BERICHTE
aus dem
**INSTITUT FÜR MEERESKUNDE**
an der
CHRISTIAN-ALBRECHTS-UNIVERSITÄT KIEL

Nr. 310
1999

# An adjoint method for the assimilation of statistical characteristics into eddy-resolving ocean models



von

Armin Köhl



Kopien dieser Arbeit können bezogen werden:
Institut für Meereskunde an der Universität Kiel
Abt. Theoretische Ozeanographie
Düsternbrooker Weg 20
24105 Kiel, -FRG-

ISSS 0341 - 8561

## Abstract

The work investigates perspectives of the parameter estimation problem with the adjoint method in eddy resolving models. Sensitivity to initial conditions resulting from the chaotic nature of this type of models limits the direct application of the adjoint method to the forecast range. Beyond this range an increasing number of secondary minima of the cost function emerges and prevents convergence of this method. In the framework of the Lorenz model it is shown that averaged quantities are suitable for describing invariant properties of the attractor and that secondary minima are for this type of data transformed into stochastic deviations.

An adjoint method suitable for the assimilation of statistical characteristics of data and applicable on time scales longer than the forecast range is presented. The approach assumes a larger predictability for planetary scales which are here defined by spatial and temporal averaging. The adjoint to a prognostic model for statistical moments is invented for calculating cost function gradients that ignore the fine structure resulting from secondary minima. Coarse resolution versions of eddy resolving models are used for this purpose. Identical twin experiments were performed with a quasigeostrophic model to evaluate the performance and limitations of this approach for improving models by estimating parameters. The wind stress curl is estimated from simulated mean stream function and sea surface height variability. Even very simple closure schemes for second order moments are shown to give reasonable estimations of gradients that perform efficiently in minimizing cost functions.

The method is applied in the second part to the $1/3^{o}$-CME-model for the assimilation of SSH variance derived from satellite altimeter data from TOPEX/POSEIDON und ERS1 in association with climatological data to estimate the underlying mean circulation. A parametrization of SSH variance for the adjoint equations is derived from the parametrization approach of Green (1970) and Stone (1972). On times scales of one year an almost consistent state with the altimetric and climatological date is achieved by estimating initial conditions for temperature and salinity. The assimilation of SSH variance data introduces complementary informations about the main frontal structures consistent with climatological observations. The estimated state is found to be unstable and to return back to the first guess quasi equilibrium state for longer integration periods.

## Zusammenfassung

Es wird untersucht, inwieweit die adjungierte Methode zur Parameterschätzung in wirbelauflösenden Modellen eingesetzt werden kann. Die Sensitivität gegenüber den Anfangsbedingungen, die aus der dieser Art von Modellen unterliegenden chaotischen Dynamik resultiert, begrenzt den Assimilationszeitraum durch die Vorhersagbarkeit. Die Vergrößerung des Zeitraums ist von der Enstehung einer wachsenden Anzahl von Nebenminima begleited, die die Konvergenz dieser Methode verhindern. Am Beispiel des Lorenz-Modells wird gezeigt, daß statistische Momente zur Charakterisierung invarianter Attraktoreigenschaften geeigent scheinen und daß Nebenminima für diese Art von Größen in stochastische Abweichungen der Momente transformiert werden.

Es wird eine auf Zeitspannen größer als die Vorhersagbarkeit anwendbare, adjungierte Methode vorgestellt, die für die Assimilation statistischer Momente geeignet ist. Der Ansatz geht von einer größeren Vorhersagbarkeit für die großräumigen Skalen aus, die hier durch zeitliche und räumliche Mittelung beschrieben sind. Die Methode nutzt ein Adjungiertes zu einem separaten Model der statistischen Momente, um Kostenfunktionsgradienten zu berechnen, die die aus den Nebenminima resultierende Feinstruktur ingnorieren. Grobauflösende Zwillingsmodelle dienen als Approximationen für diese Modelle. Identische Zwillingsexperimente wurden zu Untersuchung der Einsatzmöglichkeit der Methode für Parameterschätzungen mit einem quasigeostrophischen Modell durchgeführt. Der Windantrieb wurde aus simulierten mittleren Stromfunktionswerten bzw. deren Varianz bestimmt. Es zeigt sich, daß im Fall zweiter Momente sehr einfache Schließungsansätze zur Berechung von Gradienten geeignet sind.

Die Methode wird in dem zweiten Teil zu Assimilation von Oberflächenauslenkungsvarianzen aus Satellitenaltimeter-Daten von TOPEX/POSEIDON und ERS1 im Zusammenhang mit klimatologischen Daten in das 1/3°-CME-Modell verwendet, um die zugehörige mittlere Zirkulation zu bestimmen. Die Parametrisierung der Varianzen im Adjungierten wird aus dem Ansatz von Green (1970) und Stone (1972) abgeleitet. Für Integrationszeiträume von einem Jahr kann ein annähernd mit klimatologischen und den Altimeter-Daten konsistenter Zustand durch Bestimmung der Anfangsbedingung für Salz und Temperatur gefunden werden. Dabei können die Variabilitätsdaten zusätzliche, zur Klimatologie komplementäre Informationen über die Struktur der Fronten liefern. Der geschätzte Zustand ist instabil und kehrt für längere Integrationsperioden zu dem Referenzzustand zurück.

# Contents

# Chapter 1

# Introduction

Our knowledge and the understanding of the dynamics of the oceans result in the first place from observations. In general, the number of observations is too sparse for giving a complete picture of the physical processes involved. Theoretical considerations have always been included and both sources of informations, inspiring each other mutually, need to be combined to form a more complete description.

Nowadays, numerical models play a role of growing importance in supplementing informations on the ocean's evolving state. Data assimilation represents in this context a formal method for combining observational data and theoretical concepts formulated as numerical models.

It might be tempting to conclude that models of growing realism may replace partly the need to observe the ocean or, that data from a more complete observational network need no longer be integrated by an assimilation method. But for the present situation it is still on the contrary. Additional data sources such as data from satellites and an enhanced model realism are just at the beginning to be sufficient for using them in context of an assimilation method.

Particularly systematic model errors were and are still one mayor obstacle that hinder a successful application of these methods. Using higher resolu-

tion in numerical models for resolving the transfer due to mesoscale eddies explicitly is one of the lines, that is followed to improve the simulations. However, the presence of mesoscale eddies introduce at the same time non-linear chaotic dynamics into these models.

Since most methods for the data assimilation are based on a linear assumptions they are now found of only limited applicability. From this, the problem arises that important applications are no longer practicable with these type of models.

The present study deals with the extension of a particular method, the adjoint method, in association with statistical moments. In this method, limitations concerning the length of the time interval, from which data may be integrated, is one of the mayor obstacles. The aim is to establish a new assimilation scheme, specially designed to increase this interval.

## 1.1 Limitations and perspectives

In almost all applications, where ocean models are used for the prediction of unobserved quantities, the assumption of a perfect model is crucial. In order for using a model to predict the effects of a change in certain parameters, e.g. a doubling of atmospheric $CO_2$, the model has to be able to simulate the most important features of the present day climatological state (Gates et al., 1990).

One goal of data assimilation is to use observational data for improving model parametrizations and for the evaluation of this condition. Statistical quantities such as the mean state and the variability are a common way to intercompare models and data. Data assimilation in meteorology and oceanography with strong constraint variational inverse methods however, tries to find a particular solution of a dynamical system that best matches the observations in a certain time interval. In this formulation the dependence of the special solution to certain parameters is used to find the best

match by solving a parameter optimization problem. With the development of the adjoint model technique, this method becomes tractable in applications with realistic models (Le Dimet and Talagrand, 1986).

In oceanographic applications of this inverse technique, the idea, that parameters can be determined from a steady oceanic circulation, has so far played an important role for the design of the experiments. In one of the first applications Tziperman and Thacker (1989) determined model parameters such as momentum fluxes and mixing parameters from simulated data in a model with a steady solution. Results from assimilation experiments with low resolution models demonstrate that models with much higher resolution are needed in order to find an oceanic state consistent with real climatological data (Marotzke and Wunsch, 1993; Schiller and Willebrand, 1995; Yu and Malanotte-Rizzoli, 1998).

Applications of the adjoint method in high resolution models came into fashion particularly with the use of satellite altimeter data. Although an extension seems to be straightforward, there are substantial difficulties connected with applications in present-day models. The circulation simulated by eddy resolving models is subject to chaotic dynamics, which restricts the applicability of the adjoint method to very short time ranges. For present ocean models this range is of the order of a few months. A period considerable larger is however needed in order to estimate parameters correctly (Schröter et al., 1993).

In the meteorological literature the initial condition is mainly regarded as a parameter for optimization with the aim to predict the flow for the time that follows the assimilation period. From this point of view it became clear that the assimilation time is limited by the predictability. Tanguay et al. (1995) investigate the ability of the variational assimilation scheme to fill in small scale detail in accord with large scale observational information. They argued that the advantage of this method, which allows to recover unmeasured detail in a dynamically consistent way, is limited by two conflicting requirements. The assimilation period must on the one hand be

of the order of the time scale of growth of synoptic eddies. At the same time, increasing the period beyond this time is accompanied with reduced convergence rates or even divergence of the optimization which starts first in the smaller scales. The reason for this behavior is the emergence of an increasing number of secondary minima of the cost function when increasing the period of assimilation (Li, 1991; Stensrud and Bao, 1992). The length scale dependence of the nonlinear time scale determines an optimal duration or resolution at which to perform assimilation. A way to circumvent this problem and to increase the assimilation time is described by Pires et al. (1996). Their quasi-static variational assimilation algorithm tries to avoid the solution getting trapped in secondary minima by tracking the absolute minimum over progressively longer assimilation periods. Swanson et al. (1998) investigated the effect of this method on the possibility to prolongate the assimilation period. They found that error growth caused by imperfect model settings still limits the period to 3-5 days in quasigeostrophic (QG) atmosphere models.

For linear dynamics the sequential Kalman filter produces the same state at the end of the assimilation period as variational methods, provided the error statistic is perfectly known (Thacker, 1986). In spite of this relation, sequential methods, which are widely and successfully used for the assimilation in high resolution models (Killworth et al., 1999; Fox et al., 1998), do not show obvious problems connected with the finite predictability. Limitations through finite time predictability enter however the Extended Kalman Filter (EKF) along with linearized equations that predict error covariances. For computational reasons most applications of sequential methods do not include a prognostic calculation of the error and therefore avoid the unbounded error variance growth that Evensen (1992) found by implementing the extended Kalman filter in a QG model.

As eddy-permitting and even eddy-resolving models (Bryan and Smith, 1998) are becoming state-of-the-art, and the computational burden involved with the assimilation of observations spanning a time range of more than

one year is no longer prohibitive, the request for an appropriate inverse method capable for improving parameters emerges. Arguments exist for a potential chance of predicting the ocean climate state in presence of a chaotic atmospheric subsystem (Griffies and Bryan, 1997) and this can be explained by regarding subsystems with different characteristic time scales (Boffetta et al., 1998).

Lorenz (1975) introduced the concept of predictability of the second kind that regards a response of statistical properties to a change in forcing. Palmer (1993) regards the climate as the attractor of a nonlinear dynamical system in a quasi-stationary regime. He added a forcing function to the Lorenz (1963) system and found that the probability density function (PDF) changes with the forcing. From this he concludes vice versa that a change in the PDF's will hint the influence of the forcing. The present work will follow in his footsteps searching for an adjoint method that involves only statistical properties such as mean values and variability. The idea is to no longer simulate the corresponding trajectory of the observations, but instead to constrain only statistical properties, in order to extend the limits of the adjoint method beyond the forecast range.

In the next chapter the connection between secondary minima and the exponential growth of adjoint variables is shown. In Chapter 3 the idea for an algorithm is illustrated by regarding the behavior of cost functions that are based on statistical quantities. The method is described in Chapter 4. An application to a QG model is presented in Chapter 5. The assimilation of sea surface height variance and climatological data into a realistic model of the North Atlantic is described in Chapter 6.

# Chapter 2

# Singular Vectors and Multiple Minima

This section presents a review of the theoretical framework connected with the limitations of variational assimilation in chaotic systems. The model is expressed as a coupled set of differential equations

$$\frac{dx}{dt} = f(x,t) \tag{2.1}$$

where $f$ is a nonlinear function of the state vector $x$. For a given reference solution $x(t)$, the evolution of an infinitesimal small initial perturbation $y(t_o)$ superposed on $x(t_o)$ obeys the following tangent linear equation

$$\frac{dy}{dt} = L(x(t))\, y(t), \tag{2.2}$$

where $L(x(t))$ is the Jacobian matrix of $f$ at $x(t)$. The integral of (2.2) from $t_o$ to $t$ yields

$$y(t) = M(x(t_o), t)\, y(t_o). \tag{2.3}$$

The eigenvalues $\nu_i(t)$ and eigenvectors $\zeta_i(t)$ of $M(x(t_o), t)$ describe the normal mode stability of the system. In general the advection term in ocean and atmosphere models makes $M(x)$ non-normal, that is $M^+M \neq MM^+$. The growth of error

$$\|y(t)\|^2 = \|My(t_o)\|^2 = \langle y(t_o), M^+My(t_o)\rangle \tag{2.4}$$

involves the adjoint $M^+$ to the linear propagator $M$. The corresponding eigenvalues $\mu_i(t)$ and eigenvectors $\eta_i(t)$ of the adjoint operator $M^+$ satisfy the biorthogonality condition

$$(\nu_i - \mu_j^*)\langle\zeta_i, \eta_j\rangle = 0. \tag{2.5}$$

The relation shows that eigenvalues of pairs of eigenvectors that are not orthogonal must be complex conjugated. Since $M^+M$ is symmetric, an orthonormal set of eigenvectors $v_i(t)$ and real eigenvalues $\gamma_i(t)$ exist. The lack of orthogonality of the eigenvector set of nonnormal operators may lead for short times to a transient growth even though all normal modes are damped (Lacarra and Talagrand, 1988; Farrell and Ioannou, 1996). In the limit of infinite integration times

$$\lim_{t\to\infty} \gamma_i(i) = \lim_{t\to\infty} \frac{1}{t} \log \|Mv_i(t_o)\| \tag{2.6}$$

the definition of the Lyapunov exponents is retrieved from (2.4) that gives a quantitative measure of chaos in dynamical systems (Oseledec, 1968). A positive value of $\gamma_i$ indicates an exponential separation of two nearby trajectories which is a characteristic feature of a chaotic system. In the limit $t \to \infty$ the error growth is dominated by the eigenmode associated with the eigenvalue $\nu_i(t)$ with maximum real part, the most unstable mode, which in this limit corresponds to the largest Lyapunov exponent. The adjoint eigenmodes are closely related to the sensitivity analysis the adjoint method provides. The mode with the largest real part of the corresponding eigenvalue is the most singular vector at initial time. That means if the initial condition is regarded as a parameter, this mode describes in the limit of very large periods the most sensitive portion with respect to observations at time $t$.

This can be illustrated in terms of the adjoint assimilation formalism by defining a cost function that measures a single component of x at time t

$$J_x = \frac{1}{2}(x_i(t) - x_i^{obs})^2 \tag{2.7}$$

and the Lagrangian with the dynamical equation (2.1) formulated as strong constraint

$$\mathcal{L}(x_o) = J_x + \int_{t_o}^{t} dt' \lambda(t')(\dot{x}(t') - f(x(t'))). \tag{2.8}$$

The Euler Lagrange equations then read

$$\lambda_j(t) = (x_j(t) - x_j^{obs}) \tag{2.9}$$

$$\dot{\lambda}(t') = -L^+(x(t'))\,\lambda(t') \tag{2.10}$$

From the gradient of the cost function with respect to the initial condition

$$(x_i(t) - x_i^{obs})\frac{\partial x_i(t)}{\partial x_j(t_o)} = \frac{\partial \mathcal{L}(x_o)}{\partial x_j(t_o)} = \lambda_j(t_o) = M_{ji}^{+}(x_i(t) - x_i^{obs})$$

the sensitivity

$$\frac{\partial x_i(t)}{\partial x_j(t_o)} = M_{ji}^{+} \tag{2.11}$$

can be calculated.

In the limit of long time periods the correspondence between the largest Lyapunov exponent and the real part of the dominant eigenvalue of the modal growth also holds because of the biorthogonality relation for the adjoint eigenvectors. It follows that the adjoint state vector, the Lagrangian parameter $\lambda$, will finally be dominated by the most unstable eigenvector of the adjoint propagator and therefore show exponential growth.

Since $\lambda(t)$ mediates the gradient of the cost function with respect to any parameters, the exponential growth of $\lambda(t)$ is transferred to an exponential growth of any gradient. Note that even in cases where the limit of validity of linearization is by far exceeded, the validity of the gradient obtained from tangent linear adjoint equations by the variational formulation still holds. In physical systems state variables and cost functions are normally bound by an upper limit. It follows that infinitely growing gradients are necessarily accompanied by an infinitely growing number of secondary minima. Vice versa, the known fact that an infinitely growing number of secondary minima emerges when the assimilation period is increased results in an infinitely growing sensitivity of the model to parameters (Pires et al., 1996).

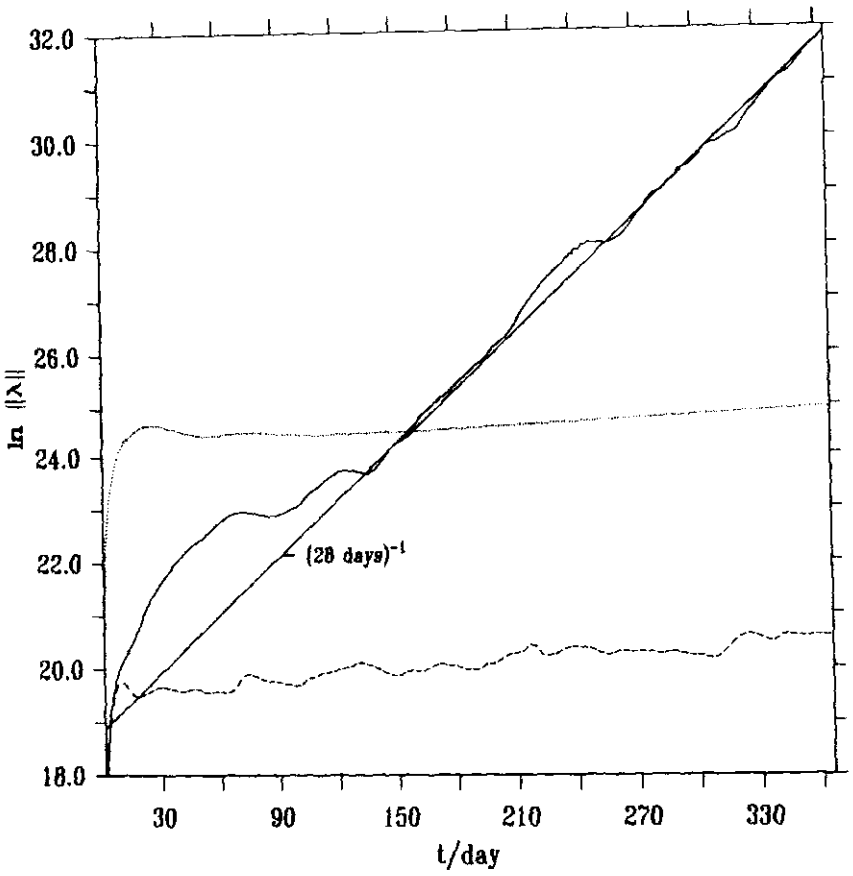


Figure 2.1: *Temporal evolution of the norm of Lagrangian variables for the standard adjoint of an 1/6° QG model described in detail in Chapter 5 (solid line). The dashed curve is obtained when only the advection of model data differences is retained in the adjoint and the dotted curve corresponds to applying the formulation (4.4)-(4.6) to the QG model.*

Figure 2.1 shows the evolution of the adjoint variables of an 1/6° QG model described in Section 5. Two different regimes can be distinguished. The period of 3 months separates the regime of linear growth form exponential growth thereafter, when the amplitudes of the unstable modes have reached the amplitude of linear growing background. For this model an eddy-turnover time scale $\tau$ = 30 days connected with a forecast range of about 2-3 $\tau$ can be calculated. The limits of variational assimilations are

clearly visible when regarding linearized equations.

