again this is a direct consequence of the construction of this B matrix. Or put differently, the variance structure shown in Fig. 7.13c is one of the 48 realizations used for computing Fig. 7.13b.

Fig. 7.14 shows correlations at 800hPa of one particular gridpoint located in the mid–Atlantic at 500 hPa with all other gridpoints for the same configurations as Fig. 7.13.
Figure 7.13: Geopotential height variances [m$^2$] (i.e., diagonal elements of $P^f$) for (a) the static B, (b) the "best static" B, and (c) the EKF (cycle #25). Contour interval is given between panels b and c (shaded regions correspond to variances > 200m$^2$). Same contours and shading used in all panels.
Figure 7.14: Correlation structures at 800 hPa implied by $P^f$ of one selected gridpoint (at 500 hPa) with all other gridpoints on the globe for (a) the static $B$, (b) the "best static" $B$, and (c) the EKF (cycle #25). Contour interval is 0.2 (negative values dashed blue).
The static B matrix is block-diagonal by construction and therefore shows an (exactly) isotropic correlation structure (Fig. 7.14a). On the other hand, the correlation obtained from the EKF (Fig. 7.14c) shows a dipole structure with a strong negative correlation upstream. This correlation pattern is to some degree present in Fig. 7.14b. However, in view of Fig. 7.13 the results suggest that a reasonable approximation of the off-diagonal elements of $P_f$ by means of a static B matrix is more difficult than the specification of the variances. This point might be even more important in the case when the model state vector consists of different physical variables (i.e., temperature and height) and thus correlations between these variables have to be taken into account.
This section describes experiments involving the RRKF in different configurations. As in the previous section, an overview of the experiments carried out is available from table 7.3 and selected results about the performance of the system are given in table 7.4. The performance of the RRKF is compared to the static and the EKF system. From theoretical considerations it is expected that the static B configuration forms the lower and the full EKF the upper bound for the performance of the RRKF (see, sec. 6.4).

![Figure 7.15: Analysis and forecast errors for RRKF in terms of total energy [J/kg] as a function of forecast time [h]: k = 10 (Xbd; red), k = 100 (Xbj; light green), k = 200 (Xbo; dark green), and k = 1000 (Xbm; blue). Also shown are the corresponding curves for the static system (Xat; black) and the EKF (Xbx; orange) as shown in Fig. 7.1 (mean over 36 cases).](image)

First, the performance of the RRKF is summarized in Fig. 7.15 showing analysis and forecast errors in terms of TE [J/kg] for RRKF configurations with different numbers \( k \) of HSVs. The configurations shown in Fig. 7.15 are: \( k = 10 \) (Xbd; red), \( k = 100 \) (Xbj; light green), \( k = 200 \) (Xbo; dark green), and \( k = 1000 \) (Xbm; blue). In addition, the corresponding curves for the static system and the EKF (see, Fig. 7.1) are included for comparison. It is evident from Fig. 7.15 that the RRKF shows a nonmonotonic behavior with respect to the number of HSVs \( (k) \) in terms of analysis errors: The RRKF system using \( k = 10 \) vectors (red curve) has larger analysis errors than the static system (black curve), although the difference is fairly small. Using \( k = 100 \) vectors (light green curve) reduces
7.4 Role of the background formulation and RRKF

analysis (and forecast) errors compared to the static system by roughly 10%. Increasing further the number of HSVs the quality of the analysis improves substantially (i.e., $k = 200$). A very large number of HSVs (more than 1000) is necessary to have an RRKF performing almost as good as the EKF.

![Graph showing the comparison between static system and RRKF for analysis and forecast errors](image)

Figure 7.16: as Fig. 7.15, but in terms of RMS Z500 [m] for gridpoints (a) $>30$N (northern extratropics) and (b) $<30$S (southern extratropics).

Fig. 7.16 again shows the performance of the RRKF but in terms of RMS errors for geopotential height at 500hPa evaluated for (a) the northern hemisphere extratropics and (b) the southern hemisphere extratropics. Compared to Fig. 7.15 the overall conclusions are similar although in Fig. 7.16a the behavior of the RRKF is monotonic with respect to increasing $k$. The striking feature of Fig. 7.16 is the difference in analysis errors and the subsequent error growth between the two hemispheres. Again, the difference in the quality of the analysis (between NH and SH) is a consequence of the observational network chosen. The error growth evident from Fig. 7.16b suggests that for analysis errors of about 50 meters (in an RMS sense) the error growth is strongly influenced by the nonlinear model dynamics resulting in a very small error growth rate compared to Fig. 7.16a: The errors in Fig. 7.16b grow by a factor of 1.5 over 48 hours (in an RMS sense) whereas for Fig. 7.16a a factor of more than 4 is obtained over the same time interval (see, table 7.4 for the individual errors).

Fig. 7.17 shows analysis increments at the three model levels in terms of geopotential height [m] for RRKF using $k = 100$ vectors $(X_{bj})$ compared to the static system and the EKF. As in Figs. 7.9 and 7.10 one particular assimilation cycle (here, cycle #36) has been chosen randomly. The contour interval is 25m. It is evident from Fig. 7.17 that both the static system and the RRKF (using $k = 100$ HSVs) have larger increments compared to the EKF (panel c). Generally speaking there seems to be only a small agreement among the three configurations. However, closer inspection reveals some important features. For example, let us focus on the increment structure south east of the African continent between 30° and 60° eastern longitude. The structure in Fig. 7.17a (i.e., for the static configuration)
Figure 7.17: Analysis increments in terms of geopotential height [m] for (a) the static system (Xat), (b) RRKF using $k = 100$ HSVs (Xbj), and (c) the EKF (Xbx) for one particular assimilation cycle (cycle #36). Contour interval is 25m, negative values dashed blue.
is essentially barotropic without any significant tilt with height. In comparison, the RRKF (and the EKF as well) shows a westward tilt with height of the increment structures. Due to the poor vertical discretization of the QG model (having three levels only) this baroclinic features do not show up very strongly.

Looking at differences in RMS Z500 between the static system and the RRKF ($k = 100$) a picture similar to the comparison with the EKF is obtained. Fig. 7.18 shows differences between 4D–Var RMS analysis increments [m] for the static configuration (Xat) and the RRKF using $k = 100$ vectors (Xbj). As in Fig. 7.8, the largest differences occur in data sparse areas over the Oceans and the southern hemisphere. This property suggests that the RRKF is indeed capable of approximating the dominant structures of $P_f$. Focusing on the northern Pacific and comparing Fig. 7.18 to Fig. 7.8 a good agreement is found among RRKF and EKF, respectively.

Figs. 7.19 and 7.20 show analysis and forecast error variances (i.e., diagonal elements of $P^a$ and $P^f$, respectively) at 500hPa for (a) the static system (Xat), the RRKF using $k = 100$ HSVs (Xbj), and (c) $k = 1000$ HSVs (Xbm). Also shown are corresponding (true) analysis errors ($x^a - x^t$) as well as the background field ($x^b$ for analysis cycle #25). Same contour and shading conventions are used in all panels. The spatial distribution of analysis error variance as shown in Fig. 7.19 reflects the impact of the observations (lower variance over the continents). By increasing the number of HSVs from 0 to 100 and 1000 (top to bottom panel) the overall magnitude of variance is reduced and the spatial structures become more prominent. This is a direct consequence of the more sophisticated specification of the background error covariances within the assimilation cycle. The agreement between the variances and the "true" errors is good for the configurations involving the RRKF (pan-
Figure 7.19: Analysis error variances \([m^2]\) (implied by \(P^a\)) and corresponding errors \((x^a - x^l)\) [m] in terms of geopotential height at 500hPa for (a) static system (Xbx), (b) RRKF \(k = 100\) (Xbj), and RRKF \(k = 1000\) (Xbm) for one particular cycle (# 25). Contour interval for variances (and shading) is given in the legend. For analysis errors the contour interval is 50m.
7.4 Role of the background formulation and RRKF