In the limit of long integration times, sensitivities estimated by the adjoint method only reflect the chaotic nature of the model and are different to sensitivities that were derived from finite perturbations of model parameters in sensitivity studies. However, the chaotic nature of high resolution models blurs also the assessment of finite perturbations in the way that a reliable influence of the parameter is only detectable if the perturbation is large enough. Otherwise it is impossible to separate the effect of changing the attractor from a macroscopic change of the trajectory on a nearly unchanged attractor. Small perturbations cause a macroscopic change in the trajectory but in general a smooth change of the underlying dynamics expressed by the attractor (Eckmann and Ruelle, 1981). An exception from this are bifurcation points, where the topological nature of the attractor may change when the parameter crosses certain critical values.

The same argument can be derived for the response of PDF's to parameter changes. Adding a small stochastic forcing that may represent unresolved subgrid scale interactions to the nonlinear system allows one to describe the evolution of its PDF by the Fokker-Planck equation, a linear advection-diffusion equation defined in the phase space of the original equation. The solution depends smoothly on the parameters as long as the coefficients of the original differential equation depend smoothly on the parameters (Liptser and Shiryayev, 1974).

These observations describe the focal point of the limitations of the adjoint method in applications with chaotic models and form the basis for the experiments described in the following chapter.

# Chapter 3

# Lorenz Model

In order to extend the variational method beyond the forecast range, it is clearly not sufficient to just define the cost function on basis of statistical quantities, as long as the original formulation of the adjoint method is used to find a special trajectory, that optimally represents the statistical constraints. This follows immediately from the considerations of the last chapter. Starting with the standard formalism and regarding the cost function will illustrate the idea for an alternative approach.

The Lorenz (1963) model is a widely referenced chaotic system. It approximates the Rayleight-Bénard convection problem by a spectral truncation of the Boussinesq equations,

$$\dot{x} = -\sigma x + \sigma y, \tag{3.1}$$

$$\dot{y} = \rho x - y - xz, \tag{3.2}$$

$$\dot{z} = -\beta z + xy, \tag{3.3}$$

where $\sigma$ is the Prandtl number, $\rho$ is a normalized Rayleight number and $\beta$ is the domain aspect ratio. A standard forth order Runge-Kutta scheme with a stepsize $\Delta t = 0.01$ is used for integration. To characterize the attractor of this system a cost function based on the mean position

$$\bar{x} = \frac{1}{t}\int_o^t x(t')dt' \tag{3.4}$$

is chosen

$$J = (\bar{x} - \bar{x}^o)^2 + (\bar{y} - \bar{y}^o)^2 + (\bar{z} - \bar{z}^o)^2, \tag{3.5}$$

where the reference values $\bar{x}^o$, $\bar{y}^o$, $\bar{z}^o$ are taken from experiments with the classical parameters: $\sigma = 10.0$, $\rho = 28.0$, and $\beta = 8/3$. For $\rho > \rho_H = 24.74$ the system is chaotic with Lyapunov exponents of (0.93,0,-14.60) for these parameters (Nese et al., 1987). The trajectory orbits the two unstable fix-

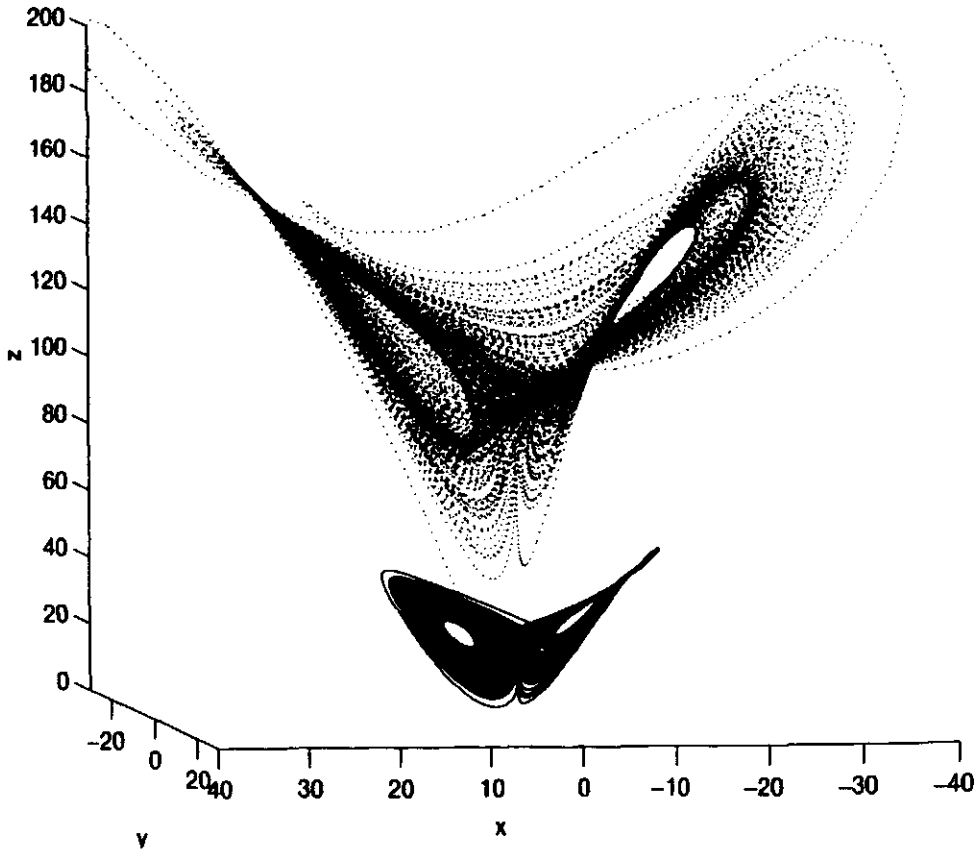


Figure 3.1: *Graphs of the Lorenz attractor for* $\rho = 28$ *(solid) and* $\rho = 120$ *(dotted).*

points, z=$\rho - 1$, $x = y = \pm\sqrt{\rho - 1}$ and is moved when the parameter changes. The movement of the attractor illustrated in Figure 3.1 is accompanied by a magnification of the amplitudes of the oscillations. The evolution of the cost function for different values of $\rho$ and an increasing

integration period $t$ is shown in Figure 3.2. For small intervals $t$ the cost function seems to have an unique minimum and is significantly nonparabolic due to the nonlinearity of the system. Prolonging the interval is accompanied with the emergence of an increasing number of secondary minima which is generic in chaotic systems.

The threshold for the emergence of secondary minima depends on the width of the parameter interval in regard, a detail of which the quasi-static variational assimilation of Pires et al. (1996) takes advantage. In the limit of very long integration times the cost function approaches a limit shape which has a parabolic form, that reflects an almost linear dependence of the position of the attractor on the parameter $\rho$. Secondary minima are transformed into stochastic deviations. A detailed investigation of the distribution of the mean and cost function values from a small interval around $\rho = 31$ shown in Figure 3.3 reveals a Gaussian distribution for the mean values. The gradient calculated by the standard variational approach correctly describes the topology of the cost function but is of very little help in finding the absolute minimum. Being almost vertical, it will only lead to the neighboring minimum. Sensitivites calculated with the adjoint method give no information on the dependence of the mean position on the parameter. In contrast to this, finite parameter perturbations as applied by investigating macroscopic sensitivity approximately describe the correct dependency.

A sensible algorithm must be able to calculate a "mean" gradient that disregards the stochastic deviations of the limit curve of Figure 3.2. This curve may also be described by the maximum values or moments of the PDFs in Figure 3.3. A reasonable gradient could thus be estimated from finite differences, if a proper parameter difference is chosen, but would be very expensive in high dimensional parameter spaces. An estimation by the adjoint method is possible if it is based on a prognostic equation for the moments that describes the mean values independently of special realizations of the trajectory.

To explain the stochastic behavior of moments calculated from chaotic sys-

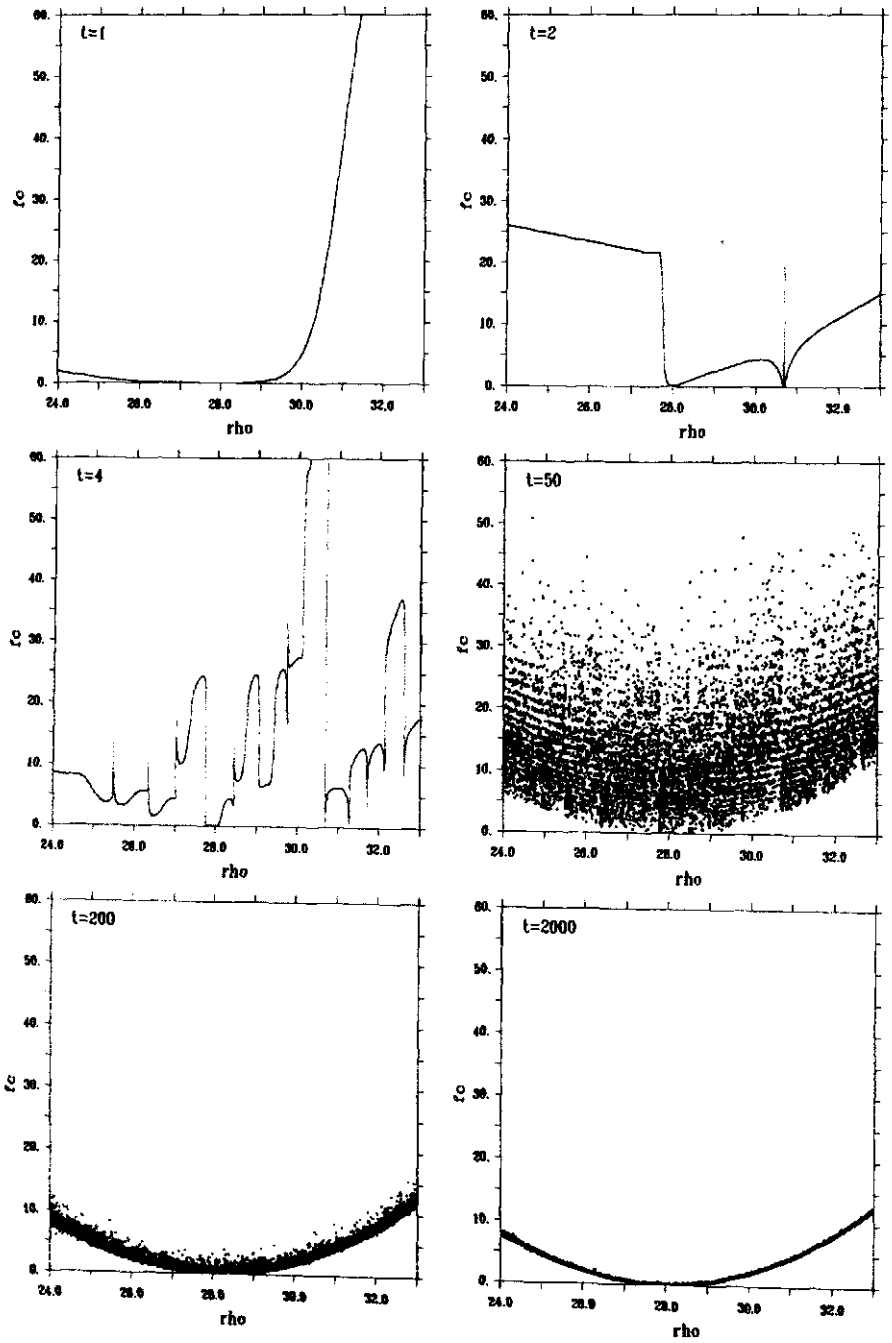


Figure 3.2: *Evolution of the cost function (3.5) from the Lorenz model in dependence of the Rayleigh number $\rho$ when increasing the integration period $t$.*

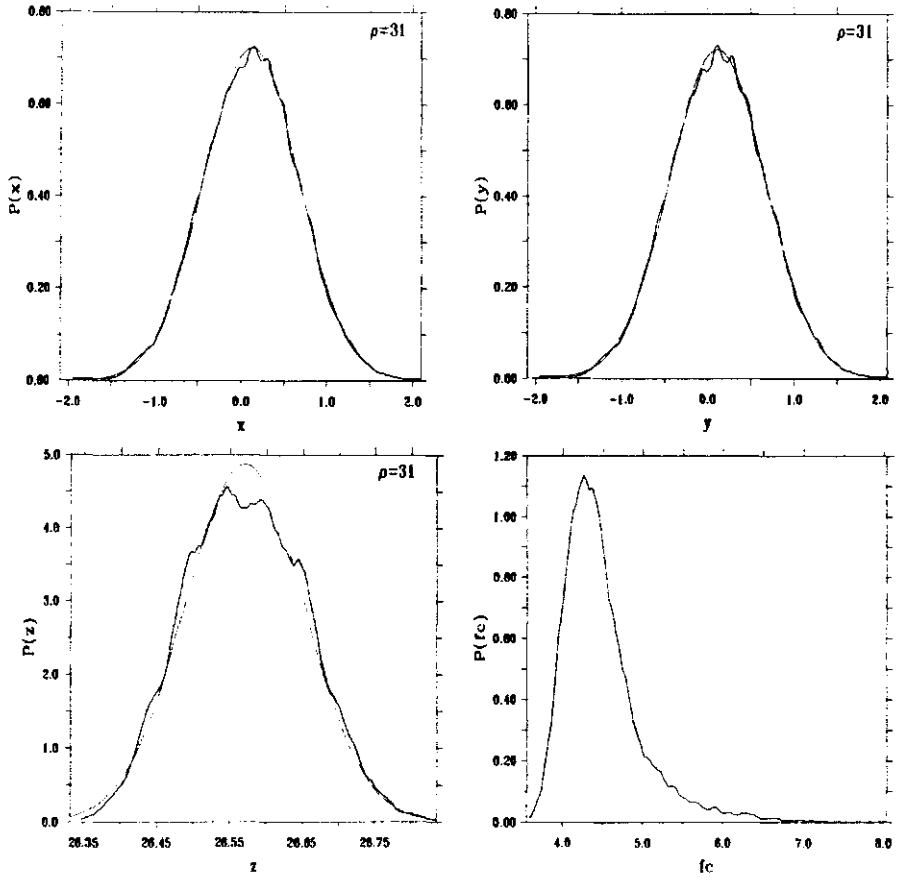


Figure 3.3: *Probability density functions describing the distribution of the mean values $\bar{x}, \bar{y}, \bar{z}$ and the cost function $fc$ of the Lorenz model together with Gaussian approximations. The ensemble of realizations is constructed by varying $\rho$ slightly around a mean value of $\rho = 31$.*

tems in the general framework, we return to the notation of Chapter 2. The system (2.1) which is regarded as statistically stationary is decomposed into a slow mean component, $\bar{x}$, which might be an ensemble or a temporal mean, and the deviation, $\delta x$, which describes the fast and transient part

$$x = \bar{x} + \delta x. \tag{3.6}$$

The equation then reads

$$\dot{\bar{x}} \;=\; \overline{f(\bar{x} + \delta x)} = \bar{f}(\bar{x}, \delta x) \tag{3.7}$$

$$\delta\dot{x} = f(x) - \overline{f(\bar{x}+\delta x)} = \delta f(\bar{x}, \delta x). \tag{3.8}$$

For any reasonable physical system $x$ can be assumed to be bounded, $\|x\| < K$. The relation

$$\|\dot{\bar{x}}\| < \frac{2K}{T} \tag{3.9}$$

motivates a scaling for the time, $t \to t/T$, with the averaging time interval $T$. After scaling the rate of change increases in (3.8) with the prolongation of the averaging interval $T$.

Following Hasselmann (1976) the slow component can then again be divided into a "mean" component $\langle\bar{x}\rangle$ and the deviation $\bar{x}'$. The expectation value $\langle\cdots\rangle$ is regarded as a very long time mean or as a solution of the equation for moments $\hat{x}$

$$\frac{d\hat{x}}{dt} = \hat{f}(\hat{x}), \tag{3.10}$$

which would involve a closure hypothesis for expressing $\delta x$ in terms of $\bar{x}$. The evolution equation for the deviation then yields

$$\frac{d\bar{x}'}{dt} = \bar{f}(\bar{x}, \delta x) - \hat{f} = f'. \tag{3.11}$$

According to Hasselmann (1976) $f'$ can be assumed to represent a stationary random process. It appears in correspondence to the behavior of the cost function in case of the Lorenz model that the transient components act on mean values as a Gaussian stochastic process. In the limit of long times the large number of irregular and statistically independent short time oscillations were superposed which enables the applicability of the central limit theorem for the distribution of the mean values.

# Chapter 4

# The Method

## 4.1 Different approaches

A recent approach to circumvent limitations in strongly nonlinear systems is to define sequential assimilation procedures in terms of PDF's. Miller et al. (1994) and Evensen (1994) derived a generalization of the EKF based on Monte-Carlo estimates of covariance statistics. This method is found to work well in QG models by Evensen and van Leeuwen (1996). It can be regarded as a second-order moment approximation of the more general method of Miller et al. (1999) who applied the Bayes theorem directly to PDF's. The evolution of the PDF's are calculated by the Fokker Planck equation or in more complex models estimated by a Monte-Carlo approach. A description for statistical moments can be derived from the PDF evolving according to the Fokker-Planck equation, when stochastic forcing is added. A possible algorithm would thus involve the adjoint to the Fokker-Planck equation. Since in higher dimensional systems a direct method is impracticable, PDF's have to be estimated by a Monte-Carlo approach. This would imply that a generating Langevin system has to be found for the adjoint to the Fokker-Planck equation. A somewhat similar approach, among which an ensemble average of cost function gradients of short time slices of one long integration period filters the effect of the secondary minima, was found

by Lea et al. (1999) to work well in the Lorenz system with an intermediate time scale of t=0.44 to estimate the macroscopic sensitivity. If the slices cover longer periods the ensemble size must be increased very fast to capture the increasing steepness of the gradients due to the increasing number of secondary minima. This method would probably not be applicable on annual time or longer scales for ocean models since slicing the trajectory does not allow the model-data difference information to propagate in the adjoint far enough to reach the origin causing the difference.

The unbounded growth of the adjoint variables results in ocean and atmosphere models from the advection term

$$u_t + uu_x. \tag{4.1}$$

The corresponding term

$$\lambda_t + u\lambda_x - u_x\lambda \tag{4.2}$$

of the adjoint equation retrieves the advection of the model-data differences along the streamlines of the forward model plus an extra term. This term may cause exponential growth with a time scale in order of the nonlinear eddy-turnover time $T \sim L/U$, where $U$ is a typical velocity and $L$ a typical length scale of the motion. The growth rate of 28 days for the norm of the adjoint variables from the QG model shown in Figure 2.1 corresponds fairly well with an estimation of the turnover time $T = 30$ days.

Evensen (1992) shows that the growth of the errors estimated by the Kalman filter is caused by an analogous term in his QG model. This term was neglected to avoid unbounded growth. The errors are then shown to decrease in time due to the neglect of the inherent dynamical instability. Assimilation experiments are then shown to perform better than with the original formulation. Likewise, a naive way for trying to obtain a description for the "mean" gradient proposed in Section 3 would be to neglect the instability terms and to keep just the advection of the misfit information. This was tested with the adjoint of the QG model. The exponential growth of the adjoint variable is evaded as it is visible from the Lagrangian variables in

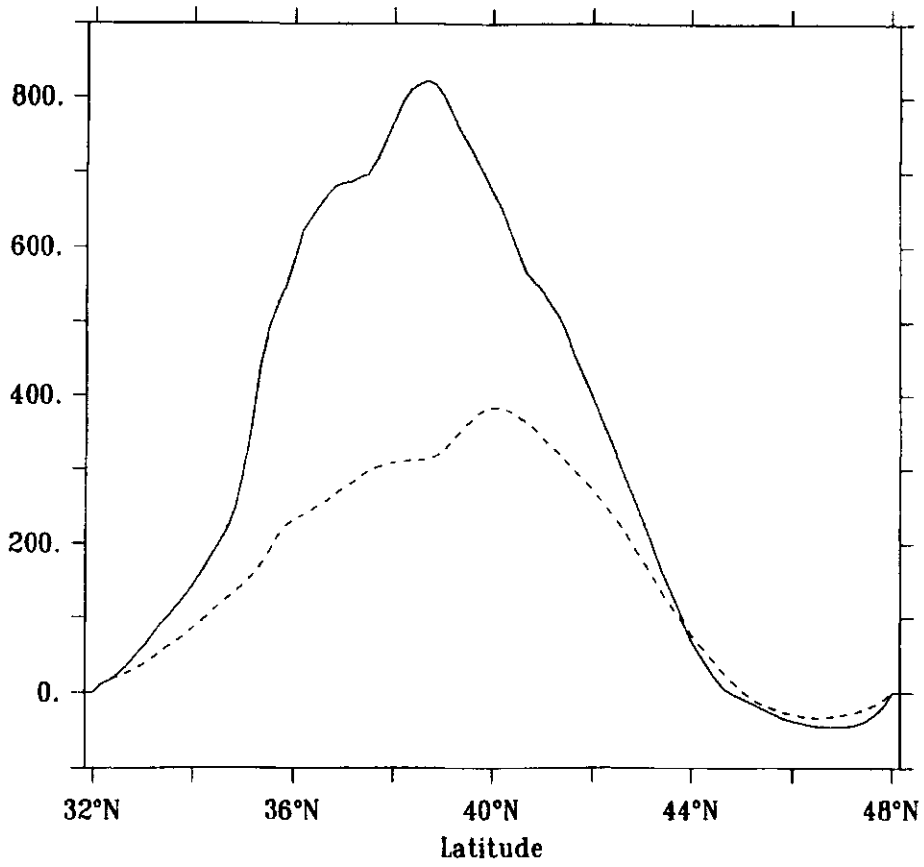


Figure 4.1: *Cost function gradient with respect to the zonal wind stress curl for an integration time of 30 days from standard adjoint formalism (solid) and from a formulation with pure advection of the model-data-misfit information (dashed).*

Figure 2.1. However, Figure 4.1 shows that for an integration period of one month, which is clearly within the forecast range, the amplitude of the gradient is considerably underestimated. Although the shape of the two gradients look similar, this remains not true when the period of assimilation is increased. It is thus not possible to obtain a useful gradient by this approach. An analog result for the Kalman filter was given by Evensen (1994). The time evolution of the error covariances calculated with the reduced formulation were also considerably underestimated in comparison to

the errors calculated from a Monte Carlo forecast.

## 4.2 Description of the method

Since statistical moments are used as potential elements in the cost function, it is sufficient to regard only equations that describe the evolution of moments as they can be obtained by Reynolds decomposition. The moment equations are then used instead of the high resolution model as a strong constraint in the variational formulation. This means, the high resolution model is no longer needed. From the nonlinearity of the evolution equation a closure problem arises and it becomes necessary to express higher order correlations of the transient components in terms of lower order moments.

As a first step, the standard way is followed and the closure problem is handled by introducing a model twin on a coarser grid

$$\frac{dX}{dt} = F(X) \tag{4.3}$$

with an increased eddy viscosity that mimics transient processes with short term predictability. This kind of parametrization usually has severe shortcomings in regions where nonlinear processes are dominant (see also Figure 5.1). Using this model for the calculation of the moments would reduce the method to the assimilation in coarse resolution models. Any improvement in association with the use of a higher resolution is lost and there is no chance for extending the method to higher order moments.

Judged on basis of their results, coarse resolution models can not be regarded as sufficient approximations to a model for the moments. On the other hand, state-of-the-art realizations of any moment, $\overline{x^n}$, are available from the solution of high resolution models. It is thus possible to follow a mixed approach. Since only tangent linear equations are needed in the variational formalism, an expansion of the moment model at the best available approximation for the moments is possible without knowing a solution from the moment model. Courtier et al. (1994) invented an incremental

formulation of the variational assimilation to reduce the cost of the method for an operational implementation. They expand the forward model at the first guess solution and use a simplification of the first order tangent linear term for the optimization steps afterwards and found that the algorithm works well.

Following a similar approach, the coarse resolution model twin is expanded at a realization of the moment that is calculated from the solution of the high resolution model and averaged to the coarse grid. The resulting tangent linear equations are employed to approximate the first order of the moment model and are then used in the variational formulation as a strong constraint for the statistical moments. This means, the high resolution model is still applied for the calculation of approximations of the moments.

Introducing a parameter $\alpha$ of the high resolution forward model and the adjoint operator of the coarse resolution twin model, $F^+(X,\alpha)$, the scheme of the assimilation algorithm then reads:
minimize

$$J(\alpha) = \frac{1}{2}(\alpha - \alpha_b)^T B^{-1}(\alpha - \alpha_b) + \frac{1}{2}(H\bar{x} - y)^T O^{-1}(H\bar{x} - y) \qquad (4.4)$$

with the high resolution forward model

$$\frac{dx}{dt} = f(x, \alpha, t), \qquad (4.5)$$

and the adjoint equations

$$\frac{d\lambda}{dt} = F^+(\bar{x}, \bar{\alpha})\,\lambda + O^{-1}(H\bar{x} - y), \qquad (4.6)$$

where $\bar{x}$ is the time mean calculated from the solution of (4.5). The moment is after a spatial averaged to the coarse grid inserted into the adjoint equation (4.6). $O$ is the error covariance of the observations $y$, $B$ the error covariance of the a priori information, $\alpha_b$, of the parameter $\alpha$ and $H$ the observation operator.

This scheme allows the adjoint variables to be propagated on the mean stream lines of the forward model, although the transient part, e.g. the

eddy flux terms, are only represented by simple parametrizations. The approximation will cause the Lagrangian variable and therefore the gradients to be only approximative. But as in the outer loops in the application of the incremental method by Rabier et al. (1998), no approximations are made within the cost function and the forward calculation. Since the gradient is only a means for finding the minimum, errors in the Lagrangian variables will possibly reduce the performance with an eventual failure in convergence of the method. The shape of the cost function and the position of the minimum will not be affected.

Figure 2.1 shows that the Lagrangian variable does not increase exponentially in this formulation, but approaches an asymptotic value, which is much higher than in the above case where the instability mechanism is removed in the adjoint.

Since (4.6) are linear autonomous equations for calculation of the Lagrangian variable $\lambda$, the system represents a simple method for the calculation of the stationary solution. The adjoint step therefore can be speeded up by using a more efficient scheme such as a Gauss-Seidel solver.

The error covariance $O$ is a function of the measurement errors but represents in the same way limited representativeness of the observations due to variability of the physical system. Statistical moments appear in high resolution models as stochastic quantities. Under the prerequisite that the modeled variance of the statistical moment is a sensible approximation for the variance of the data analog and the measurement errors are negligible in comparison to this variance, an estimation of the covariance $O$ is possible from an ensemble of model integrations.

Calculating the Lagrangian variables on a coarser grid implies an interpolation step. Since the coarse grid represents only a subspace of the original system the method can only estimate the part of the parameters that projects onto this subspace. This is the coarse resolution and time averaged part. The computationally more economic way is to perform the descent

algorithm on the coarse grid. In order to keep the fine scale complement unchanged, this part has to be isolated and added after the descent step. In all numerical experiments presented later we follow the simpler way and interpolate the gradient to the high resolution grid to avoid a decomposition.

The method is so far only applicable for mean values. The extension of this method to higher order moments is not straightforward. A simple approach that introduces additional closure schemes into the adjoint equations is presented in Section 5.4.

# Chapter 5

# Application to the QG Model

In order to investigate if the ideas developed from the Lorenz model hold in a more complex framework and to evaluate the potential performance of the method in respect to the assimilation of real data a QG model is regarded in the following. As suggested by Schröter et al. (1993), the new method will be first tested in an identical twin configuration where the perfect model assumption holds and results can be judged by the true fields and parameters. The general strategy of the experiments is to reconstruct the zonal wind stress pattern from simulated statistical moments such as mean stream function and stream function variance expressed as sea surface high (SSH) variance. The model is considered to be in a statistically stationary state and the realizations of the moments are considered as being characteristic for the dynamical behavior. That means, parameters can in principle be recovered from investigating macroscopic sensitivity.

## 5.1 Model description

The three layer QG model is based on the Holland (1978) model and basically identical to that described in Vogeler and Schröter (1995). A limited area double gyre configuration is set up on the $\beta$-plane with the Coriolis

parameter at central latitude of $40^o$ to mimic a simple model for midlatitude jets. The resolution is $1/6^o$ in the zonal and meridional direction and the area extends meridionally from $32^o$ N to $48^o$ N. The layer thicknesses are from top to bottom 300m, 700m and 4000m, respectively and reduced gravities at the layer interfaces are $0.0357ms^{-2}$ and $0.0162ms^{-2}$. Friction coefficients are chosen as $10^{-7}s^{-1}$ for the bottom friction and $10^2m^2s^{-1}$ for harmonic lateral friction. The reference experiment is forced with zonal wind stress given by $\tau = \tau_o \cos(2\pi y/L_y)$ with $\tau_o = 10^{-4}m^2s^{-2}$. A flat bottom is prescribed and free slip conditions are applied at the closed boundaries. The equilibrium solution is a statistical stationary double gyre with a maximum zonal velocity of about $1m/s$ and a maximum SSH rms variability of about $60cm$ (Figure 5.1).

## 5.2 Adjoint model

The prognostic model used for the construction of the adjoint for stream function moments is identical to that described above except of a decrease in resolution to $2/3^o$ and an increase of the lateral friction to $10^4m^2s^{-1}$. The solution of this configuration is a stationary two gyre system with a reduced penetration scale and velocity of the zonal jet (Figure 5.1). The adjoint to the QG model is described in Moore (1991) and Schröter et al. (1993). As descent algorithm the quasi-Newton routine, MINIM, is chosen which is based on the Davidon-Fletcher algorithm and which was successfully applied by Jung et al. (1998). The discretisation in time was changed to a finite difference of the adjoint analogue to Sirkes and Tziperman (1997) in order to suppress the computational mode. Computing the stationary solution of (4.6) by a Gauss-Seidel iteration as described in Section 4.2 lead to gradients that were found to perform more efficient in the optimization and were used throughout the following described experiments.

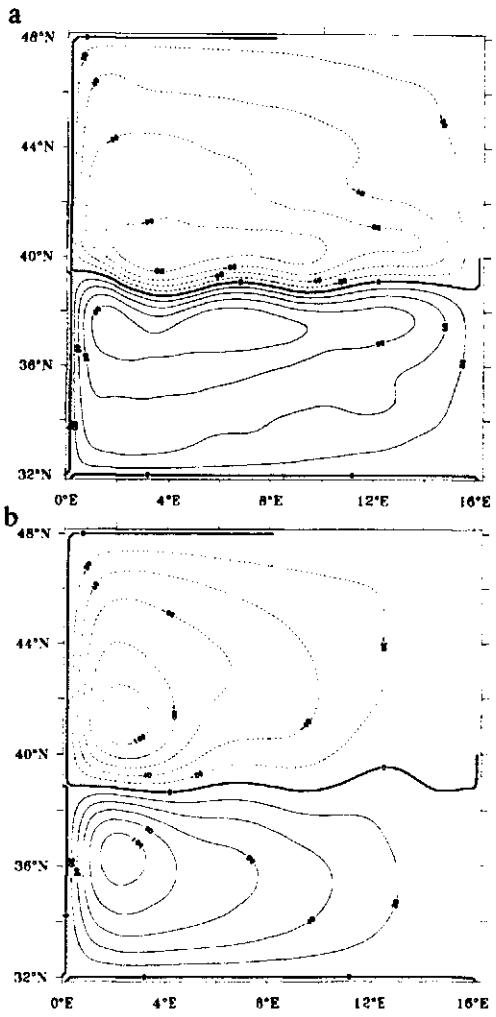


Figure 5.1: *Annual mean upper layer stream function (in Sv) from the reference (a) and a coarse resolution experiment (b).*

## 5.3 Assimilation of stream function data

### 5.3.1 Behavior of the cost function

To measure the least-square distance between the actual upper layer mean stream function $\bar{\psi}_1$ and the simulated observations a quadratic cost function

$$J(\psi(\tau)) = \frac{1}{N} \sum_i \frac{\left(\bar{\psi_{i,1}}(\tau) - \bar{\psi_{i,1}}(\tau_o)\right)^2}{\epsilon^2} \tag{5.1}$$

is introduced where $i$ counts the number of horizontal grid points $N$. All experiments are started from the same initial condition, which is the state after 20 years of forward integration with reference wind forcing. Mean values were derived from an integration period of one year. The control parameter is the curl of the zonal wind stress $\tau$.

The error covariance matrix is assumed to be diagonal and spatially homogeneous. $\epsilon^2$ is then defined as the spatial average of variance of an ensemble of mean stream functions. The ensemble members are derived from varying the reference wind stress pattern by an amount of less than 5%. Due to this choice of the error covariance the cost function scales in a way that it shows values of about one in the vicinity of the minimum. Since the moments of high resolution models are regarded as stochastic quantities no additional noise was added to the simulated observation.

For comparing with the results from the Lorenz model, a section through the cost function is made. The wind stress patterns that correspond to this section are obtained by decomposing the reference function into discrete wavelet modes (Press et al., 1993) and tuning the amplitude of the fourth mode which causes mainly large scale variations of the wind stress. A decomposition into wavelet modes was chosen to allow for an easy reduction of the number of degrees of freedom, realized in an experiment described below. It was shown by Farge (1992) who also gives an excellent review, that data form turbulence may be reconstructed fairly well by using only a limited number of wavelet modes. The modes depend on a scale and a position parameter. The first two modes describe the mean value of the function. The remainder of the space is classified by a sequence of N sets, n=1..N. Each set contains $2^n$ modes of an identical wavelength proportional to $2^{-n}$, starting for n=1 with the wavelength that corresponds to lenght of the whole interval. The references wind stress is thus described essentially by the third and forth mode. The reference wind stress and the pattern resulting from doubling the forth coefficient is shown in Fig. 5.9.

Figure 5.2 shows the section through the cost function. General character-

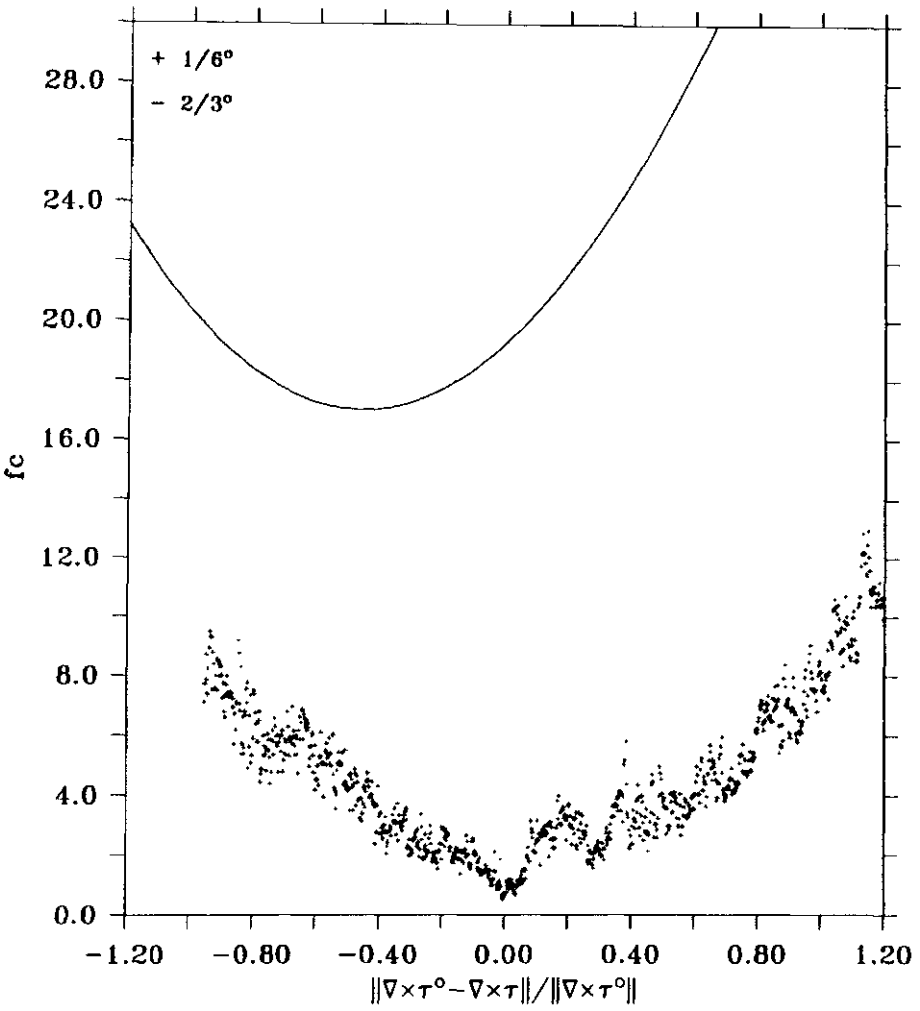


Figure 5.2: *Samples from sections through cost functions based on annual mean upper layer stream functions. Simulated observations are from the high resolution model with the reference wind forcing $\tau^o$. The dependence on the parameter $\tau$, which is effectively the forth mode of a wavelet decomposition, is displayed by the normalized Euclidian difference to $\tau^o$. The crosses correspond to experiments with the high resolution $1/6^o$ forward model. The smooth curve is the costfunction that results from simulations with the low resolution $2/3^o$ model.*

istics of the Lorenz system are retrieved and the curve resembles one with stochastic deviations superposed on a parabolic shape. The smooth black curve in Figure 5.2 describes the costfunction that results from an attempt to simulate the stream function with the low resolution $2/3^o$ model on which the adjoint is based. The large displacement of this curve indicates that no consistent solution can be found by optimizing the parameter. The shifted minimum shows that even the parameters estimated from the minimum are significant in error. This illustrates the general problem of variational assimilation in presence of systematic model errors which project onto the estimated parameters. However, the resemblance of the shape of the cost function to the high resolution analogue suggests that a linearization approach about the mean calculated from the solution of an eddy resolving model may work fairly well.