Figure 7.20: as Fig. 7.19, but for forecast error variances and corresponding background fields $\mathbf{x}^b$. 
7.4 Role of the background formulation and RRKF

els b and c) but lacks resolution of the spatial location of individual structures in case of the static system (panel a). Note that the true errors (i.e., the differences $x^a - x^t$) are computed independently and thus provide a fairly strict test for the error variances computed within the assimilation system itself. Fig. 7.20b and Fig. 7.20c show forecast error variance predicted from the corresponding $P^a$ in Fig. 7.19. Consequently, these $P^f$ matrices are the background error statistics for the next assimilation cycle. Also shown in Fig. 7.20 are the corresponding background fields $x^b$ as a results of the (nonlinear) prediction starting from the analysis $x^a$. Fig. 7.20a shows the (static) background error variance as implied by the static $B$ matrix. Obviously, this variance field does not reflect the uncertainties associated with the background field shown. However, for the RRKF configurations shown in panels b ($k = 100$) and c ($k = 1000$) the error variance does reflect the potential uncertainties of the atmospheric flow. Focusing again on the northern Pacific it is clearly seen that the largest variance structures are associated with frontside of the troughs and regions of strong cyclogenic development. Similar and even more pronounced features are also present in the southern hemisphere. Comparing the predicted variance field from the RRKF using $k = 100$ vectors (panel b) to the RRKF using $k = 1000$ vectors (panel c) it becomes obvious that despite incorporating 100 HSVs the variance structure in panel b still resembles the static background error variance to a large degree.

![Figure 7.21](image)

Figure 7.21: Difference in RMS Z500 analysis errors between the static system (Xat) and the RRKF using $k = 100$ HSVs (Xbj; left panel). The right panel shows corresponding difference in terms of RMS forecast errors at day-4 over Europe.

To illustrate the impact of the RRKF on the quality of the forecast Fig. 7.21 shows differences in RMS analysis and forecast errors between the static system (Xat) and the RRKF using $k = 100$ HSVs (Xbj). Focusing on the northern hemisphere the largest improvement in analysis quality is obtained over the Pacific (as well as over the Atlantic). Fig. 7.21b shows a zoom at the European area at forecast day 4 (after 120h of forecast) corresponding to the analyses shown in Fig. 7.21a. It is evident from Fig. 7.21b that at day-4 the RRKF is able to improve the forecasts over Europe substantially on the basis of the more accurate analyses over the Oceans four days before.
7.5 Further experiments

In this section further experiments are described that address issues relevant in the context of both RRKF and EKF. These experiments have been carried out to study the impact of the model dynamics on the predicted error covariances (section 7.5.1) as well as the influence of scaled (static) background error covariance matrices on the performance of the RRKF (section 7.5.2).

7.5.1 Nonlinear covariance prediction

In the context of the EKF the time–propagation of the error covariances is carried out using the TL model operator $M$ although the model state itself is propagated in time with the nonlinear (NL) model operator $\mathcal{M}$. To quantify the impact of this assumption on the $P_f$ matrices obtained the prediction step for the error covariances has been carried out using a MC approach thus using the full nonlinear model $\mathcal{M}$. Specifically, this computation is carried out by assuming the initial pdf as normal with given mean $x^a$ and covariance structure $P^a$. Note that this is basically the same approach as in the sampling technique (see, eq. 4.2.2 in section 4.2) but using the full $P^a$ matrix and not its SV–decomposition. In the linear case (i.e., the EKF) the basic state trajectory is also started from the analysis $x^a$. Hence this configuration is a reasonable counterpart to the linear EKF propagation. The forecast error covariance matrix $P_f$ is then estimated on the basis of a time–evolved sample of size NSVMC from this initial pdf. Note that the numerical model error term as discussed in section 7.1.1 is also used within the NL EKF for fair comparison.

![Figure 7.22](image-url)  
**Figure 7.22:** Analysis and forecast errors for EKF and NL EKF in terms of (a) TE [J/kg] and (b) RMS Z500 [m] for gridpoints >30N (solid) and <30S (dashed) as a function of lead time [h].

To study the impact of the NL EKF and thus the assumption of linear error propagation cycling experiments have been carried for different sample sizes. Here we discuss results for
two specifications, namely NSVMC=5000 (experiment Xbv) and NSVMC=20000 (experiment Xbs), respectively. Table 7.4 summarizes the performance of these two experiments compared to the (linear) EKF (experiment Xbx). Let us first look at the impact of the sample size on the performance of the system. The results suggest, that on the basis of 5000 realizations the resulting covariance structures lack important features. In other words, this under-sampling is suboptimal in terms of analysis and forecast quality.

Fig. 7.22 shows analysis and forecast errors in terms of (a) TE [J/kg] and (b) RMS Z500 [m] as a function of lead time [h] for NL EKF and EKF. Comparing the NL EKF (Xbs; magenta curve) to the (linear) EKF (Xbx; orange curve) suggests that the nonlinear model dynamics has a significant impact on the predicted $P^f$ matrices and thus the performance of 4D–Var. The NL EKF reduces the analysis error to below 10 J/kg (see, table 7.4). In terms of RMS errors the largest reduction of errors occurs in the southern hemisphere (dashed curves in Fig. 7.22b) although the small reduction of analysis errors in the northern hemisphere (solid curves in Fig. 7.22b) leads to significantly more accurate forecasts.

Figure 7.23: Differences between 4D–Var RMS analysis increments [m] for EKF (Xbx) and NL EKF (NSVMC=20000, Xbs) at 500 hPa, averaged over 30 cases. The shaded areas indicate negative values, where NL EKF has smaller increments. Contours: +/- 10/20/40/80/160 m (solid red, positive values; dashed blue, negative with areas shaded).

Fig. 7.23 shows differences in RMS analysis increments between the EKF and the NL EKF. It is evident that the largest differences occur in those regions where the specification of $P^f$ has the largest impact on the performance of the 4D–var system (see, sections 7.3 and 7.4).

Investigation of individual assimilation cycles does not reveal substantial differences between $P^f$ obtained from either EKF or NL EKF. As an example, Fig. 7.24 shows geopotential height variances [$m^2$] as well as corresponding background errors ($x^b - x^f$) from
assimilation cycle #25 for EKF (Xbx) and NL EKF (Xbs). The variance fields obtained are very similar in both configurations however the NL EKF (lower panel) has smaller background errors. Careful inspection of the variance structures reveals small differences in the northern Pacific to the west of the American continent as well as in the South Pacific storm tracks. It is probably due to this small features that the NL EKF is more successful in reducing analysis and forecast errors. The structure of the off–diagonal elements (i.e., correlations) does not show any significant difference between EKF and NL EKF (not shown).

However, the NL EKF configuration shows a small but consistent improvement of the quality of the individual analyses and thus (on average) results in far more accurate forecasts. It is however unclear if this improvement is solely due to the nonlinear model
7.5 Further experiments

dynamics or also as a consequence of improved numerical stability. In view of these considerations, the results presented in this section are of preliminary nature and need to be investigated in greater detail.

7.5.2 Scaling issues in RRKF

The motivation for introducing a scaling parameter into the RRKF is twofold: First, this scaling allows for adjusting potential differences between the overall variance levels induced by the static $B$ and the set of $k$ HSVs. However, using $k < n$ (i.e., a subset of HSVs) amounts to neglecting part of the total variance contained in the full set of HSVs. Therefore the second motivation for using a scaling parameter is to be able to take into account the effect of neglecting variance through limiting the number of HSVs.

To motivate the impact of the scaling we repeat eq. (6.4.25) omitting cross–covariances by setting $\alpha = 0$. It is evident from eq. (7.5.1) that the scalar scaling factor $\beta$ has an impact on the estimated forecast error covariance matrix $\hat{P}_k$:

$$\hat{P}_k = \begin{bmatrix} \beta^{-1}E & 0 \\ 0 & 1 \end{bmatrix}$$ (7.5.1)

This impact does not amount to a simple scaling of the overall covariance matrix as it acts only on one block of the (now) block-diagonal matrix. Including cross–covariances (i.e., $\alpha \neq 0$) does not change this property of the algorithm. The following results present evidence that indeed this scaling has a significant impact on the performance of the RRKF in terms of analysis and forecast errors obtained. Moreover the results suggest that this scaling can be used to adjust the ratio between the two sources of covariance information, namely the set of $k$ HSVs and the static $B$, thereby making the RRKF system more powerful. On the other hand it is seen that an ill–specified $\beta$ deteriorates the performance of the system resulting in analysis errors larger than those obtained from the purely static system.

To study the impact of the scaling parameter $\beta$ RRKF experiments have been carried out for $k = 10$ and $k = 100$ HSVs and different specifications of $\beta$. As in the previous sections, table 7.3 (see, section 7.1) lists the experimental configurations described below. In addition, table 7.4 (see, section 7.2) shows selected results for the configurations considered in terms of analysis and forecast errors.

Fig. 7.25 shows the performance of the RRKF for different choices of $\beta$ and $k$ (i.e., the number of HSVs used) in terms of TE [J/kg] as a function of lead time [h]: These configurations include $k = 10/\beta=1.0$ (Xbd; dashed red), $k = 10/\beta=2.0$ (Xbf; dashed brown), $k = 10/\beta=5.0$ (Xbg; dashed maroon) as well as $k = 100/\beta=1.0$ (Xbj; green), $k = 100/\beta=0.5$ (Xbk; yellow), $k = 100/\beta=2.0$ (Xbl; turquoise), and $k = 100/\beta=5.0$ (Xbn; indigo). From Fig. 7.25 it is evident that the scaling factor $\beta$ does have an impact on the performance of the RRKF and thus the 4D–Var system. Comparing configurations involv-
Further experiments

7.5 Further experiments

Figure 7.25: Impact of scaling factor $\beta$ on RRKF performance. Show are analysis and forecast errors for configurations $k=10/\beta=1.0$ (dashed red), $k=10/\beta=2.0$ (dashed brown), $k=10/\beta=5.0$ (dashed maroon) as well as $k=100/\beta=1.0$ (green), $k=100/\beta=0.5$ (yellow), $k=100/\beta=2.0$ (turquoise), and $k=100/\beta=5.0$ (indigo) in terms of TE [J/kg] as a function of forecast time [h].

With $\beta = 2.0$ (dashed brown and cyan curves) to those using $\beta = 1.0$ (dashed red and green curves) it becomes evident that the performance of the system is improved significantly. On the basis of $k = 10$ HSVs the analysis error (in terms of TE) is reduced by roughly 20% (configurations Xbd and Xbf in table 7.4).

To investigate further the performance of the 4D–Var system with respect to the scaling parameter Fig. 7.26a shows analysis and forecast errors for RRKF with $k = 10$ HSVs in terms of RMS Z500 for northern (solid) and southern extratropics (dashed curves): $\beta = 1.0$ red curves and $\beta = 2.0$ brown curves, respectively. It is evident from Fig. 7.26a that the largest improvement is obtained for the southern hemisphere whereas the systems performs almost neutral for the northern hemisphere. Again, this is a direct consequence of the spatial density of the observations and thus the importance of background formulation in data sparse regions. Fig. 7.26b shows analysis errors in terms of TE [J/kg] for selected assimilation cycles (i.e., cycles #30 to #46). The coloring conventions are the same as in previous figures and corresponding curves for the static system (black) and the EKF (orange) are included for comparison. It is evident from Fig. 7.26b that for the set of assimilation cycles shown the RRKF using $k = 10$ HSVs and $\beta = 1.0$ (red curve) vectors performs worse
7.5 Further experiments

Figure 7.26: (a) Analysis and forecast errors for configurations $k=10/\beta=1.0$ (red) and $k=10/\beta=2.0$ (brown) in terms of RMS Z500 for gridpoints $>30\text{N}$ (solid) and $<30\text{S}$ (dashed curves) as a function of forecast time [h]. (b) Analysis errors in terms of TE [J/kg] as a function of assimilation cycle for configurations $k=10/\beta=1.0$ (red), $k=10/\beta=2.0$ (brown), and $k=100/\beta=1.0$ (green), $k=100/\beta=2.0$ (turquoise) as well as static (black) and EKF (orange) systems.

compared to the static system (black curve). However, by choosing $\beta = 2.0$ (and $k = 10$) the RRKF (brown curve) is able to improve the analysis compared to the static system (black curve). Similar conclusions are found for RRKF using $k = 100$ HSVs with the configuration with $\beta = 2.0$ (cyan curve) outperforming the (default) configuration $\beta = 1.0$ (green curve).

The reduction of analysis and subsequent forecasts errors through the "correct" choice of $\beta$ is also evident in terms of differences in RMS analysis (Fig. 7.27) and day–1 forecast errors (Fig. 7.28). As in Fig. 7.26a, the RRKF using $k = 10$ HSVs is considered for $\beta = 1.0$ and $\beta = 2.0$. Fig. 7.27 highlights the large improvement in analysis quality obtained for the southern hemisphere but also shows small but consistent improvements (10m in RMS sense) for the northern hemisphere. From Fig. 7.28 it is evident that these small improvements (in the NH) reduce the day–1 forecast errors by as much as 40m in the east Pacific as well as the Euro–Atlantic region.

In view of eq. (7.5.1) it is obvious that $\beta = 2.0$ amounts to decreasing the size of the $E$ block (i.e., the covariance structures implied through the HSVs). Fig. 7.29 shows the average over the 10 leading final–time HSVs computed for one particular assimilation cycle chosen randomly (cycle #43). Comparing the spatial locations of the final–time HSVs to the difference field shown in Fig. 7.27 suggests that the performance of the RRKF is closely related to the location and size of the final–time HSV structures.
7.5 Further experiments

Figure 7.27: Differences between 4D–Var RMS analysis errors [m] for RRKF with k=10 HSVs for $\beta=1.0$ (Xbd) and $\beta = 2.0$ (Xbf) at 500 hPa, averaged over 36 cases. The shaded areas indicate negative values, where $\beta = 2.0$ experiment (Xbf) has smaller RMS errors. Contours: +/-10/20/40/80/160 m (solid red, positive values; dashed blue, negative with areas shaded).

Figure 7.28: as Fig. 7.27, but RMS errors for day–1 forecasts.
Figure 7.29: Final–time HSV for one particular experiment randomly chosen from the 36 cases. Shown is the average over the first 10 leading final–time HSVs in terms of geopotential height. Contour interval is 60 m (positive values red, negative dashed blue).
Chapter 8

Covariance prediction with an operational forecasting system

This chapter summarizes experiments with the SV–based MC technique carried out within the framework of the ECMWF ensemble prediction system (EPS; Molteni et al. 1996). The ECMWF EPS is run operationally since December 1992 (daily run since May 1994) and major changes have been made to the system (see, Buizza and Palmer 1998; Buizza et al. 1998; Buizza et al. 1999; Buizza et al. 2002). For recent evaluations of the performance of the ECMWF EPS with special emphasise on severe weather events reference is made to (Mullen and Buizza 2001; Mullen and Buizza 2002; Chessa and Lalaurette 2001; Buizza and Chessa 2002).

The sampling technique (Ehrendorfer 1999) has been implemented by the author within the ECMWF EPS and is thus available as an alternative to the operational rotation technique (Molteni et al. 1996; Buizza et al. 1998). Since 22 January 2002 (model cycle CY24r3) the ECMWF EPS also includes perturbations in the tropics (Lalaurette 2002). These perturbations are generated with the sampling technique adopted here on the basis of the five leading (diabatic) SVs (i.e., M=50/N=5 in our notation). For properties of these tropical SVs reference is made to Barkmeijer et al. (2001).