The general problem in this method emerging from the stochastic structure of the cost function is the sensitivity of the cost function values and its gradient. It becomes clear, that finding the absolute minimum which is a singular point, is an impractical task. The precision up to which a parameter can be relocated, is expected to depend on the noise level. A criterion for terminating the optimization is given by a reduction below the noise level, e.g. a cost function value of one. This was not enforced during most of the following experiments to investigate the order of magnitude of the possible reduction.

### 5.3.2 Assimilation period of one year

A period of one year is not sufficient for the model to equilibrate to changed wind stress. However, the pattern, derived from the difference between the annual mean of the first year after changing the parameter and the reference stream function, is found to be qualitatively analogous to the difference between the corresponding two quasi-stationary states. It is interesting to see how the method performs under this condition, since this experiment

also indicates whether the response of the statistical moments to a change in parameters shows enhanced predictability in comparison to the time scale that is associated with the inherent chaotic dynamics at the mesoscale.

The optimization is started throughout the paper from a control parameter that is obtained by tuning the fourth mode of the reference wind stress. Paths from minimizing the cost function (5.1) by estimating the zonal wind stress are shown in Figure 5.3 in dependence of the normalized Euclidian distance to the reference parameter. The paths generally leave the straight line in wavelet mode space that is marked by the section through the cost function. Cost function values therefore may show considerably higher or lower values for the same Euclidean distance, even if the cost function would be free of noise. The optimization is assessed in the figure simultaneously on basis of the distance in parameter and observation space.

For the red path in Figure 5.3 an ensemble of 10 realizations was used to investigate if a reduction of the noise level results in a higher precision of the parameter estimation. The members where constructed by varying slightly the parameter by an amount of less than 5%. The expected effect on the precision is not captured and the final state is virtually of the same quality as in the experiment on basis of one realization. The Euclidian distance between the final parameters are of the same order as the distance to the reference value.

A nullspace within the parameter space that is unobservable from mean stream function data may account for the remaining difference to the reference parameter. This is investigated by restricting the parameter space to the 4 first modes of the wavelet decomposition. Since the optimization visualized by the green path in Fig. 5.3 ends up with the same distance to the minimum, it is not likely that a nullspace accounts for the limitation. The precisions may mainly be limited by a general error of the gradient due to the approximation necessary within this method or trapping into larger secondary minima (such as the one visible in Fig. 5.3) that where not ignored by the calculation of the gradient with the coarse resolution

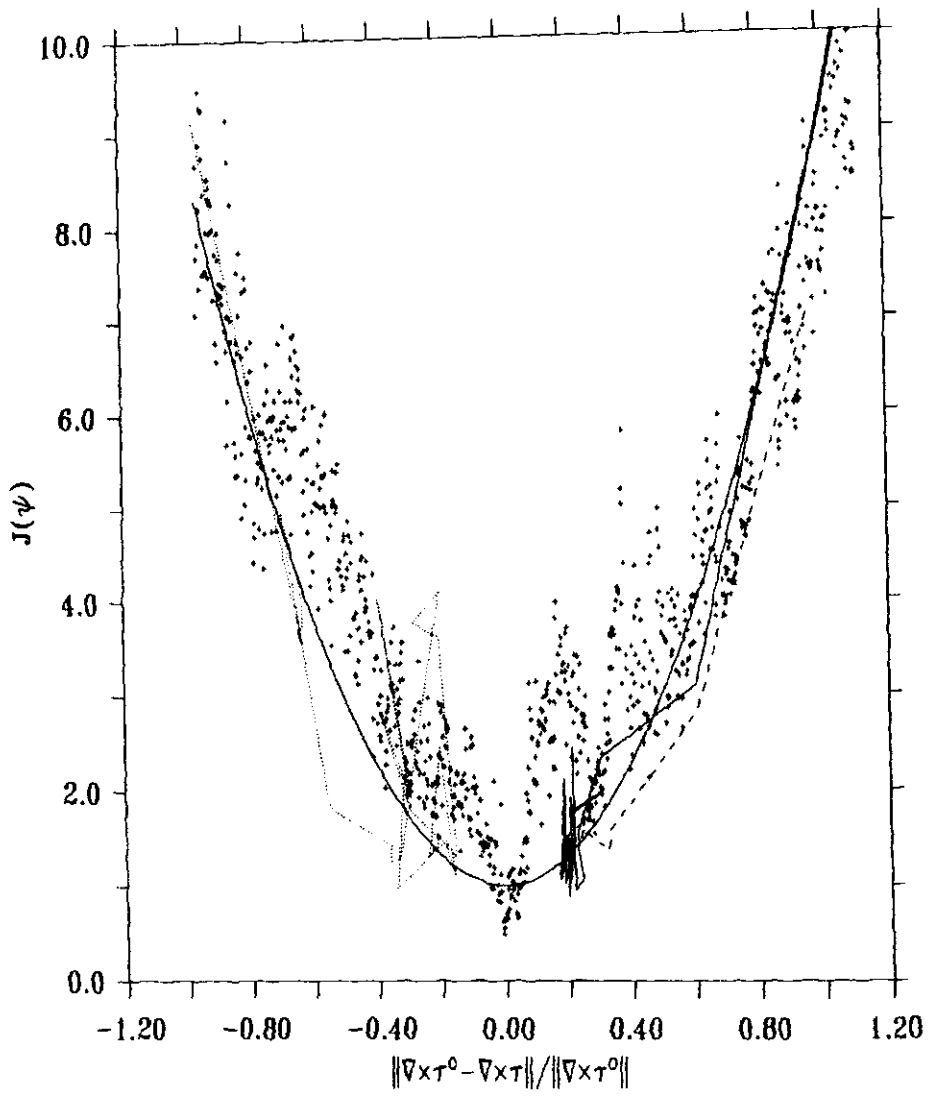


Figure 5.3: *Samples from a section through the cost function based on annual mean upper layer stream functions together with paths from optimizing the zonal wind stress* $\tau$. *The solid path is from an experiment that employs one realization, and an ensemble 10 realizations is used for the dashed path. The parameter space is reduced to the 4 first wavelet modes for the dotted path shown on the left hand side just for clarity. The parabola is used for a posterior error estimate.*

adjoint.

### 5.3.3 Error estimation

The important part in estimating the posterior error covariance of parameters is the Hessian matrix of the cost function. Disregarding the a priori information term in (4.4) the Hessian with respect to the parameter $\alpha$ reads

$$A = \left(\frac{\partial Hx}{\partial \alpha}\right)^T O^{-1} \left(\frac{\partial Hx}{\partial \alpha}\right) + \left(\frac{\partial^2 Hx}{\partial \alpha^2}\right)^T O^{-1}(Hx - y). \qquad (5.2)$$

According to Thacker (1989) the error covariance of $\alpha$ is described by the inverse $A^{-1}$ of the Hessian. An approximation to the inverse of the Hessian can be obtained from the decent algorithm if it employs a quasi-Newton method. Methods for calculating the Hessian as the Davidon-Fletscher algorithm used by MINIM utilize informations from the gradient. For our method it is expected that at the final stage the estimation of the Hessian is corrupted by the variability of the gradient which is then strongly subject to the stochastic nature of the cost function.

However, since paths from all optimizations trace the shape of the section through the cost function, it seems sensible to assume isotropy for the dependence of the cost function on the parameter. This means that the displayed section is assumed to be representative for any other section. A simple approach is then to estimate the Hessian form fitting the shape of the cost function in Figure 5.3 to a parabola. All parabolas drawn together with sections through cost functions are obtained by adjusting the curve by hand under the constraint that the value in the minimum is one. This constraint takes the applied scaling of the cost function into account. By this way, a relative error of 25% is estimated which is in correspondence with the achieved distance of the parameter estimation.

### 5.3.4 Assimilation period of 5 + 5 years

A period of 5 years is chosen to precede the assimilation to allow for equilibration to the changed wind forcing, and the period of assimilation is extended to a further 5 years to reduce stochastic deviations caused by different eddy realizations. The amplitude of the stochastic component of the cost function shown in Figure 5.4 is markly increased comparison to the one year experiment. Parameters with an rms difference to the reference value in the order of norm of the reference value, may correspond to cost function values as low as values from the vicinity of the control parameter. This is because the QG model involves a much wider spectrum of different time scales than the Lorenz model. The cost function values are, in contrast to the one year integration, now subjected to long term variability of the jet stream position.

Despite the fact that a small improvement of the parameter leads to an expected reduction of the cost function, it may still increase, if a realization higher than one from the ensemble mean occurs. Therefore small improvement steps are generally unfavorable. They will obscure the minimization algorithm, because the gradient does not contain informations about the fine structure of the cost function. Usually this is only a problem if it happens during the first steps, where the Hessian is only unsufficiently determined by prior steps. The path from the optimization in Figure 5.4 shows a final cost function value well below the noise level in the vicinity of the minimum. All cost function evaluations are plotted in the figure. Considering the last 10 iterations, it arises that the minimization does not take place through just choosing realizations which have successive lower cost function values by chance. In terms of cost function values the estimated parameter is of the same quality as one, that is much closer to the reference parameter. The final parameter value has a normalized Euclidian distance to the correct parameter of about 27%. The error estimation gives a relative error of 17% which is lower than before, because the noise level of the 1 year period assimilation experiment was very high in the vicinity of the minimum. All

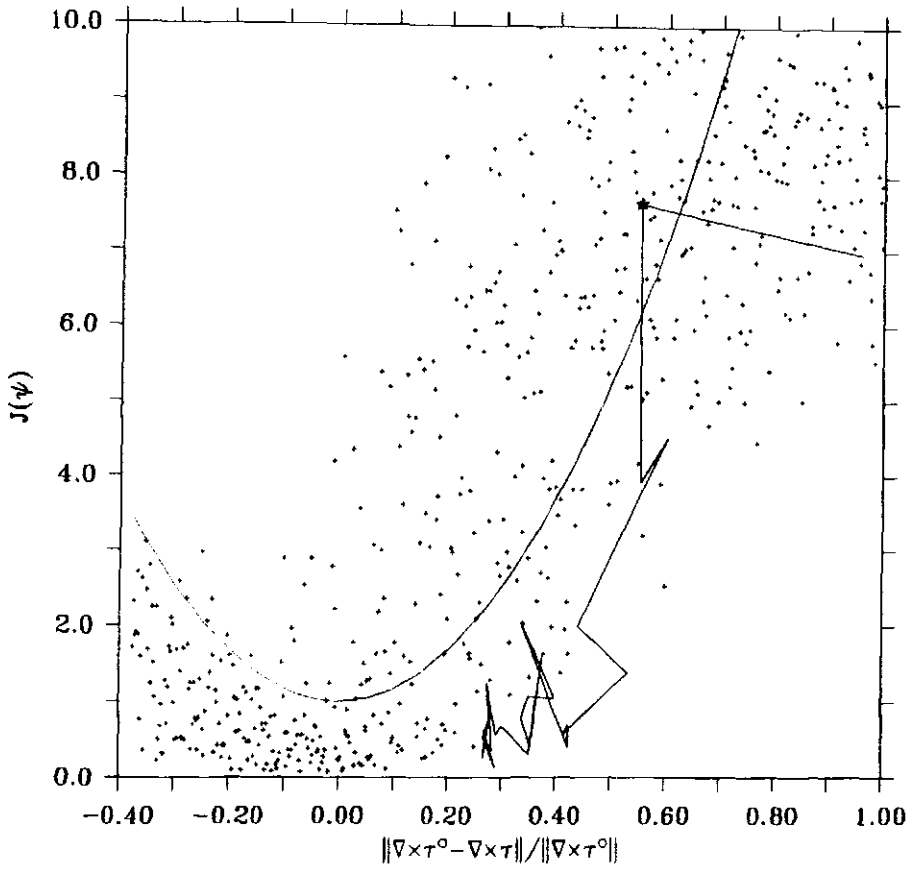


Figure 5.4: *Samples from a section through the cost function based on mean upper layer stream functions derived from an integration period of 5 years together with a path from optimizing the zonal wind stress $\tau$. The star denotes a restart of the descent algorithm. The parabola is fitted to the cost function by hand.*

points on the path of the optimization are significantly below the cost function, showing that the isotropy argument does only hold marginally, which hints at different sensitivities to the observational data in the parameter space.

## 5.4 Assimilation of variance data

Sophisticated closure schemes and models for higher order moments are only available for idealized geometries and forcing functions, e.g. Holloway and Hendershott (1977), and would be impracticable because of their cost. In this section it is shown that even the utilization of very simple "closure" schemes may produce useful gradients for minimization algorithms.

The definition of the cost function based on the stream function variance, expressed as SSH variance $\sigma^2$ through the relation SSH $= \frac{f_o}{g}\psi_1$, is analogous to section 5.3.1

$$J(\sigma^2(\tau)) = \frac{1}{N}\sum_i \frac{\left(\sigma^2_{i,1}(\tau) - \sigma^2_{i,1}(\tau_o)\right)^2}{\epsilon^2_\sigma}, \tag{5.3}$$

and the same strategy is used for estimating the error $\epsilon^2_\sigma$. Different to experiments concerning mean stream function data, it appears that within a period of one year the stream function changes too little for significantly influencing the variance (Figure 5.5). For this reason the same period of 5 years for assimilation and equilibration as in the last section is chosen.

### 5.4.1 "Closure" from velocity

Eddy variability in ocean and atmosphere models is generated due to instability processes that transfer energy from the mean velocity to the transient part. Typical mechanisms are baroclinic and barotropic instability arising from vertical and horizontal shear of the mean velocity. In order to avoid contamination by noise from differentiation, the mean upper layer velocity is used instead of velocity shear to parameterize variability:

$$\sigma^2 = \kappa|\bar{u}|^2 = \kappa([\partial_x\bar{\psi}_1]^2 + [\partial_y\bar{\psi}_1]^2). \tag{5.4}$$

The mean velocity is connected with mean shear since typical horizontal and vertical scales are associated with the zonal jet stream. The proportionality to the square of the velocity was assumed in analogy to the parametrization

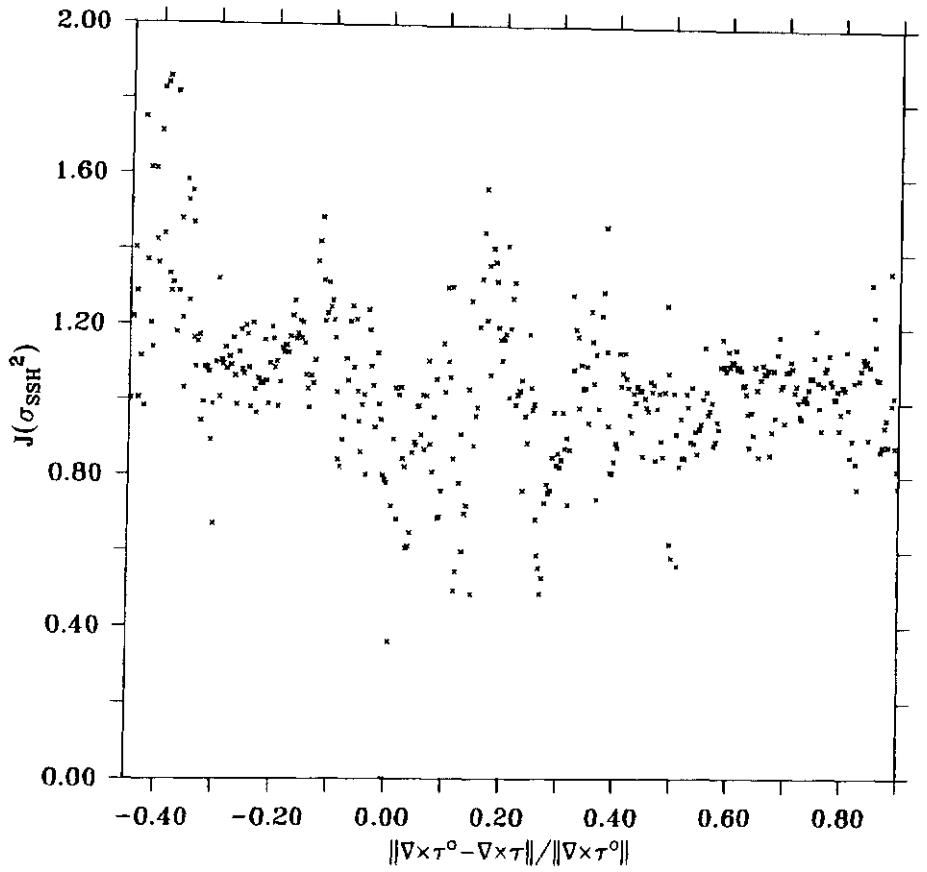


Figure 5.5: *Samples from a section through the cost function based on stream function variances taken from an integration period of one year. The dependence on the wind stress $\tau$ is displayed by the normalized Euclidian difference to the reference value $\tau^o$ as in Fig. 5.2.*

of eddy fluxes derived by Green (1970) and Stone (1972) for the baroclinic instability. A similar relation, that involves the mean kinetic energy derived from thermal wind, was used by Stammer (1997) to characterize regions of high eddy kinetic energy.