8.1 Experiments with ECMWF EPS

To investigate the performance of the sampling technique within the framework of the ECMWF EPS ensemble experiments have been carried out. Specifically, these experiments are designed to address the impact of different configurations with regard to ensemble size $M$ and the number of SVs $N$ used for approximating the analysis error covariance matrix at initial time. A summary of the configurations considered is given in table 8.1. Sampling experiments have been carried out for selected dates during summer 2001 as well as December 2001. Emphasis has been given to the comparison of the sampling
As an example for the error dynamics in the ECMWF EPS results from a sampling experiment for 2001121912 are presented. This experiment uses $M = 50$ perturbations sampled from $N = 25$ SVs, thus this choice of parameters mimics the operational configuration. Fig. 8.1 shows temperature variance [K$^2$] at 850hPa at initial time. Also shown is 500hPa height field [dam] obtained from control forecast. The contour interval for the variances is given one the right–hand side of the figure (regions $> 2$K$^2$ shaded). It is seen from Fig. 8.1 that the (nonzero) structures are associated with regions of strong cyclogenic activity and that the largest variance obtained is about 10K$^2$ (at 175°W). Over the next 24 hours of the forecast the error growth is fairly small in terms of absolute values but the variance structures have spread substantially thus covering the baroclinic regions of the individual
Figure 8.1: Initial time (2001121912) temperature variance at 850hPa obtained from M=50/N=25 sampling experiment (contour interval indicated on right hand side of figure). Also shown is Z500 height field from control forecast (black contours).

Figure 8.2: as Fig. 8.1, but T850 variance for day–1 forecast as well as corresponding Z500 height field.
8.2 EPS: results

Figure 8.3: as Fig. 8.1, but T850 variance for day–2 forecast as well as corresponding Z500 height field.

Figure 8.4: as Fig. 8.1, but T850 variance for day–4 forecast as well as corresponding Z500 height field.
cyclones across the globe. Figures 8.2 to 8.4 show T850 variances (and corresponding height fields) for forecast days–1, 2, and 4. The largest variance is associated with the cyclone over the north–east Pacific (about 60°N, 170°W) showing variances of about 30 K$^2$ at day–2 and almost 45 K$^2$ at forecast day–4, respectively. Hence, the error doubling time estimated is slightly larger than 2 days. This value corresponds nicely to the curves shown in Fig. 1.3 where an error doubling time of about 2 days is obtained. The spatial structure of the variances fields at days–3 and 4 still shows the largest variances corresponding to regions of cyclogenic activity but with variances between 2 and 8 K$^2$ covering most of the extratropic flow.

Figure 8.5: Day–4 temperature variance [K$^2$] at 850hPa obtained from ECMWF operational forecast starting from 2001121912 (contour interval indicated on right hand side of figure). Also shown is Z500 height field [dam] from control forecast (black contours).

Fig. 8.5 shows day–4 temperature variances for T850 estimated from the operational ECMWF EPS for the same initial date as in Fig. 8.1. Comparing Fig. 8.5 to Fig. 8.4 reveals that the error variances are smaller in the operational run compared to the corresponding sampling experiment. The two figures are very similar in an overall sense however the structures associated with the largest variances have a larger magnitude in Fig. 8.4.

Fig. 8.6 shows the corresponding figure but for sampling using $M = 100$ ensemble members on the basis of $N = 25$ SVs. It is evident that again the structures for day–4 variance is very similar compared to figures Fig. 8.4 and Fig. 8.5.

Fig. 8.7 shows correlations in terms of geopotential height of one particular gridpoint (50°N, 0°E) with all other gridpoints on the globe. This correlation structure is estimated from the operational EPS forecast starting from 2001121912. It is evident from Fig. 8.7 that estimating off–diagonal elements of the forecast error covariance matrix $P^f$ from only
50 ensemble members gives highly unrealistic correlation structures. Fig. 8.7 suggests that geopotential height of one point over Europe (i.e., 50°N, 0°E) is highly correlated with geopotential height over Australia or the South Pacific. Obviously, such correlations are highly unrealistic in the real atmosphere and are simply the result of a systematic undersampling of the correlation structure. To overcome this undersampling problem the number of ensemble members has to be increased.

It is however evident from Fig. 8.8, that the correlations can be improved substantially on the basis of only 50 members if the sampling technique is used for the generation of initial perturbations. As Fig. 8.7, Fig. 8.8 shows Z500 correlations but estimated from a sampling experiment using \( M = 50 \) members (and \( N = 25 \) SVs). This improvement in the estimation of off-diagonal elements is due to the fact that the operational rotation technique (underlying Fig. 8.7) imposes perturbations with opposite sign at the initial time (see, Molteni et al. 1996).

Fig. 8.9 again shows Z500 correlations but for sampling using \( M = 100 \) members on the basis of \( N = 25 \) SVs. It is evident from Fig. 8.9 that this configuration ensures that the unrealistic correlation structures found in the previous figures disappear, i.e., Fig. 8.9 does not shown any significant correlations with points in the southern hemisphere.
Figure 8.7: Day–4 geopotential height correlations at 500hPa obtained from ECMWF operational forecast starting from 2001121912. Also shown is Z500 height field from control forecast (black contours).

Figure 8.8: as Fig. 8.7, but for sampling experiment $M = 50/N = 25$ starting from 2001121912.
Figure 8.9: as Fig. 8.7, but for sampling experiment $M = 100/N = 50$ starting from 2001121912.
8.2.1 EPS: skill scores

The performance of the sampling technique has been investigated in terms of skill scores. For the verification of probability forecasts reference is made to the excellent book by Wilks (1995).

Fig. 8.10 shows Z500 anomaly correlation as well as RMS error [m] computed with respect to the ensemble mean and evaluated for the northern hemisphere. The curves correspond to the operational EPS (red) and sampling experiments $M = 50/N = 25$ (blue), $M = 50/N = 50$ (green), and $M = 100/N = 50$ (black), respectively. All curves are computed as mean over 8 cases in December 2001. The top panel in Fig. 8.10 shows the difference in skill between the ensemble mean and the control forecasts. The lower panel shows corresponding difference in RMS errors between ensemble mean and the control. It is evident from Fig. 8.10 that after day–3 the ensemble mean is more skillful than the control forecast. Moreover, the ensemble mean of the sampling experiments performs slightly better after day–6.

Fig. 8.11 shows area under the Relative Operating Characteristic (ROC) curve for the same experimental configurations as in Fig. 8.10 (same coloring conventions). Curves shown correspond to T850 for the event $T > 8K$ evaluated over the northern hemisphere. The area under the ROC curve is an overall measure of the quality of an ensemble forecast. Basically, it is a generalization of the concept of contingency tables to probabilistic forecasts. A ROC area of one is perfect; 0.5 is considered as the lower bound for a useful forecast. For detailed description of the properties of this verification measure references is made to Stanski et al. (1989). It is evident from Fig. 8.11 that all EPS configurations behave similar in terms of ROC area. The lower panel in Fig. 8.11 shows the change in ROC area with respect to the operational EPS thus indicating a slight improvement for the sampling experiments $M = 50/N = 50$ (green curve) and $M = 100/N = 50$ (black curve).

Fig. 8.12 shows Brier scores as well as Brier skill scores (BSS) for the same configurations as in the previous figures. As in Fig. 8.11, the verification area is the northern hemisphere and the statistics is shown for T850 evaluated for the event $T > 8K$. The BS is essentially a mean–squared error for probability forecasts and subsequent binary events (i.e., threshold $T > 8K$: yes/no). The BSS is a skill score computed for the BS with respect to climatology. Thus BSS=1 is a perfect forecast, BSS=0 indicates no improvement compared to climatology. Fig. 8.12 suggests that in terms of BSS all configurations (as well as the operational EPS) perform very similar.