Figure 5.6 shows that except for the boundary regions a pattern similar to the variability can be obtained by the parametrization from the mean stream function. Only the relation between variations of SSH variabilities and velocity variations enter the adjoint equations. In contrast to (5.4), an

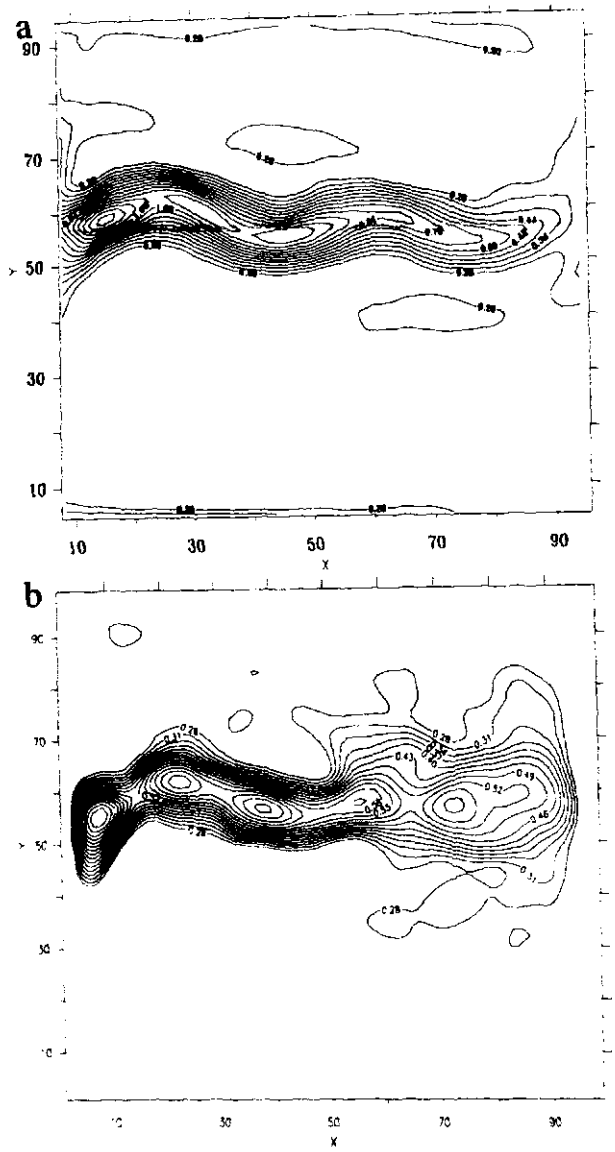


Figure 5.6: *Mean velocity (a, in m/s) and SSH variability (b, in m), derived from a model integration of 2 years.*

additional affine part can therefore be included into the estimation of $\kappa$. The coefficient $\kappa \approx 1$ is estimated from the linear regression displayed in Figure 5.7. Typical regression coefficients are between $r = 0.7$ and $r = 0.9$.

The amplitude of the stochastic component of the cost function based on

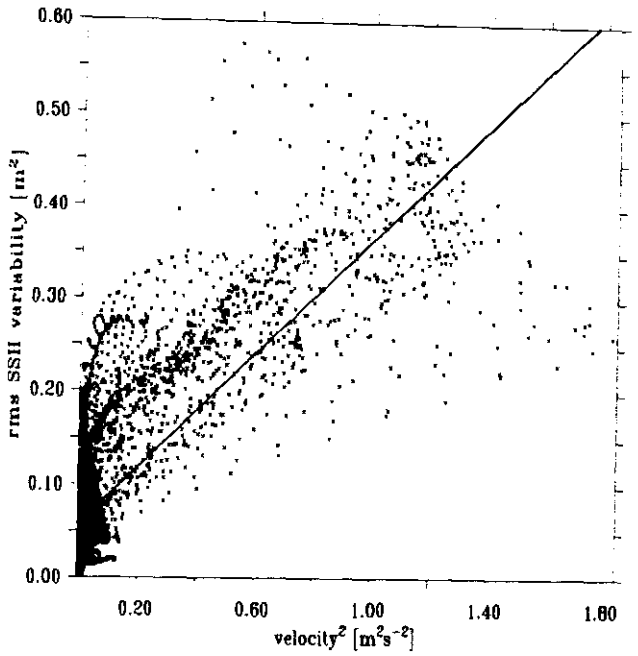


Figure 5.7: *Regression of the patterns shown in Fig. 5.6. The correlation coefficient is* $r = 0.82$.

variance data and illustrated in Figure 5.8 is very similar to the analog function based on mean stream functions. The path in parameter space obtained by optimizing the wind stress shows nearly no convergence to the correct value, although the cost function is reduced to a value below the noise level. The wind stress curl shown in Figure 5.9 reveals that the curve in the vicinity of the line of zero wind stress curl, which in linear Sverdrup theory is the position of the jet, is very well recovered, but shows nearly no convergence elsewhere. This is not surprising since Figure 5.6 demonstrates the the close relation between the mean position of high variability and the mean current. The sensitivity of the SSH variances to changes in the wind stress varies with latitude. Low sensitivities in combination with a large noise level of the cost function introduce a virtually unobservable nullspace that can not be recovered from the variance.

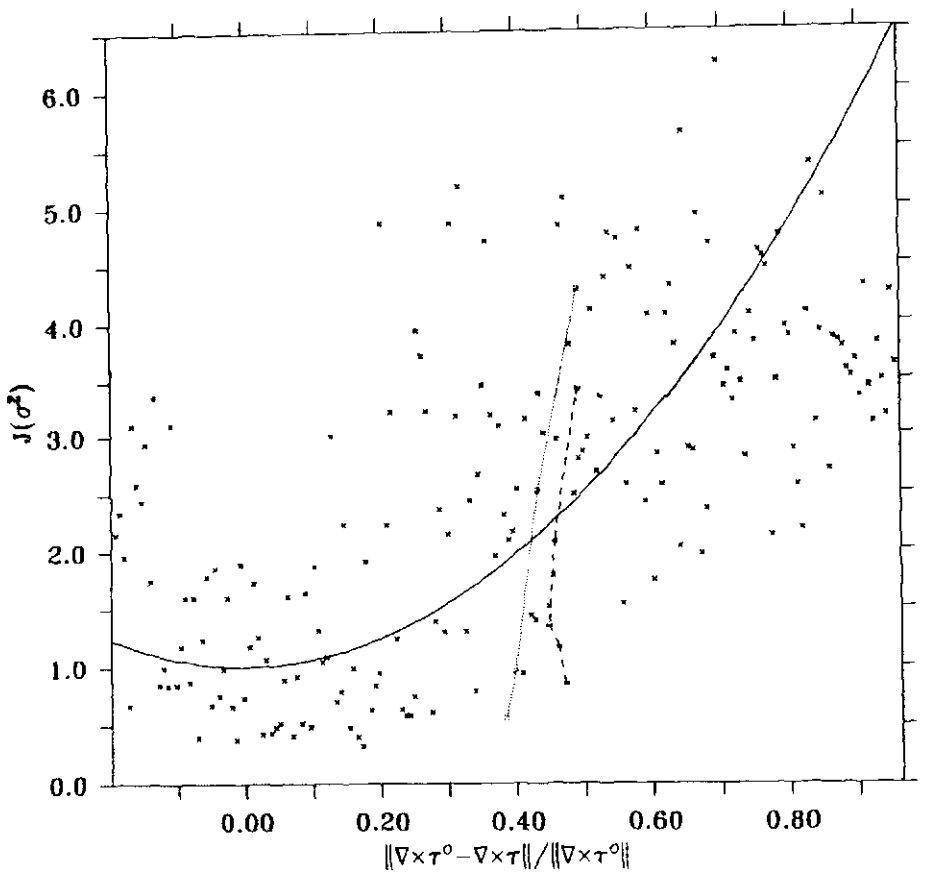


Figure 5.8: *Samples from a section through the cost function based on SSH-variances taken from a period of 5 years, together with paths from optimizing the zonal wind stress $\tau$ with the "closure" from Section 5.4.1 (dashed) and from Section 5.4.2 (dotted).*

### 5.4.2 SVD method for closing

Plotting the cost function values based on mean stream function, $J(\bar{\psi})$, against the one defined on basis of stream function variance, $J(\sigma^2)$, for slight variations in $\tau$ reveals an approximately linear relationship (Figure 5.10). This seems to be not surprising because the tangentlinear approach, this method is based on, assumes that the costfunction may be approximated by linear mappings plus stochastic components, $\tau \rightarrow J(\bar{\psi}(\tau))$ and $\tau \rightarrow$

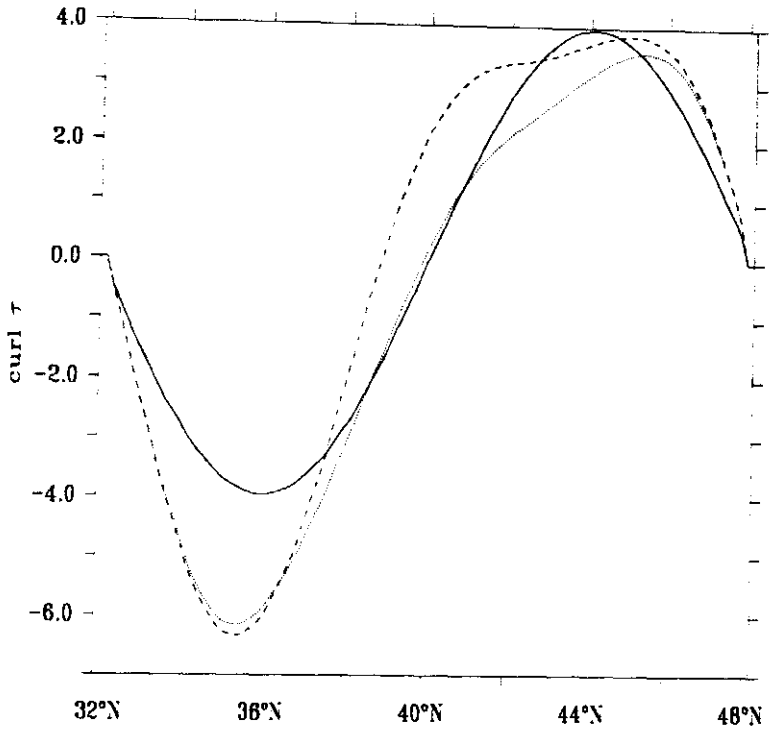


Figure 5.9: *Curl of the wind stress $\tau$: reference (solid), first guess obtained by replacing the amplitude of the forth mode of the wavelet decomposition by twice the value (dashed) and the final iteration (dotted).*

$J(\sigma^2(\tau))$. By inverting one of the maps, a linear relation between $J(\bar{\psi})$ and $J(\sigma^2)$ may be constructed. However, this notion is only correct for macroscopic variations of the parameter $\tau$, when the values of both cost functions simultanously increase with an increasing parameter error. The simple way of searching the gradient, $\partial J(\sigma)/\partial\tau$, by a regression from an ensemble of slight variations in $\tau$, that even avoids the requirement of an adjoint, is therefore found to produce very irregular gradients.

If $\sigma^2$ is locally parameterized by a linear dependency on $\bar{\psi}$, the Jacobian $\partial\sigma^2/\partial\bar{\psi}$ serves as a diagnostic closure for variability. An expression for $\partial\sigma^2/\partial\bar{\psi}$ can be derived from a Monte Carlo approach, where the ensemble is generated by perturbing the parameter $\tau$ by a small amount. With the

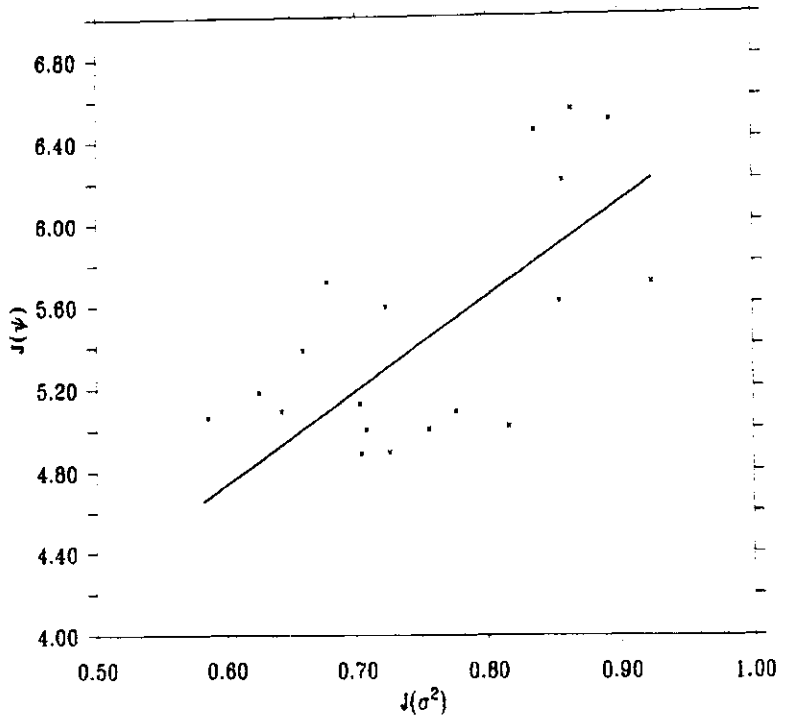


Figure 5.10: *Cost function values based on the mean stream function, $J(\bar{\psi}(\tau))$, versus the one defined on basis of the SSH variance, $J(\sigma^2(\tau))$. Values are derived form integration periods of 5 years with a preceding equilibration period of 5 years by slightly varying the wind stress $\tau$.*

ensemble average $\langle\cdot\rangle_\tau$ the autocorrelation

$$B = \langle(\bar{\psi} - \langle\bar{\psi}\rangle_\tau)(\bar{\psi} - \langle\bar{\psi}\rangle_\tau)^T\rangle_\tau \tag{5.5}$$

and the covariance

$$C = \langle(\sigma^2 - \langle\sigma^2\rangle_\tau)(\bar{\psi} - \langle\bar{\psi}\rangle_\tau)^T\rangle_\tau \tag{5.6}$$

is defined. An expression for the Jacobian is then derived from linear regression by the Gauss-Markov Theorem:

$$\frac{\partial\sigma^2}{\partial\bar{\psi}} = CB^{-1} \tag{5.7}$$

Since the number of rows in $B$ equals the dimension of $\bar{\psi}$, $B$ has for reasonable ensemble sizes, $N$, only a few nonzero eigenvalues. Their number

is always smaller than $N$. By using a singular value decomposition (SVD), it is possible to calculate the Moore-Penrose inverse of $B$ from a limited number of experiments. An ensemble of 5 realizations is used to calculate approximations of $\partial\sigma^2/\partial\bar{\psi}$ within every iteration step. The path from optimization shown in Figure 5.8 and the final curve of the wind stress curl (not shown) reveal the same behavior with the same limitations as the results from Section 5.4.1.

The calculation of the Jacobian from an order of less than 10 realizations is only a sensible task if a simple relation between $\bar{\psi}$ and $\sigma^2$ holds. The approach from Section 5.4.1 shows that such a relation indeed exists. The second scheme therefore works for the same reason as the more simple and efficient approach from the last section. The advantage is however, that no concrete relation has to be found.

An estimation for the posterior parameter error as in Section 5.3.3 gives a value for the normalized rms error of 30%. The estimated parameters are clearly outside of this range and show that an isotropic approximation for this cost function is not appropriate.

# Chapter 6

# Application to the CME Model

## 6.1 Introduction

Altimetric data have become an important source of information about the ocean circulation due to their spatial and temporal coverage of the global ocean. Although the measurements are restricted to the ocean surface, strong vertical coherence within the upper ocean extends the applicability of these data into the subsurface region. An optimal method for the extraction of subsurface information is given by assimilating the data into a general circulation model. Two principally different approaches are available and widely used, namely sequential and variational methods. Both originate from the same principle and both suffer from the imperfect knowledge of the statistical informations of the error covariances. However, following different approximations, they are limited by different obstacles. Sequential methods are based on simplifications of the extended Kalman filter and employ various extrapolation schemes to account for spatial coherence of the error covariance matrix. (Oschlies and Willebrand, 1996; Cooper and Haines, 1996; Evensen and van Leeuwen, 1996; Gavart and De Mey, 1997). The application of the method of Oschlies and Willebrand (1996) reveals that this sequential method while improving only the variability is able to keep the mean state invariant (The DYNAMO Group, 1997) which results in

an unphysical interplay of mean and transient parts of the flow. Additional independent information about the mean state is therefore necessary and was included by Killworth et al. (1999) in this scheme with great success.

Since variational methods search for the trajectory of a given dynamical system that fits the observations optimally, the consistency of the solution with the dynamical equations is guaranteed. However, since the tangent linear equations employed within this method are useful in high resolution models only within the validity of the linearization approach, the method is limited to the forecast range. Application for the assimilation of altimeter data were therefore restricted to very short time spans of a few months (Schröter et al., 1993; Morrow and De Mey, 1995), which do not allow signals to propagate into the deep ocean. In this part the method of Section 4 is applied for the assimilation of sea surface height (SSH) variance into an eddy permitting version of the Community Modeling Effort (CME) model of the North Atlantic ocean. Patterns and amplitudes of annual SSH variability are assumed to derive from an underlying quasi stationary mean circulation which is motivated by a small variance of annual SSH variability when comparing different years from TOPEX/POSEIDON (TP). The close link between mean and variable part as described by Stammer (1997) is then utilized in reversed order for trying to estimate the mean state from the variable part of the circulation by means of data assimilation. The dynamical consistency of the solution estimated by the adjoint method enables in contrast to sequential methods that the mean circulation changes in accordance with the assimilated variability.

After a brief description of the models, the choice of the error covariance matrices is discussed in detail in Section 6.4. An identical twin configuration is subsequently chosen to evaluate performance and limitations of the method before the applied parametrization approach to a closure scheme for the SSH variance is presented in Section 6.6. The results of an one year assimilation are discussed with some emphasis on the effect of the closure scheme in Section 6.7.

## 6.2 General objective and strategy

The intention of assimilating statistical moments is to find the origin of systematic differences of the climatological states calculated from ocean models and observational data and to improve the model formulation by optimizing some of the model's parameters. In ocean models, the initial state, lateral and surface boundary conditions, and coefficients in context with the formulation of mixing are potential parameters, since they are only approximately known. The influence of the initial state should be lost after integrating to a quasi-stationary state. However, most authors include the estimation of initial conditions for temperature and salinity into the optimization since only limited integration periods of a few years are practicable, during which the state is essentially controlled by this parameter (Tziperman et al., 1992; Marotzke and Wunsch, 1993; Schiller and Willebrand, 1995). Since it is found that the assimilation period is restricted to one year because of dynamical reasons as explained below, this approach will be followed and the parameter set is even restricted to the initial conditions in one experiment. The aim of truly improving the model by parameter estimation is a difficult task on short assimilation periods since there is only minor contribution from the surface boundary conditions within a period of one year. Parameter innovations are generally overestimated, when the chosen period is not appropriate for the affected processes to reach a new equilibrium. The main emphasis is therefore to show the feasibility of the method for improving modeled climatological mean states by parameter estimation and for state estimations with high resolution models. Second, it is not clear to what extent multiple stable equilibria may account for differences to the observed data (Dijkstra and Molemaker, 1999). By estimating initial conditions it is possible to construct states that more closely represent the observations and to decide on stability afterwards.

## 6.3 Numerical models

The primitive-equation ocean circulation models used in this study are based on the $1/3^o$ CME model configuration developed by Bryan and Holland (1989) and make use of the revised code described by Pacanowski et al. (1993). The domain of high resolution forward and the coarse resolution twin covers the Atlantic Ocean basin from $15^o$ S to $65^o$ N. Both models have 30 levels and share the same vertical grid spacing which increases smoothly from 35m at the surface to 250m below 1000m. Buffer zones of 5 points width are applied on the closed boundaries where salinity and temperature are restored to data taken from Levitus (1982). The northern boundary condition is supplemented by the signal of Denmark Strait overflow water (Döscher et al., 1994).

### 6.3.1 Forward model

The model configuration is essentially identical to that described by Oschlies and Willebrand (1996). It is forced with monthly mean wind stresses of Hellerman and Rosenstein (1983) and the heat flux is formulated according to the linear approximation of Han (1984). Surface fluxes of fresh water are specified by relaxation to the monthly mean values of Levitus (1982). The horizontal gridspacing is $1/3^o$ in meridional and $2/5^o$ in zonal direction. Horizontal mixing is parameterized by biharmonic friction. Constant coefficients for viscosity and diffusivity are chosen as $2.5 \times 10^{19}$ cm$^4$/s. In the vertical Laplacian mixing is used with constant coefficients of 0.3 cm$^2$/s for diffusion and 10 cm$^2$/s for viscosity. The effect of convective events are parameterized by increasing the vertical mixing coefficients to $10^4$ cm$^2$/s at places in the water column where static instability is detected.

### 6.3.2 Adjoint model

The high resolution and the coarse resolution twin model share the same source code. The horizontal gridspacing is $1^o$ in meridional and $1.2^o$ in zonal direction. Different to the forward model horizontal mixing is parameterized by harmonic friction with coefficients chosen as $5 \times 10^8$ cm$^2$/s to prevent the tangent linear model from developing unstable modes. The adjoint model was constructed with aid of the automatic code compiler TAMC after carefully modifying the code. The compiler developed by Giering and Kaminski (1998) is able to generate adjoint code on basis of the instructions of the forward model. The algorithm treats the program as a differentiable function that maps the parameters onto cost function values. The code is fragmented into basic operations to which adjoint operations have been defined. It then applies the chain rule for differentiating compositions of functions to differentiate the instructions. The code compiler has already been applied to construct adjoint code to different types of models as atmosphere, ocean and ocean wave models (Kaminski et al., 1996; Stammer et al., 1997; Hersbach, 1997). In order to prevent the adjoint from showing the computational mode that results from the Euler coupling in an adjoint to a finite difference formulation as constructed by automatic code compilers (Sirkes and Tziperman, 1997), the code is modified to perform only leapfrog steps. Divergence of the two decoupled modes is not observed since the adjoint variables asymptotically approach the stationary solution of (4.6).

### 6.3.3 Implementation of the method

The application of the method is illustrated by a flow chart in Fig 6.1. The adjoint uses only temporally and spatially averaged informations originating from the high resolution forward model. The estimated gradient and innovations calculated by the descent step therefore belong to a time independent coarse resolution subspace of the original parameter space and the

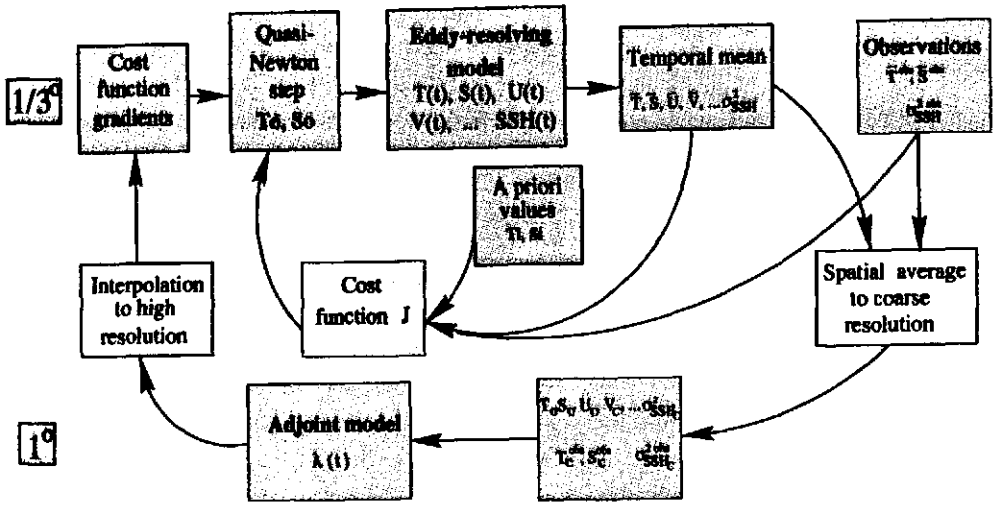


Figure 6.1: *Flow chart illustrating the assimilation method. The forward model, the observations, the descent step and the cost function are defined on the high resolution 1/3°-grid. The cost function gradiens are calculated by an adjoint to a model for the mean circulation which is approximated by a low resolution 1°-twin.*

complement of this space remains unchanged. As described in the first part, prognostic variables were averaged to the coarse grid and enter the adjoint as temporal mean values. Thermohaline and momentum flux boundary conditions are treated in the same way and the topography was constructed from the high resolution representation on basis of the same spatial averaging procedure, completed by an additional removal of holes and the restoration of the islands. The seasonal cycle is removed in the adjoint which describes the adjoint to a model for the mean circulation. This limits potential parameters to the mean component of time dependent parameters. Including a seasonal cycle into the adjoint formalism requires the formation of ensemble mean values to eliminate the transient eddy component from monthly mean values. The treatment of mixing in case of static instabilities makes the mixing coefficients effectively time dependent. By analogy to the treatment of the forcing, convection is parameterized in the adjoint by temporally and spatially averaging the mixing coefficients of the forward model. Exponentially increasing Lagrangian variables which indicate a limit of pre-

dictability were not found for periods of several years if averaged prognostic variables originate from a coarse resolution forward model run. Stronger gradients and higher extreme values of the velocity fields resulting from the $1/3^o$ model, however, limit the integration period for the adjoint to one year when using the fields from the $1/3^o$ model.

The variable storage minimization algorithm M1QN3 by Gilbert and Lemaréchal (1989) represents a good compromise in terms of storage requirements between conjugated gradient and quasi-Newton methods and was chosen as descent algorithm. Gradients calculated by the coarse resolution model are interpolated by bicubic spline to the grid of the $1/3^o$ model before descent steps were performed.

## 6.4 Data sets and error estimates

Estimations for the error covariances are given in the following for each term of the cost function. In order to calculate estimations of the covariances, an approach was followed that is similar to Evensen (1994) and Evensen and van Leeuwen (1996), who derived the forecast error covariance matrix from an ensemble of forward calculations. The ensemble was constructed by adding pseudo random fields with specified variance which were introduced to simulate the system error covariance. Although the strong constraint approach specifies the forward model to be free of error, the moments derived from averaging over finite periods of time have to be regarded as single realizations of moments since small disturbances of the parameters yield a different trajectory. As described in Section 4.2 the variance of these realizations limits the precision of localizing the minimum and should be used for the estimation of the error covariances. The internal variance depends on the length of the averaging period and can be derived from an ensemble of integrations of corresponding length that were formed by members with different but equivalent initial conditions. After integrating the model into a quasi-stationary state it is possible to take the members from slicing a

large integration period into periods of corresponding length. These periods may be regarded as independent due to the rapid decorrelation of the transient eddy signal. For the estimation of error covariances, both, the observational error and the error derived from the variance of the averaged quantities has to be taken into account. This is done by approximating the covariance by the largest of the two errors. If not otherwise specified, the error covariance matrix is chosen to be diagonal, constant in time, and spatially homogeneous.

### 6.4.1 SSH

Sea surface height variance is calculated from five-daily maps of merged data from TOPEX/POSEIDON and ERS1 (TP/ERS1) created by Oschlies and Garçon (1998) enclosing the period from October 1992 to October 1993 (Figure 6.11d). The total measurement error as given by Fu et al. (1994) is 4.7 cm for TOPEX and 5.1 cm for POSEIDON, while an overall rms difference of 3 cm to tide gauge measurements is found by Morris and Gill (1994). The variance of the SSH variability of the model from annual means of 7 years that follow each other shows values larger than 4 cm rms only in a very narrow band above the Gulf Stream where values around 8 cm rms occur. The same calculation on the basis of annual mean SSH variability from TP and ERS1/2 of the years 1993,1996 and 1997 estimated from maps produced by the CLS Space Oceanography Division as part of the project AGORA and DUACS (Traon et al., 1998) and available on CD ROMs produced by the AVISO/Altimetry operations center gives essentially the same values with a shifted region for the maxima. Because of the wrong position of the Gulf Stream in the model and too few data from altimetric measurements, an approximation for a spatially dependent weighting may only be described by an analytical function approximating the distribution from the data. However, a constant value of $\epsilon_{d,SSH} = 4$ cm throughout the experiments was assumed. The effect of introducing of such a distribution would not be large and very local, affecting only 4 % of the data points where

values higher than 4 cm occur.

### 6.4.2 SST

The sea surface temperature (SST) is taken from the 9 km resolution daily nighttime maps of the AVHRR Oceans Pathfinder Program. The mean SST is calculated from maps covering the same period as the SSH data after building monthly mean values to reduce the bias due to cloud cover. The SST available from the Pathfinder Program is an estimate from the brightness temperature which is related to skin temperature of the ocean surface. The coefficients of the algorithm that converts the brightness temperature into SST have been estimated by a regression against in situ observations that are approximately cotemporal and colocated (Podesta et al., 1995), the complete database is included in the AVHRR Matchup Database. The satellite derived temperature therefore more closely resembles the mixed layer SST measured by the buoys. The rms difference between the gridded Pathfinder SST data and SST from the database is 0.94°C for the daytime and 0.97°C for the nighttime matchups (Smith et al., 1996). It also reflects the internal variability of the temperature within a gridbox due to the local character of the measurement by the buoys. Internal variability for one year mean values of SST shown by the model over the same 7 years as regarded above is overall less than 1°C except for the Gulf Stream area where it exceeds slightly the 2°C mark. The error value for temperature observations is chosen as $\epsilon_{d,SST} = 1$°C, since similar limitations hold as for the SSH variability.

### 6.4.3 Climatological data

The depth-dependent rms difference of temperature and salinity values between climatologies of Boyer and Levitus (1997, hereafter WOA97) and Gouretski and Jancke (1998, hereafter SAC) is depicted in Figure 6.2 together

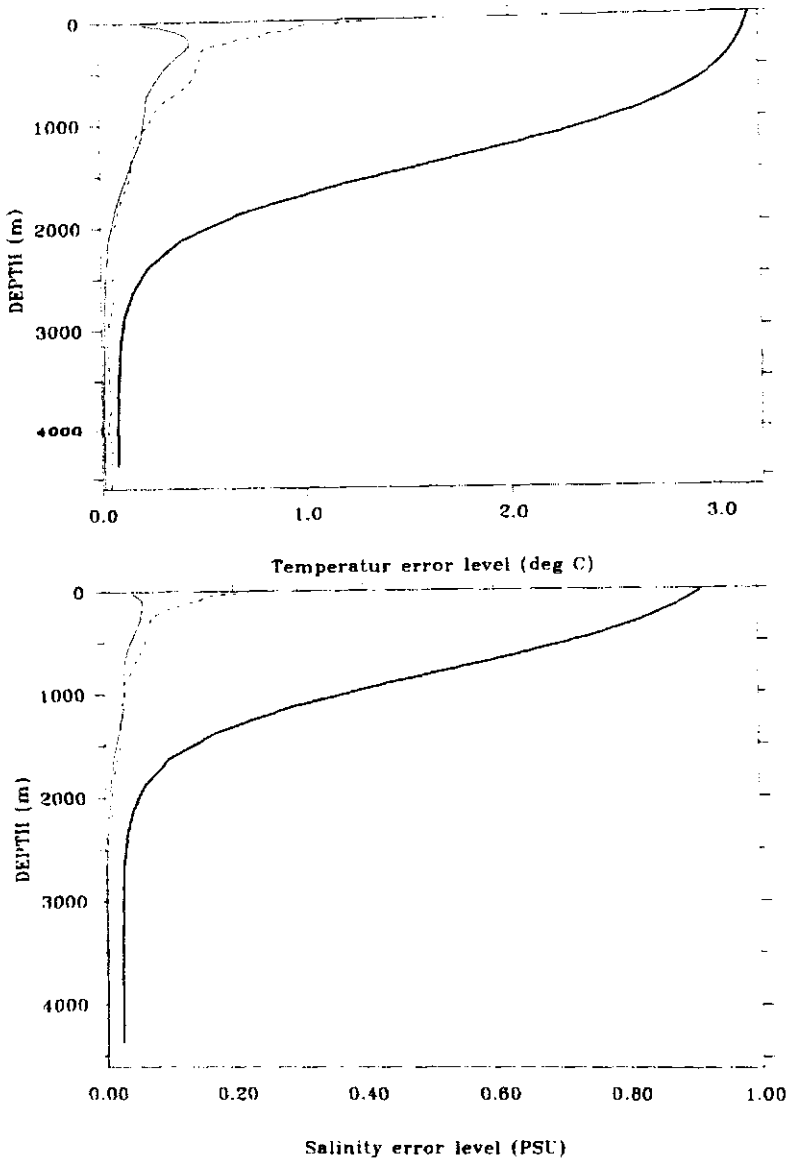


Figure 6.2: *Square root of the horizontally averaged variance of annual mean values (thin solid) together with the averaged rms difference between WAO97 and SAC climatology (dashed). The bold profiles are the specified error functions* $\epsilon_T^d(z)$ *and* $\epsilon_S^d(z)$ *used for the error covariance. They contain estimations of the reduction in degrees of freedom due to vertical correlation within the T and S data and are therefore larger than estimations of the variances.*

with values calculated from the variance of an ensemble of seven annual mean values of a model solution. Both approaches provide similar estimates for the error profiles except in the surface region where variability is underestimated by the model due to the restoring to monthly mean values. Eddy-like spots are visible all over the model region within the WOA97 data below 2000 m due to insufficient smoothing. The variability estimated by the model is thus lower in this region than the error of the data as approximated by the rms difference. Above 2000 m error covariances are specified as functions of depth approximating the variance curves but increased by a factor of 7. Below 2000 m, profiles approximate curves of about twice the data error. This choice and the same argument is used for the parameter penalty term described later, takes vertical correlations in the temperature and salinity fields into consideration which effectively reduces the number of independent degrees of freedom. It includes the effect of disregarding spatial correlations for the error covariances by a simple estimation of the reduction in the degrees of freedom. Typical vertical correlation radii are 350 m above and 500 m below 1000 m depth, taken at correlation coefficients of r=0.5 which are calculated from the above described ensemble. The estimations of the errors $\epsilon_{d,T}(z)$ and $\epsilon_{d,S}(z)$ in Figure 6.2 contain this vertical correlation and could not be used for the judgement on consistency. Horizontal correlations were assumed to be of the same order for all types of data and were therefore consistently left out of consideration.

### 6.4.4 Quasi cyclic stationarity

To avoid a linear drift which may seriously affect the meaning of SSH-variability, a term penalizing temporal drift as proposed by Marotzke and Wunsch (1993) should be introduced into the cost function. For being consistent with the formulation of the adjoint model problem, only statistical stationarity is demanded and the initial and final state should therefore also be a mean state. This requires an integration period of at least 3 years and was not possible because the integration time of the adjoint is limited as

explained above to less than one year. Unlike coarse resolution models the presence of eddys in high reolution models prevent them from being exactly stationary. Even after 20 years of "spin-up" when an approximately quasi-stationary state is reached, states that were separated by exactly one year, differ due to eddies locally up to more than 10°C in temperature and more than 2 PSU in salinity. The gradient calculated from this state is dominated by differences between initial and final state due to different eddy locations. Since it is senseless trying to find realizations which are more stationary than these, the error for the steadyness term has to be in the order of the difference between initial and final state. It was chosen as $\epsilon_{c,T} = 20°C$ and $\epsilon_{c,S} = 10$ PSU. For this reason the term has on the other hand very little effect to avoid the drift in the mean values resulting from imbalances caused by improper surface fluxes.

### 6.4.5 Initial conditions

When the number of parameters exceeds the number of observations the parameter estimation problem is ill conditioned. The problem may be turned into a well posed one with a unique solution when additionally a priori informations about parameters are added. It should be formed by the distance of the parameters to a first guess in a norm that takes the uncertainties into account (Thacker, 1988; Evensen et al., 1998). The first guess is the initial condition after 20 years of "spin-up". The metrics of this norm, the error covariance matrix, was approximated by the derivation of the annual mean values of the reference experiment in comparison with temperature and salinity data from the WOA97 under the assumption that the error is of the same order for the initial conditions. The differences are locally inhomogeneous and the largest errors occur for the upper 1000 m in the Gulf Stream region. The comparison result was integrated into an expressions for the error profiles of temperature and salinity that approximates the profiles of the differences multiplied by a factor to account for vertical correlations. The approximations for the error of the a priori information

then reads, $\epsilon_{i,T}(z)$= 14°C exp(-z/800m) and $\epsilon_{i,S}(z)$= 4 PSU exp(-z/800m), respectively.

### 6.4.6 Surface flux

Weighting coefficients for the a priori parameter values of heat and freshwater fluxes estimated as restoring temperature and salinity were chosen to be only slightly larger than the rms error concerning the initial condition of the surface layer even although surface fluxes are only poorly known. They are fixed as $\epsilon_{f,T}$ = 4°C and $\epsilon_{f,S}$= 0.5 PSU. Without adaption of the surface layer to the changed fluxes an error of the heat flux of about 175 $Wm^{-2}$ would derive from a mean transfer coefficient of 44 $Wm^{-2}$ $°C^{-1}$ used in the Han (1984) formulation. The adaption however, reduces the effect in the CME model to about 8 $Wm^{-2}$ $°C^{-1}$ which yields an error value of about 32 $Wm^{-2}$. The same calculation for the salinity results to an error of the freshwater flux of about 1 m $yr^{-1}$.

## 6.5 Identical twin test

In variational assimilation usually none of the parameters can be regarded as perfect. Even when the model formulation is free of errors, the number of parameters, that one has to estimate to reconstruct perfectly a model analogue of observations which enclose only certain aspects of the state, is usually too large to be determined from the available data alone. In identical twin experiments, data is constructed from a twin model run with specified change in parameters. In this way, the deviation of the solution from the known "truth" can perfectly be traced back to the influence of the parameter in regard. This experimental design is used to evaluate the performance of the methodology. In the first part it was found from identical twin experiments with a quasigeostrophic model that due to high variability of mean and cost function values, the precision up to which parameters

could be recovered is quite poor. However, the prospect of greater precision is expected for regions with lower variability.