Fig. 8.13 shows the number of outliers computed for six verification areas for sampling experiments with different choices of $M$ and $N$ as well as the operational EPS (red curve). The basic idea of this score is as follows: The $M$ ensemble members are used to partition the real line into distinct intervals. Assuming a random sample of the initial
Figure 8.10: Z500 anomaly correlation and RMS error [m] with respect to the control forecast as function of forecast day computed for northern hemisphere (8 cases): operational EPS (red), sampling $M = 50/N = 25$ (blue), $M = 50/N = 50$ (green), and $M = 100/N = 50$ (black).
Figure 8.11: ROC area T850 (top panel) and relative change with respect to the operational EPS (lower panel) for the threshold 8K computed for northern hemisphere (8 cases): operational EPS (red), sampling $M = 50/N = 25$ (blue), $M = 50/N = 50$ (green), and $M = 100/N = 50$ (black).
Figure 8.12: Brier (skill) score T850 for the event T>8K computed for northern hemisphere (8 cases): operational EPS (red), sampling $M = 50/N = 25$ (blue), $M = 50/N = 50$ (green), and $M = 100/N = 50$ (black).
Figure 8.13: Number of outliers in terms of T850 for 6 verification areas (8 cases): operational EPS (red), sampling $M = 50/N = 25$ (blue), $M = 50/N = 50$ (green), and $M = 100/N = 50$ (black).
pdf and neglecting model errors each interval is equally likely to contain the analysis (averaged over the verification area). For details refer to Strauss and Lanzinger (1996) and Buizza (1997). Fig. 8.13 shows percentage of outliers relative to the expected value of $\frac{\sqrt{2}}{M+1} \times 100\%$. It is evident from Fig. 8.13 that the sampling technique reduces the number of outliers for all configurations. The largest reduction is obtained for the ensemble using $M = 100$ members. Interestingly, the sampling configuration using $M = 50$ members on the basis of $N = 50$ (green curve) SVs performs almost as good as the configuration $M = 100/N = 50$ (black curve). This is probably due to the fact that including a larger number of SVs leads to more large scale error structures and thus lower probabilities of the analysis lying outside of the two extremes of the ensemble.
Chapter 9

Summary and conclusions

This chapter summarizes the main results from the cycling experiments carried out within the QG/4D–Var system as well as the experiments within the ECMWF EPS. A short outlook addresses open questions and motivates topics that could form the basis of future research related to this thesis.

9.1 QG/4D–Var system

We have studied the impact of the background formulation in a QG/4D–Var system. The background formulations considered include specifications using a static background error covariance matrix ($P^f \equiv B$) as well as dynamic background error covariance matrices $P^f$ available from either EKF or RRKF. The formulation of both EKF and RRKF is explicit in terms of the matrix algebra involved and thus allows for a comparison in great detail. The RRKF system is designed to blend (dynamic) information about the present forecast error statistics as contained in the leading HSVs and static information available from the B matrix. For a summary of the configurations as well as and illustration of the implementation reference is made again to figures 6.1 and 6.3, respectively.

Cycling experiments within this QG/4D–Var system have been carried out to assess the performance of the 4D–Var system using different background specifications. As a measure of the performance of the system analysis and forecast errors have been computed with respect to a truth–run (i.e., a pre–specified solution of the nonlinear QG system) for either total–energy or RMS errors in terms of geopotential height for the northern and southern extratropics, respectively. For a summary of the performance of the cycling experiments described reference is made to table 7.4. Additionally, differences between RMS errors of two sets of cycling experiments have been computed as a means for extracting the relevant differences from series of 48 consecutive assimilation cycles.

We stress that all cycling experiments carried out share a common experimental setup besides the specification of $P^f$. 
9.1 QG/4D–Var system

- 12h assimilation window
- pre-specified sets of observations with time-independent errors
- no re-scaling of the background errors
- background field is updated using the nonlinear model: $x^b = M(x^n)$
- each cycling experiment is initialized with a static assimilation run (i.e., no difference in assimilation cycle #1 between different configurations)

Because of this initialization, assimilation cycles #1 to #12 are not considered for the calculation of analysis and forecast errors as well as differences in RMS errors.

The results show that the QG/4D–Var system is indeed sensitive to the background specification in terms of analysis errors obtained. The overall level of variance as implied through the static $B$ matrix has been adjusted to get a reasonably balanced system in terms of analysis errors. This adjustment is necessary in order to avoid systematic trends in analysis quality with respect to the truth run (see, section 6.2.1).

First, the performance of the static system has been compared to the EKF. Within the EKF a (small) numerical model error term has been included in order to avoid numerical instabilities. This numerical model error term prevents individual entries of the $P_f$ matrix from becoming too small (i.e., perfect observation of single wavenumbers) and thus causing the matrix to become (almost) singular. Compared to the static system the EKF is superior in terms of analysis and forecast errors obtained with largest improvements over data sparse areas (i.e., the Oceans). This improvement is a direct consequence of the more realistic background error statistics since the structure of analysis increments is dominated by the background in data sparse areas (see, e.g., Fig. 7.2). The EKF reduces analysis errors to about one third in terms of TE (Fig. 7.1). To assess the importance of the day-by-day variations of the $P_f$ matrix, a configuration denoted "best static" system has been investigated. This "best static" system uses a static background error covariance matrix computed as the mean of the $P_f$ matrices obtained from the EKF (for the same series of assimilation cycles). The "best static" system performs remarkably well compared to the (ordinary) static system (analysis errors are reduced to one half in terms of TE) but is unable to approach the performance of the EKF. As discussed in section 7.4, these results are strongly influenced by the choice of the observational distribution and thus seem to be over-optimistic. On the basis of this comparison it is evident that only if the day–by–day variations of $P_f$ are taken into account the information provided to the 4D–Var system (i.e., background and observations) is used in the most efficient way. The $P_f$ matrices for the three configurations have been investigated in terms of variances (Fig. 7.13) and correlation structures implied. The static $B$ matrix, by construction, corresponds to a flat background
error specification with isotropic correlations. Hence, the static \( B \) does not reflect the uncertainties associated with the present atmospheric flow. The \( P^f \) matrices obtained from the EKF clearly reflect the distribution of observations through reduced variances over the continents as well as the uncertainties corresponding to the areas of cyclogenic activity especially over the Oceans. The nonzero structures of the correlation patterns reflect (as implied by the EKF) the importance of the flow–dependent specification of \( P^f \) (Fig. 7.14).

The RRKF has been the focus of the second set of experiments carried out within the QG/4D–Var system. Comparing the performance of the RRKF in terms of analysis quality obtained to both the static system and the EKF reveals that the RRKF is able to improve the analyses (see, e.g., Fig. 7.15). This improvement is however strongly dependent on the number of HSVs used within the RRKF. Using only \( k = 10 \) HSVs the RRKF performs neutral or even slightly worse compared to the static system. For \( k > 100 \) the RRKF is able to improve the quality of the analyses (compared to the static system) but a (very) large number of HSVs (i.e., more than 1000) is necessary for the RRKF to perform similar compared to the EKF. The structures of analysis increments differ substantially among the configurations considered. There is however some degree of similarity in terms of spatial locations of the increments between RRKF and EKF: The largest increments are found in the southern hemisphere and in data sparse areas in general (see, Fig. 7.17) corresponding to the areas where the largest improvement (in terms of analysis quality) is found (Fig. 7.18). Comparison of analysis errors to analysis variances implied by \( P^a \) reveals a good agreement for both RRKF and EKF. For the static system this agreement is slightly worse (Fig. 7.19). In view of the performance of the RRKF, the overall conclusion is that the RRKF is able to improve the quality of the analyses and subsequent forecasts if more than \( k = 100 \) HSVs are used. Using \( k = 100 \) HSVs results in an overall reduction of analysis of roughly 20% (in terms of RMS errors) with the largest improvement corresponding to the Pacific Ocean as well as the southern hemisphere in general (see, table 7.4 and Fig. 7.18).

The cycling experiments using nonlinear propagation of the error covariances (NL EKF) have demonstrated the impact of the nonlinear model dynamics on the performance of the QG/4D–Var system. The NL EKF is able to improve the analyses significantly compared to the (linear) EKF (Fig. 7.23). The investigation of the impact of the nonlinear dynamics on the propagation of error covariances and possible implications for 4D–Var are beyond the scope of this work and need to be addressed in more detail in subsequent studies.

Further experiments have been carried out to investigate the impact of scaling the static \( B \) in the RRKF through a scaling parameter \( \beta \). It has been demonstrated that this scaling has an impact on the performance of the QG/4D–Var system (Fig. 7.25). Moreover, an appropriate choice of the scaling parameter \( \beta \) is necessary to use the RRKF in the most efficient way. The results suggest, that the scaling might be important to achieve a certain balance between the dynamic information as contained in the set of HSVs and the static
9.1 QG/4D–Var system

B. This balance is necessary to improve the performance of the RRKF in terms of analysis and forecast errors (see, figures 7.27 and 7.28, respectively). The spatial regions where the largest differences occur are clearly related to the nonzero structures implied through the set of HSVs (Fig. 7.29).

9.1.1 Comments on the role of the RRKF in operational 4D–Var

In operational data assimilation systems, the RRKF has been seen as a promising approach for including more influence of the dynamics into 4D–Var (see, e.g., Rabier et al. 1997). Although extensive experimentation has been carried out at ECMWF the RRKF has an entirely neutral impact on forecast scores (quoted from Fisher and Andersson 2001) in the operational ECMWF 4D–Var system. This section is an attempt to shed some light on the role of the RRKF within operational 4D–Var systems. Keeping in mind the large gap in resolution and physical realism between the QG/4D–Var systems used within this work and the current operational practise the nature of this section is highly speculative.

The results from the QG system suggest that the impact of the RRKF is small and, on average, slightly negative for O(10) HSVs compared to the static system. Although significant differences are found on a case–to–case basis the average scores in both TE and RMS sense are slightly negative. This decay in analysis (and forecast) quality might be related to the blending of the two sources of covariance information (i.e., static $B$ and the HSVs) and thus lead to a (physically) unbalanced $P_f$ matrix in regions where HSVs are present. It seems however possible to overcome, at least partly, this deficiency by means of scaling of the $B$ matrix within the RRKF algorithm. By introducing this scaling factor $\beta$ (see, section 7.5.2) the performance of the RRKF (using 10 HSVs) is improved significantly with a consistent reduction of both analysis and forecast errors compared to the static system. Similar characteristics are observed for the experiments using $k = 100$ HSVs. The performance of the system is strongly influenced by the choice of the scaling parameter resulting in reduction of analysis (and forecast) errors between 10 and 20 % depending on the specification of $\beta$. Judging from these results the blending of the dynamic and static covariance information within the RRKF is a crucial step for the success of the RRKF especially when the number of HSVs is small. The results suggest that for increasing number of HSVs the impact of $\beta$ becomes smaller. Comparison of the RRKF to the EKF suggests that only on the basis of a very large number of HSVs (say, 500 to 1000) the RRKF performs similar to the EKF. In view of the theoretical properties of the HSVs this performance is discouraging and might be related to the specifications necessary for the computation of the HSVs themselves.
9.2 Sampling experiments

We have investigated the properties of the sampling technique within the operational ECMWF EPS. The experiments carried out within the EPS were mainly concerned with comparing the properties of the sampling method to the operational rotation method. These methods for generating initial–time ensemble members are very similar but the sampling technique allows for generating an arbitrary number of perturbations $M$ from a given set of $N$ SVs. In view of this property the sampling method is superior to the rotation method as it allows for greater flexibility. The theoretical motivation for the sampling is that it generates the ensemble of $M$ perturbations consistent with the information about analysis errors as contained in the set of $N$ SVs.

To illustrate the error growth within the EPS forecast error variances have been shown for different lead times. In view of the estimated forecast error variances the impact of increasing the number of ensemble members is very small. In addition, the sampling performs almost indistinguishable compared to the operational rotation method. However, inspection of structures implied through the off–diagonal elements of the forecast error covariance matrix reveals large differences between the configurations investigated. In view of estimating correlations structure the sampling technique performs better compared to the operational system. Increasing the size of the ensemble (from 50 to 100 members) results in improved reconstruction of the correlation patterns (see, Figs. 8.7 to 8.9).

The performance of the individual configurations has been compared in terms of probabilistic skill scores. The results suggest that the performance of the sampling is neutral to slightly positive compared to the operational EPS and almost independent of the configuration (i.e., $M$ and $N$) chosen. The main differences in terms of performance is found in the number of outliers where the sampling experiments performed systematically better than the operational EPS (see, Fig. 8.13).

A potential weakness of all experiments carried out is the use of total–energy SVs and not HSVs since the latter are not available operationally. Therefore the properties of the sampling method as described here are only of a preliminary nature and need to be investigated in greater details when HSVs become available routinely.

9.3 Outlook

The primary question of this work has been to study the role of the background formulation in 4D–Var. To that aim three possible configurations for specifying the background errors have been considered, namely a static system, EKF, and RRKF. Obviously, this choice represents only a small fraction of possible approaches towards including flow–dependent background errors in 4D–Var. At the time of writing, the most promising approaches seem to be appropriate formulations of the Reduced Order Kalman Filter as well as the Ensem-
9.3 Outlook

ble Kalman Filter. Investigating these approaches within the QG/4D–Var system might shed light on the ongoing discussion about (suboptimal) variants of the Kalman filter. In view of the RRKF considered here, further research is needed with regard to the issue of balancing the two sources of covariance information within the RRKF. Concerning the QG/4D–Var system itself several improvements towards a more realistic system might be envisaged. These improvements concern primarily the assumptions made for the specification of the observational information (e.g., spatial distribution and error statistics) as well as the model formulation itself. Especially the poor vertical discretization suppresses important baroclinic effects. Finally, the assumption of linearity for the propagation of error covariances has to be studied in much more detail.

The properties of the sampling technique need to be investigated in much more detail when HSVs are available operationally. It is thought that on the basis of HSVs the sampling technique might be an enhancement of the operational EPS.
Appendix A

List of acronyms and abbreviations

<table>
<thead>
<tr>
<th>acronym</th>
<th>meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>3D–Var</td>
<td>Three–Dimensional Variational data assimilation</td>
</tr>
<tr>
<td>4D–Var</td>
<td>Four–Dimensional Variational data assimilation</td>
</tr>
<tr>
<td>BSS</td>
<td>Brier Skill Score</td>
</tr>
<tr>
<td>DA</td>
<td>Data Assimilation</td>
</tr>
<tr>
<td>DAO</td>
<td>Data Assimilation Office</td>
</tr>
<tr>
<td>ECMWF</td>
<td>European Centre for Medium–Range Weather Forecasts</td>
</tr>
<tr>
<td>EKF</td>
<td>Extended Kalman Filter</td>
</tr>
<tr>
<td>EPS</td>
<td>Ensemble Prediction System</td>
</tr>
<tr>
<td>HSV</td>
<td>Hessian Singular Vector</td>
</tr>
<tr>
<td>KNMI</td>
<td>Koninklijk Nederlands Meteorologisch Instituut</td>
</tr>
<tr>
<td>MC</td>
<td>Monte Carlo</td>
</tr>
<tr>
<td>NCAR</td>
<td>National Center for Atmospheric Research</td>
</tr>
<tr>
<td>NL</td>
<td>Non–Linear</td>
</tr>
<tr>
<td>NMC</td>
<td>National Meteorological Center</td>
</tr>
<tr>
<td>NRL</td>
<td>Naval Research Laboratory</td>
</tr>
<tr>
<td>NWP</td>
<td>Numerical Weather Prediction</td>
</tr>
<tr>
<td>PV</td>
<td>Potential Vorticity</td>
</tr>
<tr>
<td>PPS</td>
<td>Probabilistic Prediction System</td>
</tr>
<tr>
<td>pdf</td>
<td>probability density function</td>
</tr>
<tr>
<td>QG</td>
<td>quasigeostrophic</td>
</tr>
<tr>
<td>ROC</td>
<td>Relative Operating Characteristic</td>
</tr>
<tr>
<td>RRKF</td>
<td>Reduced–Rank Kalman Filter</td>
</tr>
<tr>
<td>RMS</td>
<td>Root–Mean–Square</td>
</tr>
<tr>
<td>SV</td>
<td>singular vector</td>
</tr>
<tr>
<td>TE</td>
<td>Total Energy</td>
</tr>
<tr>
<td>TL</td>
<td>Tangent–linear</td>
</tr>
</tbody>
</table>