A simple experimental configuration, the estimation of the surface heat fluxes from annual mean simulated SST data was selected to compare the precision of the parameter estimation that can be achieved in the CME model with the results from the quasigeostrophic model. The relationship between mean surface fluxes estimated from restoring to an "effective" atmospheric temperature and mean surface temperature is very close. In principle a complete recovering of the parameters should be possible and no a priori informations were introduced. The restoring temperature was multiplyed by an arbitrary factor of 1.2 resulting in a rms temperature deviation of 3.5°C of the mean SST with respect to the reference experiment. The optimization was stopped after 8 iterations and reduced the rms error to 0.05°C with an expected further decrease. Figure 6.3a shows that differences remain at this stage at the continental boundaries. These are due to errors resulting from extrapolating the gradient from the 1° model to the domain of the 1/3° model. Further errors are visible at the northern and southern boundaries. Sensitivities to heat flux changes are much lower at the boundaries because of the lateral restoring boundary condition. The estimation of the Hessian by the minimization algorithm captures different sensitivities only after a few iterations which causes the convergence in these regions to lag behind. In contrast to the results from the QG model the absolute minimum could be precisely estimated and no limitations occur due to stochastic nature of the mean values which should be important mainly in regions of high variability.

The same experimental settings with annual mean SST from the Pathfinder project described in Section 6.4.2 as data reduces the rms difference from 1.34°C to 0.45°C after 13 iterations with no expected further reduction. Differences remaining mainly in the Gulf Stream region result obviously from differences in spurious eddy signatures which are not controllable by the estimated parameter. This low sensitivity is also the reason for the

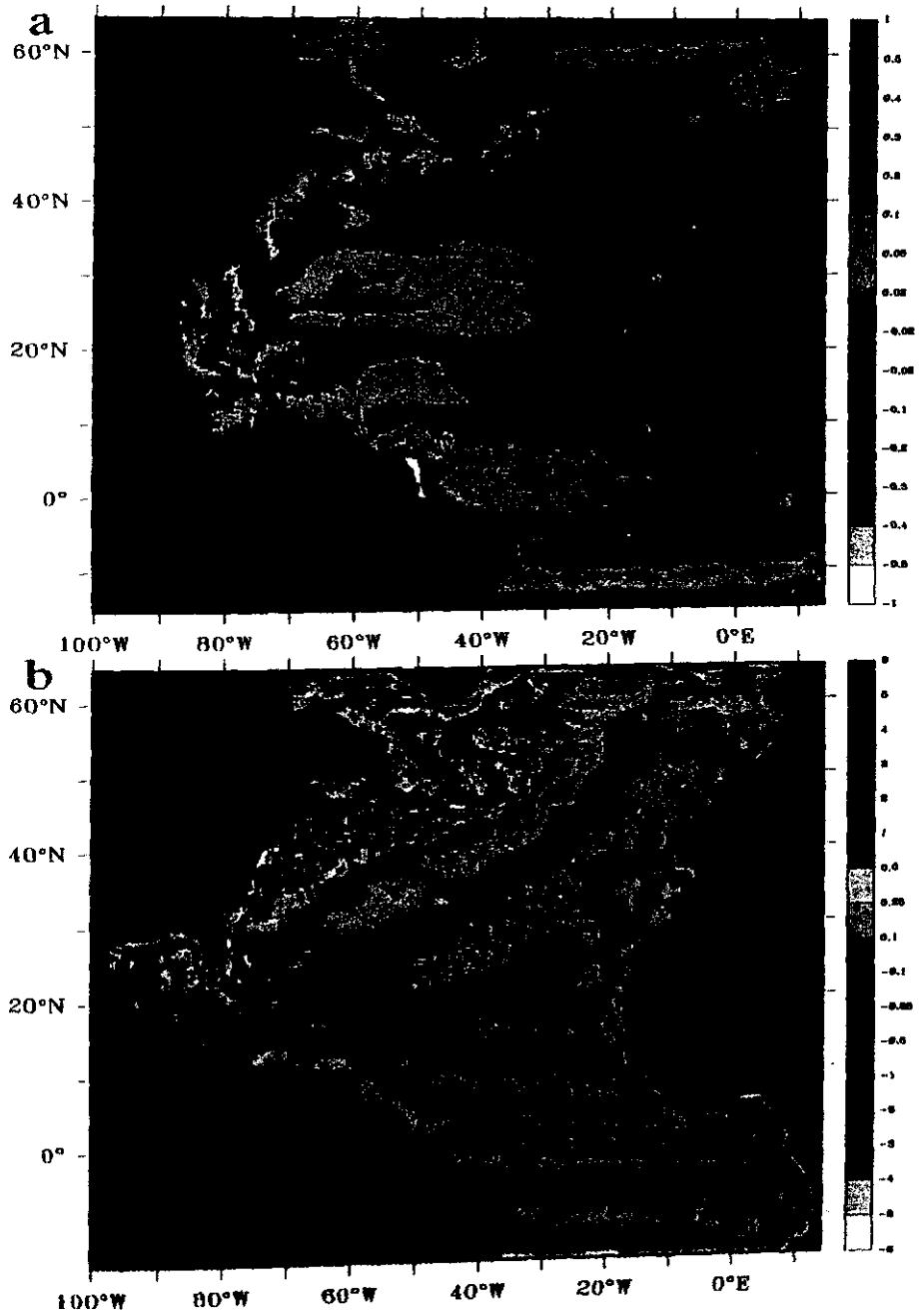


Figure 6.3: *Surface temperature difference in °C remaining after optimization of surface heat fluxes (assimilation - data): (a) identical twin experiment. Mean Pathfinder SST is usedas data in (b).*

quality of the relocation in the first experiment. Large scale disturbances in heat flux have only minor effects on the evolution of the eddy field on time ranges of one year.

Since only statistical properties and not the exact location of the trajectory were constrained by this method, the second experiment gives an estimation of the precision that could be achieved by this method when using real data.

## 6.6 "Closure" for SSH variance

The simple parametrization approach as described in Section 5.4 is followed to mimic a model for the prediction of higher order moments in the adjoint formulation. For the inclusion of SSH variance $\sigma_{SSH}$ as a second order moment, a diagnostic closure is presented that parameterizes locations of high eddy variability in terms of the density structure derived from the mean temperature and salinity distribution. The generation of eddies is closely related to the stability properties of the mean current. Eddy energies calculated from tracked drifting buoys (Richardson, 1983; Krauss and Käse, 1984) and altimeter data (Stammer, 1997) indicate the major frontal zones as the primary location for the occurrence of variability. Outside the tropical regime spectral characteristics of altimetric data from TP analyzed by Stammer (1997) suggest baroclinic instability as the dominant source of variability in accordance with the spectral relations of geostrophic turbulence. A recent Richardson number based parameterization of eddy transfer coefficients invented by Visbeck et al. (1997) uses horizontal and vertical stability, $M^2 = g/\rho_o\sqrt{\rho_x^2 + \rho_y^2}$ and $N^2 = g\rho_z/\rho_o$, to express the coefficient in terms of the isopycnal slope $M^2/N^2$ of the baroclinic front, with g the gravitational acceleration, $\rho$ the density and $\rho_o$ a reference density. This approach was originally derived on basis of the theory of baroclinic instability by Green (1970) and Stone (1972) who described the eddy velocity in terms horizontal and vertical density gradients of the mean flow.

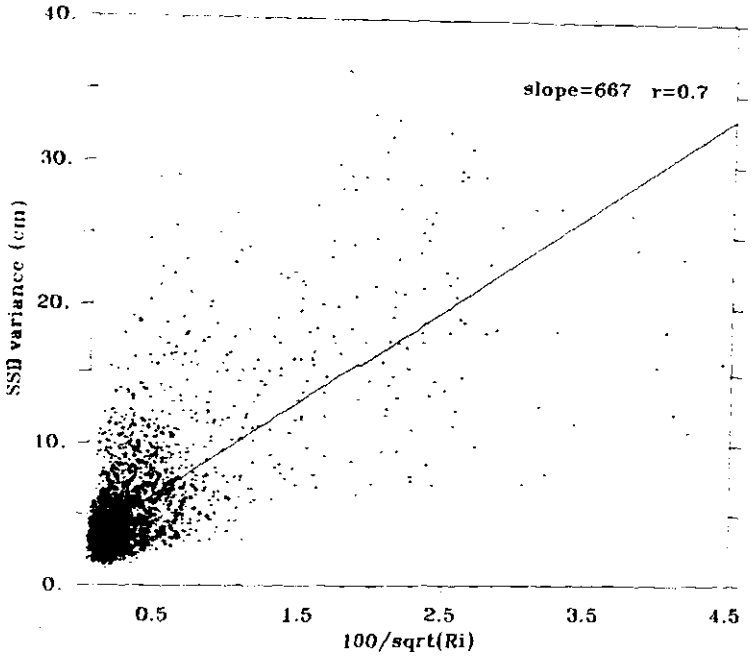


Figure 6.4: *Scatter diagram of SSH variance from TP/ERS1 versus its parameterization (6.1) used with climatological data from Gouretski and Jancke (1998). The region is identical to Figure 6.5, the slope is 667 cm and the correlation coefficient is* $r = 0.7$.

We follow along this line and apply a linear relation between SSH variance and depth integrated eddy velocity expressed in terms of the Richardson number, Ri,

$$\sigma_{SSH} = \gamma \frac{1}{H} \int_{110m}^{1250m} \frac{1}{\sqrt{Ri}} dz = \gamma \frac{1}{H} \int_{110m}^{1250m} \sqrt{\frac{g(\rho_y^2 + \rho_x^2)}{f_o^2 \rho_o \rho_z}} dz, \qquad (6.1)$$

employing the thermal wind balance for the calculation of the vertical velocity shear. $H$ is the depth interval, $f_o$ the Coriolis parameter of a central latitude and $\gamma$ the coefficient of proportionality. The relation is shown by Treguier et al. (1997) to give a reasonable representation of the main regions of variability. The impact of this relation when assimilation SSH variance in regions where the model considerably underestimates variability is to steepen the frontal structure. In this way, the available potential energy as

source for eddies generated by baroclinic instability is enhanced.

Figure 6.4 shows a regression between SSH variance and amplitudes of the parametrization, employing the thermal wind relation as in (6.1) for a calculation of the mean flow velocities from the SAC climatology. Since only the relation between variations of SSH variabilities and density variations enter the adjoint equations, an additional affine part can be included into the estimation of $\gamma$. Figure 6.5 indicates the distribution of the parameterized variability as it is described by the relation (6.1). The qualitative agreement between the pattern in Figure 6.11d and Figure 6.5 is supported by a significant correlation of 0.70. The parameter of proportionality $\gamma$ calculated from regressions with different climatologies or from the corresponding fields derived from the solution of our forward model spans a range of values between 122 cm for the model solution and values of 434 - 667 cm for the climatologies. The small value in case of the model solution may partly be caused by cutting away a large portion of the variability signal, since the parametrization is only defined where the depth is larger than 1250 m. The value for the implementation into the adjoint is chosen as $\gamma = 200$ cm between $10^o$ S and $50^o$ N and $\gamma = 0$ elsewhere to exclude tropical and high latitude regions where the correlation is very low. A relation to the horizontal gradient of the mean velocity calculated from thermal wind as Stammer (1997) presented gives also good correlations and parametrization. However, it is very similar to (6.1) and would affect the results only marginally.

The parametrization is based on a function that contains only horizontal and vertical derivations of the mean density and a constraint on absolute density values and the distribution among temperature and salinity is not provided by the assimilation of SSH variability. From this one can conclude that an application of the scheme for the estimation of initial conditions for temperature and salinity will result in a large subspace of equivalent solutions. In the subsequent sections two ways are presented to handle this problem. One is the already mentioned inclusion of a priori informations

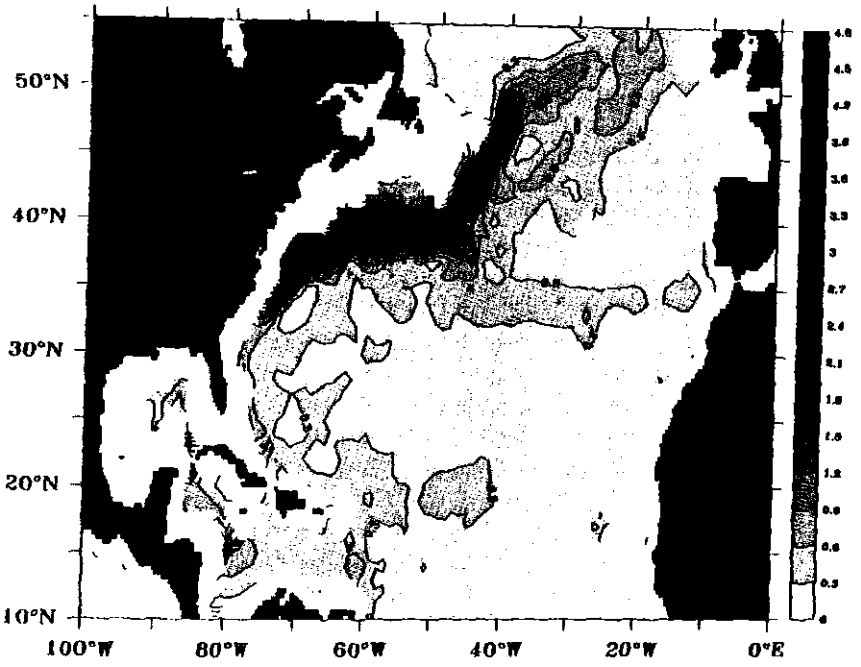


Figure 6.5: *Parameterization of SSH variance employing relation (6.1) with climatological data from Gouretski and Jancke (1998) ($\gamma = 1$).*

for the parameters and a second would be to add climatological data of temperature and salinity.

In order to explain the effect of the parameterization some results are presented from the experiment described in detail later in Section 6.7.2. Figure 6.6 shows horizontal and vertical cross-sections through the gradient of the cost function parts $J_{SSH}$, which measures the difference of the SSH variances, and $J_{WAO97}$ which measures the difference of annual mean temperature and salinity values to the WOA79 data with respect to the temperature initial condition. Although there are markly differences between the patterns, both indicate the same characteristic errors of the model which has a northward displaced Gulf Stream with too low variability and almost no Azores Current with the associated variability. The general features of $\partial J_{SSH}/\partial T_o$ confirm the above made supposition, locations of underestimated SSH variance are distinguished by spatial gradients in the pro-

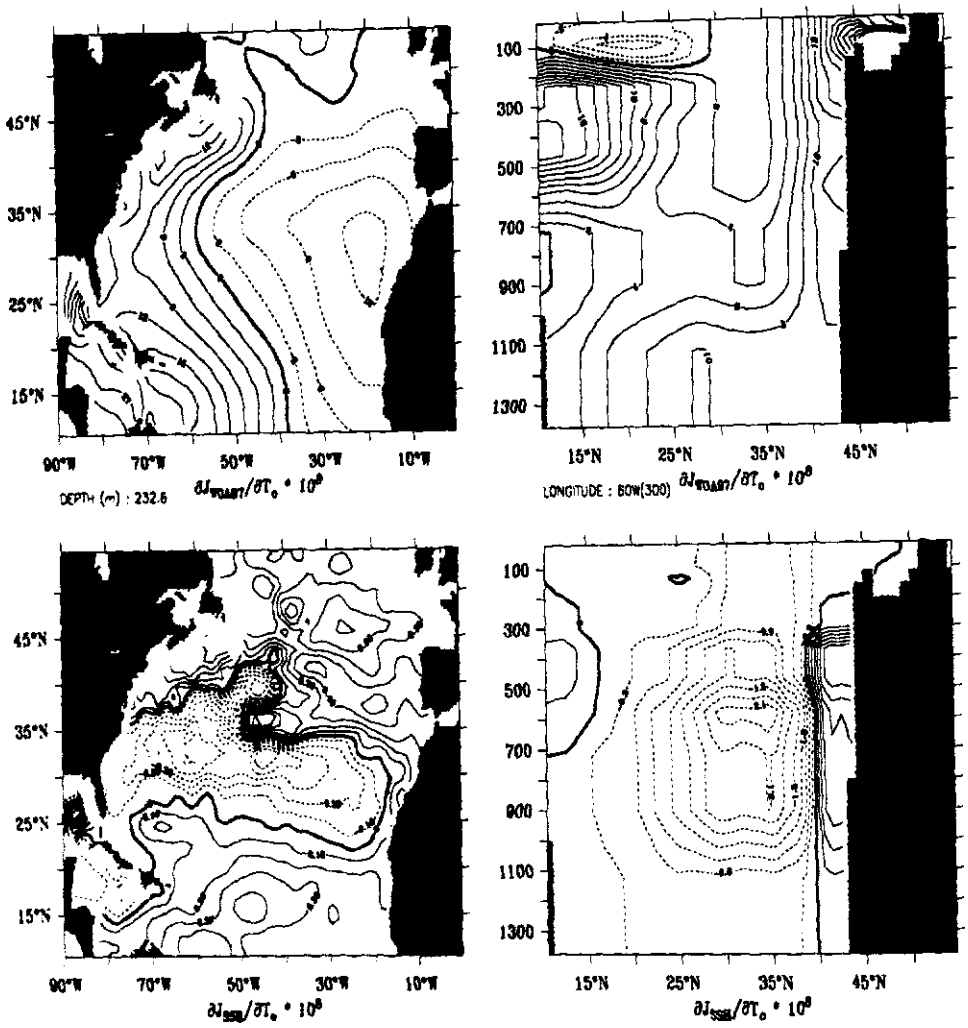


Figure 6.6: *Gradient with respect to the temperature initial condition of the cost function part $J_{SSH}$, which measures the difference of the SSH variances, and $J_{WAO97}$ which measures the difference of annual mean temperature and salinity values to the WOA79 data (in $1/^oC$). The horizontal level is at 230 m depth and vertical sections is along 60° W.*

posed temperature change. Spatial structures of the gradients of either part are consistent in suggesting warmer water south and colder water north of the modeled Gulf Stream. The vertical structure of both gradients share some similar features although $\partial J_{SSH}/\partial T_o$ is derived via the closure scheme from the temperature and salinity structure of the forward model. The

scheme therefore provides a method for the vertical extrapolation of SSH data. The precise position of the Gulf Stream is difficult to constitute from $\partial J_{WOA97}/\partial T_o$ while the gradient of $J_{SSH}$ contains spatial gradient informations that clearly marks the position. In the same way, the signature of the Azores front is only visible from the $\partial J_{SSH}/\partial T_o$ term.