Table A.1: Summary of acronyms and abbreviations.
# List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Illustration of NWP: From Observations to weather forecasts</td>
<td>2</td>
</tr>
<tr>
<td>1.2</td>
<td>4D–Var in the Lorenz 1984 model: Assimilation experiment involving both</td>
<td>4</td>
</tr>
</tbody>
</table>
|        | observation and background term with (a) large ($\sigma_b^2 = 1.0$, $\sigma_o^2 = 1.0$) and
|        | (b) small ($\sigma_b^2 = 0.1$, $\sigma_o^2 = 1.0$) background error variances.
|        | Shown are the trajectories ($Z$ component) of the first guess ($fg$), the analysis
|        | ($ana$), and the background ($bg$) as a function of time. The observations
|        | are shown as crosses ($obs$).                                               |      |
| 1.3    | Error growth in the ECMWF EPS: Show is (total) T850 variance [K$^2$] obtained
|        | from ECMWF EPS as a function of forecast time [h].                          | 6    |
| 1.4    | Temperature variance [K$^2$] (850hPa) estimated from ECMWF EPS at (a) initial
|        | time and (b) forecast day–5. Contour interval is 0.8K at initial time (a) and
|        | 8K at day–5 (b).                                                           | 6    |
| 2.1    | Illustration of 4D–Var (adapted from Bouttier and Courtier 1999).           | 9    |
| 5.1    | Illustration of the vertical discretization of the QG model.                | 22   |
| 5.2    | Illustration of assimilation window and cycling.                            | 23   |
| 5.3    | Observation locations ($\times$) at three model levels (200, 500, and 800 hPa)
|        | for one particular time–slot. The total number of observations is $p=326$ in
|        | this case.                                                                  | 24   |
| 6.1    | Illustration of background formulation for cycling in 4D–Var.               | 27   |
| 6.2    | Analysis errors in terms of TE [J/kg] as a function of analysis cycle in static
|        | system for different specifications of observations.                        | 29   |
| 6.3    | Illustration of dynamic background configurations.                          | 30   |
| 6.4    | Illustration of RRKF.                                                       | 30   |
| 7.1    | Analysis and forecast errors for static system ($Xat$; black), "best static"
|        | system ($Xbc$; violet), and EKF ($Xbx$; orange) in terms of total energy [J/kg]
|        | as a function of forecast time [h].                                         | 41   |
7.2 Differences between 4D–Var RMS analysis errors [m] for fully dynamic background (EKF) and static B at 500 hPa, averaged over 36 cases. The shaded areas indicate negative values, where EKF has smaller RMS errors than the static system. Contours: +/- 10/20/40/80/160 m (solid red, positive values; dashed blue, negative with areas shaded).

7.3 as Fig. 7.2, but RMS background errors.

7.4 as Fig. 7.2, but RMS errors for day–1 forecasts.

7.5 as Fig. 7.2, but RMS errors for day–2 forecasts.

7.6 as Fig. 7.2, but RMS errors for day–3 forecasts.

7.7 as Fig. 7.2, but RMS errors for day–4 forecasts.

7.8 Differences between 4D–Var RMS increments [m] for fully dynamic background (EKF) and static B at 500 hPa, averaged over 36 cases. The shaded areas indicate negative values, where EKF has smaller increments than the static system. Contours: +/- 2/4/8/16/32 m (solid red, positive values; dashed blue, negative with areas shaded).

7.9 Geopotential height analysis increments [m] for static system (Xat) at three model levels (200, 500, and 800 hPa). Contour interval is 25m, with negative values dashed blue.

7.10 as Fig. 7.9, but for EKF.

7.11 as Fig. 7.2, but for "best static" (Xbc) and static (Xat) B. The shaded areas indicate negative values, where configuration Xbc has smaller RMS errors than the static system. Contours: +/- 10/20/40/80/160 m (solid red, positive values; dashed blue, negative with areas shaded).

7.12 as Fig. 7.8, but for "best static" (Xbc) and static (Xat) B. The shaded areas indicate negative values, where "best static" system has smaller increments than the static system. Contours: +/- 2/4/8/16/32 m (solid red, positive values; dashed blue, negative with areas shaded).

7.13 Geopotential height variances [m$^2$] (i.e., diagonal elements of $P^f$) for (a) the static B, (b) the "best static" B, and (c) the EKF (cycle #25). Contour interval is given between panels b and c (shaded regions correspond to variances > 200m$^2$). Same contours and shading used in all panels.

7.14 Correlation structures at 800 hPa implied by $P^f$ of one selected gridpoint (at 500 hPa) with all other gridpoints on the globe for (a) the static B, (b) the "best static" B, and (c) the EKF (cycle #25). Contour interval is 0.2 (negative values dashed blue).
7.15 Analysis and forecast errors for RRKF in terms of total energy [J/kg] as a function of forecast time [h]: $k = 10$ (Xbd; red), $k = 100$ (Xbj; light green), $k = 200$ (Xbo; dark green), and $k = 1000$ (Xbm; blue). Also shown are the corresponding curves for the static system (Xat; black) and the EKF (Xbx; orange) as shown in Fig. 7.1 (mean over 36 cases).

7.16 as Fig. 7.15, but in terms of RMS Z500 [m] for gridpoints (a) $>30^\circ$N (northern extratropics) and (b) $<30^\circ$S (southern extratropics).

7.17 Analysis increments in terms of geopotential height [m] for (a) the static system (Xat), (b) RRKF using $k = 100$ HSVs (Xbj), and (c) the EKF (Xbx) for one particular assimilation cycle (cycle #36). Contour interval is 25m, negative values dashed blue.

7.18 Difference between RMS Z500 increments for static system (Xat) and RRKF with k=100 HSVs (Xbj). The shading indicates regions where the RRKF has smaller analysis increments than the static system. Contours: +/- 2/4/8/16/32 m (solid red, positive values; dashed blue, negative with areas shaded; same shading conventions as in Fig. 7.8).

7.19 Analysis error variances [m$^2$] (implied by $P^a$) and corresponding errors $(x^a - x^t)$ [m] in terms of geopotential height at 500hPa for (a) static system (Xbx), (b) RRKF $k = 100$ (Xbj), and RRKF $k = 1000$ (Xbm) for one particular cycle (# 25). Contour interval for variances (and shading) is given in the legend. For analysis errors the contour interval is 50m.

7.20 as Fig. 7.19, but for forecast error variances and corresponding background fields $x^b$.

7.21 Difference in RMS Z500 analysis errors between the static system (Xat) and the RRKF using $k = 100$ HSVs (Xbj; left panel). The right panel shows corresponding difference in terms of RMS forecast errors at day–4 over Europe.

7.22 Analysis and forecast errors for EKF and NL EKF in terms of (a) TE [J/kg] and (b) RMS Z500 [m] for gridpoints $>30^\circ$N (solid) and $<30^\circ$S (dashed) as a function of lead time [h].

7.23 Differences between 4D–Var RMS analysis increments [m] for EKF (Xbx) and NL EKF (NSVMC=20000, Xbs) at 500 hPa, averaged over 30 cases. The shaded areas indicate negative values, where NL EKF has smaller increments. Contours: +/- 10/20/40/80/160 m (solid red, positive values; dashed blue, negative with areas shaded).

7.24 Z500 background error variance [m$^2$] and background errors $(x^b - x^t)$ (contour interval: 50m) for EKF (Xbx) and NL EKF (Xbs). Shading conventions are given in the legend.
7.25 Impact of scaling factor $\beta$ on RRKF performance. Show are analysis and forecast errors for configurations $k=10/\beta=1.0$ (dashed red), $k=10/\beta=2.0$ (dashed brown), $k=10/\beta=5.0$ (dashed maroon) as well as $k=100/\beta=1.0$ (green), $k=100/\beta=0.5$ (yellow), $k=100/\beta=2.0$ (turquoise), and $k=100/\beta=5.0$ (indigo) in terms of TE [J/kg] as a function of forecast time [h].

7.26 (a) Analysis and forecast errors for configurations $k=10/\beta=1.0$ (red) and $k=10/\beta=2.0$ (brown) in terms of RMS Z500 for gridpoints $>30N$ (solid) and $<30S$ (dashed curves) as a function of forecast time [h]. (b) Analysis errors in terms of TE [J/kg] as a function of assimilation cycle for configurations $k=10/\beta=1.0$ (red), $k=10/\beta=2.0$ (brown), and $k=100/\beta=1.0$ (green), $k=100/\beta=2.0$ (turquoise) as well as static (black) and EKF (orange) system.

7.27 Differences between 4D–Var RMS analysis errors [m] for RRKF with $k=10$ HSVs for $\beta=1.0$ (Xbd) and $\beta = 2.0$ (Xbf) at 500 hPa, averaged over 36 cases. The shaded areas indicate negative values, where $\beta = 2.0$ experiment (Xbf) has smaller RMS errors. Contours: +/- 10/20/40/80/160 m (solid red, positive values; dashed blue, negative with areas shaded).

7.28 as Fig. 7.27, but RMS errors for day–1 forecasts.

7.29 Final–time HSV for one particular experiment randomly chosen from the 36 cases. Shown is the average over the first 10 leading final–time HSVs in terms of geopotential height. Contour interval is 60 m (positive values red, negative dashed blue).

8.1 Initial time (2001121912) temperature variance at 850hPa obtained from M=50/N=25 sampling experiment (contour interval indicated on right hand side of figure). Also shown is Z500 height field from control forecast (black contours).

8.2 as Fig. 8.1, but T850 variance for day–1 forecast as well as corresponding Z500 height field.

8.3 as Fig. 8.1, but T850 variance for day–2 forecast as well as corresponding Z500 height field.

8.4 as Fig. 8.1, but T850 variance for day–4 forecast as well as corresponding Z500 height field.

8.5 Day–4 temperature variance [K$^2$] at 850hPa obtained from ECMWF operational forecast starting from 2001121912 (contour interval indicated on right hand side of figure). Also shown is Z500 height field [dam] from control forecast (black contours).

8.6 as Fig. 8.5, but for sampling experiment $M = 100/N = 50$ starting from 2001121912.
8.7 Day–4 geopotential height correlations at 500hPa obtained from ECMWF operational forecast starting from 2001121912. Also shown is Z500 height field from control forecast (black contours). .................................................. 75
8.8 as Fig. 8.7, but for sampling experiment $M = 50/N = 25$ starting from 2001121912. ................................................................. 75
8.9 as Fig. 8.7, but for sampling experiment $M = 100/N = 50$ starting from 2001121912. ................................................................. 76
8.10 Z500 anomaly correlation and RMS error [m] with respect to the control forecast as function of forecast day computed for northern hemisphere (8 cases): operational EPS (red), sampling $M = 50/N = 25$ (blue), $M = 50/N = 50$ (green), and $M = 100/N = 50$ (black). ................................. 78
8.11 ROC area T850 (top panel) and relative change with respect to the operational EPS (lower panel) for the threshold 8K computed for northern hemisphere (8 cases): operational EPS (red), sampling $M = 50/N = 25$ (blue), $M = 50/N = 50$ (green), and $M = 100/N = 50$ (black). ................. 79
8.12 Brier (skill) score T850 for the event T>8K computed for northern hemisphere (8 cases): operational EPS (red), sampling $M = 50/N = 25$ (blue), $M = 50/N = 50$ (green), and $M = 100/N = 50$ (black). ....................... 80
8.13 Number of outliers in terms of T850 for 6 verification areas (8 cases): operational EPS (red), sampling $M = 50/N = 25$ (blue), $M = 50/N = 50$ (green), and $M = 100/N = 50$ (black). ....................... 81
List of Tables

7.1 Common parameters and specifications in 4D–Var experiments. . . . . . 36
7.2 Summary of specifications for static and EKF experiments. . . . . . . . 37
7.3 Summary of specifications for RRKF experiments. . . . . . . . . . . 38
7.4 Summary of 4D–Var results: analysis and forecast errors. . . . . . . . 39
8.1 Experimental configurations for sampling experiments with ECMWF EPS. 70
A.1 Summary of acronyms and abbreviations. . . . . . . . . . . . . . . . . 89


Department Memorandum, R43.2/ME/80, 24 pp. [Available from ECMWF, Shinfield Park, Reading RG2 9AX, England].


Danksagung


Unserem langjährigen Projektleiter und Betreuer Herrn Prof. Dr. M. Ehrendorfer danke ich für seine Unterstützung meiner Arbeiten und die Möglichkeit der Mitarbeit in den von ihm geleiteten Forschungsprojekte.

Ohne die Hilfe und Ratschläge von Leo Haimberger wäre diese Arbeit nur schwer möglich gewesen. Ausserdem war Leo Haimberger zu jeder Zeit ein offener und interessierter Ansprechpartner für Diskussionen aller Art und hat so wertvolle Anstösse zum Gelingen dieser Arbeit gegeben.

Bei meinen Kollegen Bodo Ahrens, Felix Hamelbeck und Christoph Matulla möchte ich mich für die vielen interessanten Diskussionen sowie das Interesse an meiner Arbeit bedanken.

Ich danke Herrn Prof. Dr. M. Hantel für das produktive Umfeld in seiner Arbeitsgruppe, sowie den Institutsvorständen Prof. Dr. W. Seiberl und Prof. Dr. R. Steinacker für die Unterstützung von Seiten des Instituts.


I would like to thank the staff at ECMWF for the hospitality during my visits to the centre. In particular, I am indepted to Jan Barkmeijer (KNMI), Roberto Buizza, Tim Palmer, Mike Fisher, and Erik Andersson for numerous discussions related to my research efforts and their continuous interest in my work.

User support at ECMWF and in particular Norbert Kreitz and John Greenaway as well as Nils Wedi did a tremendous job in dealing with numerous technical problems that have arisen on many circumstances.

Many discussions with Ron Errico (DAO) and Joe Tribbia (NCAR) provided a deeper understanding into a broad range of issues regarding atmospheric predictability and data assimilation.

Helpfull suggestions by Anthony Hollingsworth, Martin Leutbecher, and David Richardson (EMCWF), as well as, Carolyn Reynolds, and Ron Gelaro (NRL) are greatly acknowledged.

Discussions with Brian Farrell (Harvard University) and Petros Iaonnou (Athens University) provided a deeper insight into the art of Kalman filtering.
CURRICULUM VITAE
Alexander Beck

Personal data and basic education

Date and place of birth:

- 5 July 1974, Vienna, Austria; Austrian citizenship

Family:

- father: Dr. Franz Beck–Schmied
- mother: Eveline Beck

Basic education:

- 1980 – 1984: primary school
- 13 June 1992: high school leaving examination passed with distinction

Mandatory military service:

- January 2001 – September 2001: Langenlebarn Airfield, Air defense

Academic education

- October 1992 – May 1997: study of meteorology at the University of Vienna
- collaboration with the Institute for Meteorology and Geophysics (Department for Theoretical Meteorology) for an FWF–funded research project (P11234–GEO: “Forecast skill prediction for weather prediction models”): June 1996 – May 1999
- diploma thesis: *On the approximation of forecast error covariances in a barotropic model* (in German) under guidance of Dr. Martin Ehrendorfer
- 1 July 1997: degree awarded: “Magister der Naturwissenschaften”
- doctoral study at the Institute for Meteorology and Geophysics at the University of Vienna since October 1997 (title of dissertation: *Data assimilation and error covariances in atmospheric models* )
- collaboration with the Institute for Meteorology and Geophysics for an FWF–funded research project (P13729–GEO: “Variational Data Assimilation and Ensemble Prediction”): July 1999 – September 2002
• since October 2002: collaboration with the Institute for Meteorology and Geophysics for an ÖAW–founded research project (HÖ22–2002 “Scale Invariance and Down-scaling of Precipitation in Alpine Catchments”)

• several working trips to ECMWF between 1997 and 2002

• participation in various international meetings: Reading (UK), The Hague (Netherlands), Moliets (France), and Mt. Bethel (USA)

• collaboration with the Institute of Medical Physics and Biostatistics, University of Veterinary Medicine, Vienna (January 2000) and Institute of Meteorology and Physics, University of Agricultural Sciences, Vienna (December 2000)