## 6.7 Assimilation experiments

For the specification of absolute density values and the distribution among temperature and salinity, different strategies concerning the set of the included cost function terms are pursued. The cost function is defined as a sum

$$J = \sum_{\beta,\alpha} J_\alpha^\beta \tag{6.2}$$

of quadratic parts

$$J_\alpha^\beta = \frac{1}{N} \sum_n^N \frac{(\alpha_n - \alpha_n^{obs})^2}{\epsilon_{\beta,\alpha}^2} \tag{6.3}$$

with $\alpha \in \{SSH, SST, T, S\}$ representing the data types and $\beta \in \{c, d, f, i\}$. N is total number of temperature gridpoints of the model. The term for cyclic stationary and the one for annual or climatological means of observational data are labeled with c and d, and $\beta = i, f$ denotes the a priori information term for the initial condition and the surface fluxes, respectively. The notation corresponds to the naming of the weights $\epsilon_{\beta,\alpha}$ used in Section 6.4. There are only data terms for $SSH$ and $SST$. The iterations are started with the same parameter set as the control run which is the year after a twenty-year "spin-up" from the state of rest.

### 6.7.1 Including a priori informations

In the experiment described in this section informations on density values are retrived from a priori informations. The cost function includes 8 terms which are defined by the set $\{J_\alpha^\beta : \alpha \in \{T, S\}, \beta \in \{c, f, i\}\}$ supplemented

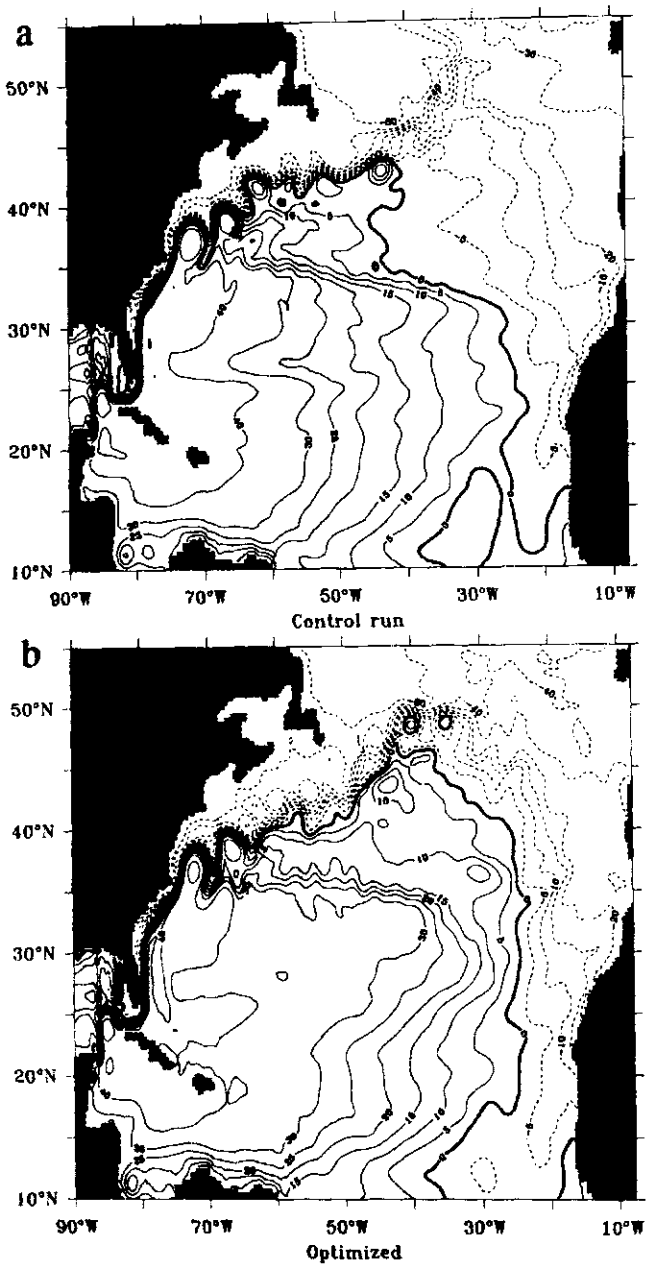


Figure 6.7: *Mean sea surface height in cm: (a) control run and the 11th iteration with optimized initial conditions for temperature and salinity and surface heat and freshwater fluxes including additional a priori informations for the parameters (b) shown together with Singh and Kelly (1997) data (c).*

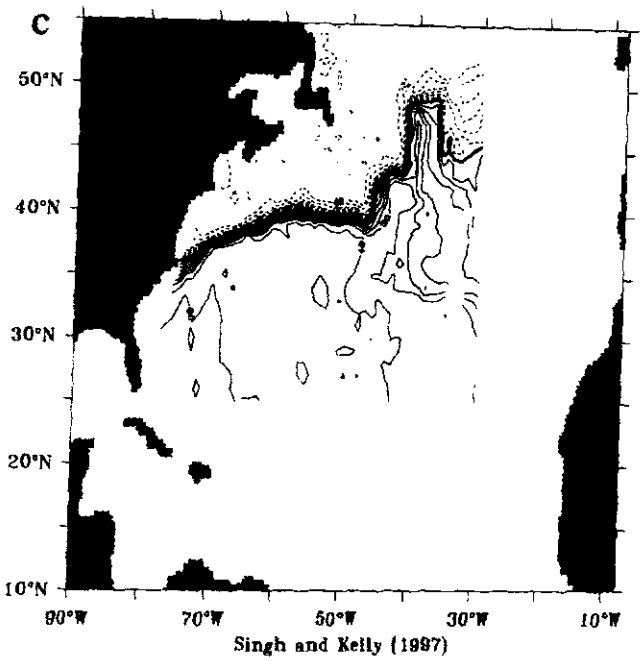


Figure 6.7: (continued)

by the data terms $J^d_{SSH}$ and $J^d_{SST}$. The estimated parameters are the initial conditions for temperature and salinity and the corresponding surface fluxes. Only a brief description of the results is presented to give insight into the reasons for the configuration chosen in the the following section. Results are presented on the basis of mean SSH because the focal point is most clearly visible in this quantity. Figure 6.7 depicts the annual mean SSH from the control run and the final iteration together with data from Singh and Kelly (1997) who estimated mean SSH from a combination of hydrographic and altimeter data. The mean front of the control run is displaced around 60°W to the north and is shifted around 42°W to the east and it is noticeably weaker than the observations suggest. The final iteration resembles more the control run than the data of Singh and Kelly (1997), although the solution is improved. This confirms partly the assumption made in introduction that it may be possible to retrieve information on the mean state by assimilating variability. However, it is assumed that the mean state predicted from the

assimilation would be more closely resemble the observation if the penalty from the a priori informations does not push the solution back to the control state.

About 20% of the cost function in the final iteration emerges from the a priori information term whereas the data terms are only reduced to 68%. The problem arises to some extend from the ratio of the number of parameters to the number of data points which is partly captured by including vertical correlations. More seriously is the violation of the assumption of a Gaussian error of the a priori informations. Since the first guess initial condition is subject to strong systematic errors, it is not possible to find an improved state consistent with the observations. From a theoretical point of view the result may under the prerequisite of proper chosen error covariances represent the optimal compromise between the position of the front emerging from a priori informations and the one which optimally corresponds to the observed SSH variance. However, the effect of assimilation is quite low and informations on systematic errors should additionally be included. A possible way to remove this systematic deficit in the initial conditions would be to subtract the modeled mean state from it and to add the data from the WOA97 afterwards. However, far better than this is to include the WOA97 data directly into the cost function to constrain the mean state and dismiss the a priori informations instead. This is done in the experiment described in following section.

### 6.7.2 Including WOA97 data

In addition to the SSH variance data, data from the WOA97 is assimilated in this experiment. The configuration of the cost function then encloses 5 terms defined by the set $\{J_{SSH}, J_\alpha^\beta : \alpha \in \{T,S\}, \beta \in \{c,i\}\}$; the cyclic stationary terms are still included although they are defacto not effective. The total value of the cost function is reduced within 10 steps from 0.43 corresponding to the value of the control to 0.21 at the final iteration. A

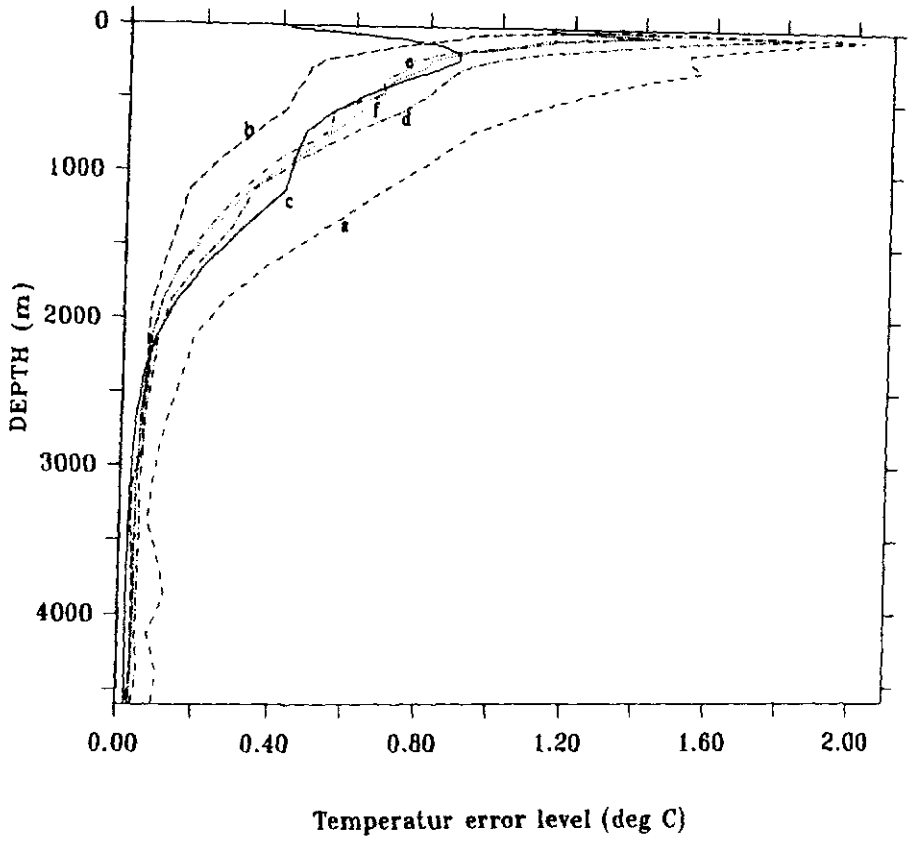


Figure 6.8: *Posterior error profiles calculated from the rms difference with temperature and salinity fields of the WAO97 climatology. Control (a), assimilation run including WOA97 and SSH-variance data (f), the following year of this assimilation run (d) and assimilating only WOA97 data (e). Curve (c) is twice the ensemble error covariance and (b) the rms difference between the SAC and the WOA97 climatologies from Figure 6.2.*

second iteration started from a different first guess was performed to investigate the sensitivity with respect to the starting point. The total cost function was then reduced from 0.90 to an identical value of 0.21 at the 10th

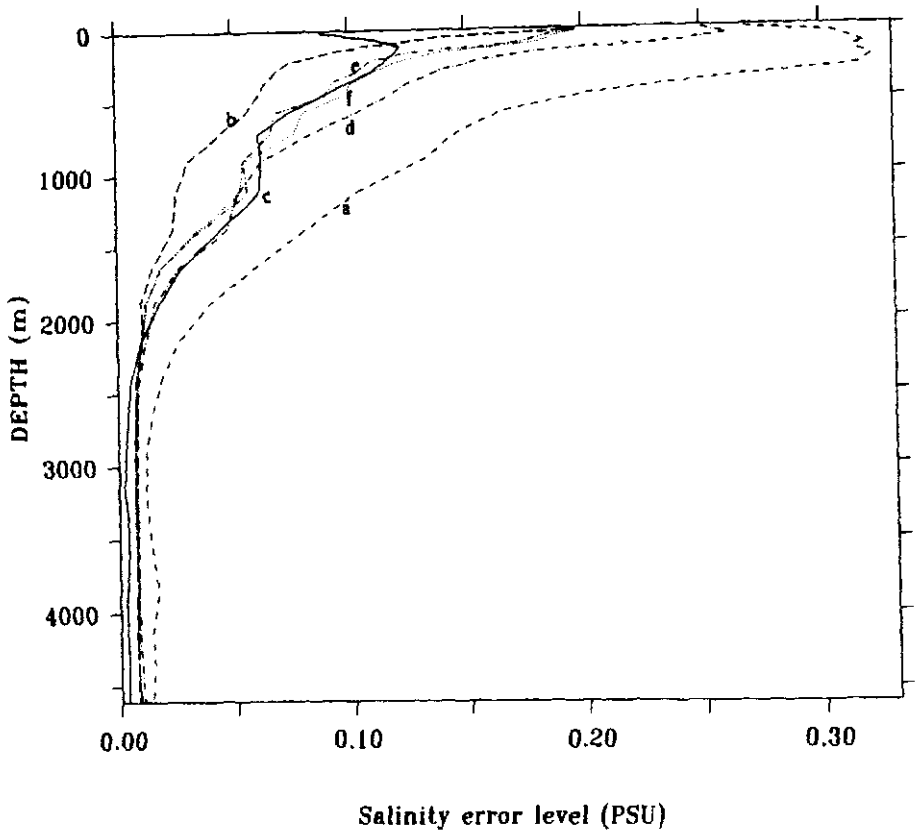


Figure 6.8: (continued)

iteration. Compared with the first iteration, an almost identical state was found that only differs due to different eddy realizations. The second iteration was continued for a further 18 iteration and all results presented in the following are taken from the 24th iteration since no further improvement was achieved during the 4 last iterations. The cost function value is then reduced to 0.18. The integration is continued afterwards for two additional years to investigate the stability of the solution. In order to evaluate the informations retrived from the variance data an experiment with the iden-

tical configuration without the SSH-term was performed. The cost function is then reduced after 21 iterations to 0.14. The value is almost identical to the corresponding cost function part of the above experiment that includes additionally SSH data. If not explicitly stated otherwise, results denoted by the assimilation run are from the assimilation that includes both, WOA97 and SSH variance data.

The effect of assimilation is assessed by posterior profiles of rms errors depicted in Figure 6.8. Except for the near surface region the profiles are generally in the order of twice the ensemble error variance corresponding to a 95% level. The usual demand for consistency is a level lower than once the error variance. It is not clear if this level could actually by reached, particularly since the eddy variability is markly enhanced (see the following) and the positions of the eddys can not be optimized by this method. The failure in the surface region results from inconsistent surface fluxes, which were not optimized. With respect to an error value of 0.45°C as achieved in Section 6.5, it is supposed that the profile would follow the twice error covariance profile even in the surface region when surface fluxes were included. The inclusion of SSH data does not affect noticeably the error profiles in the assimilation run (compare green and magenta profile in Fig. 6.8). This supports the assumption that informations introduced by the assimilation of SSH data is consistent with the WOA97 data. A slow return of the state to the state of the control is documented by an increase of the rms error for the second year (blue profile).

Figure 6.9 and Figure 6.10 demonstrate that the main source for the error in the upper 1000 m below the surface region is due to the effect of spurious eddies which is still present in the mean fields. The largest differences remain in the regions of high eddy kinetic energy, particularly in the Gulf Stream region. However, significant differences occur also east of Grand Banks and in the Caribbean Sea which have important dynamical consequences. Below 1000 m differences remain after optimization due to the already mentioned eddy like spots present within WOA97 data. The resulting mean state is

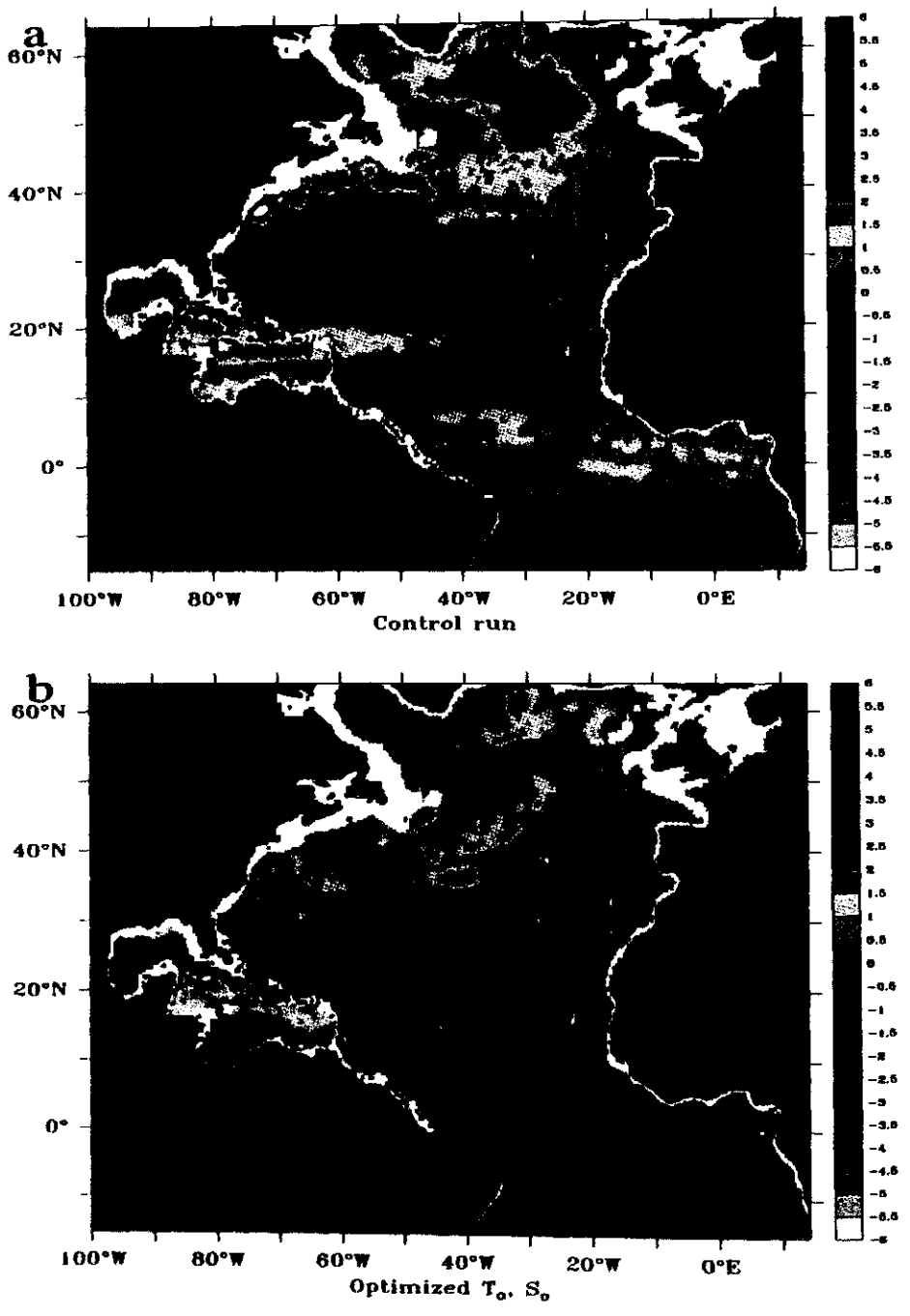


Figure 6.9: *Temperature difference in °C in 580 m: Control run - WOA97 data (a) and assimilation run - WOA97 data (b).*

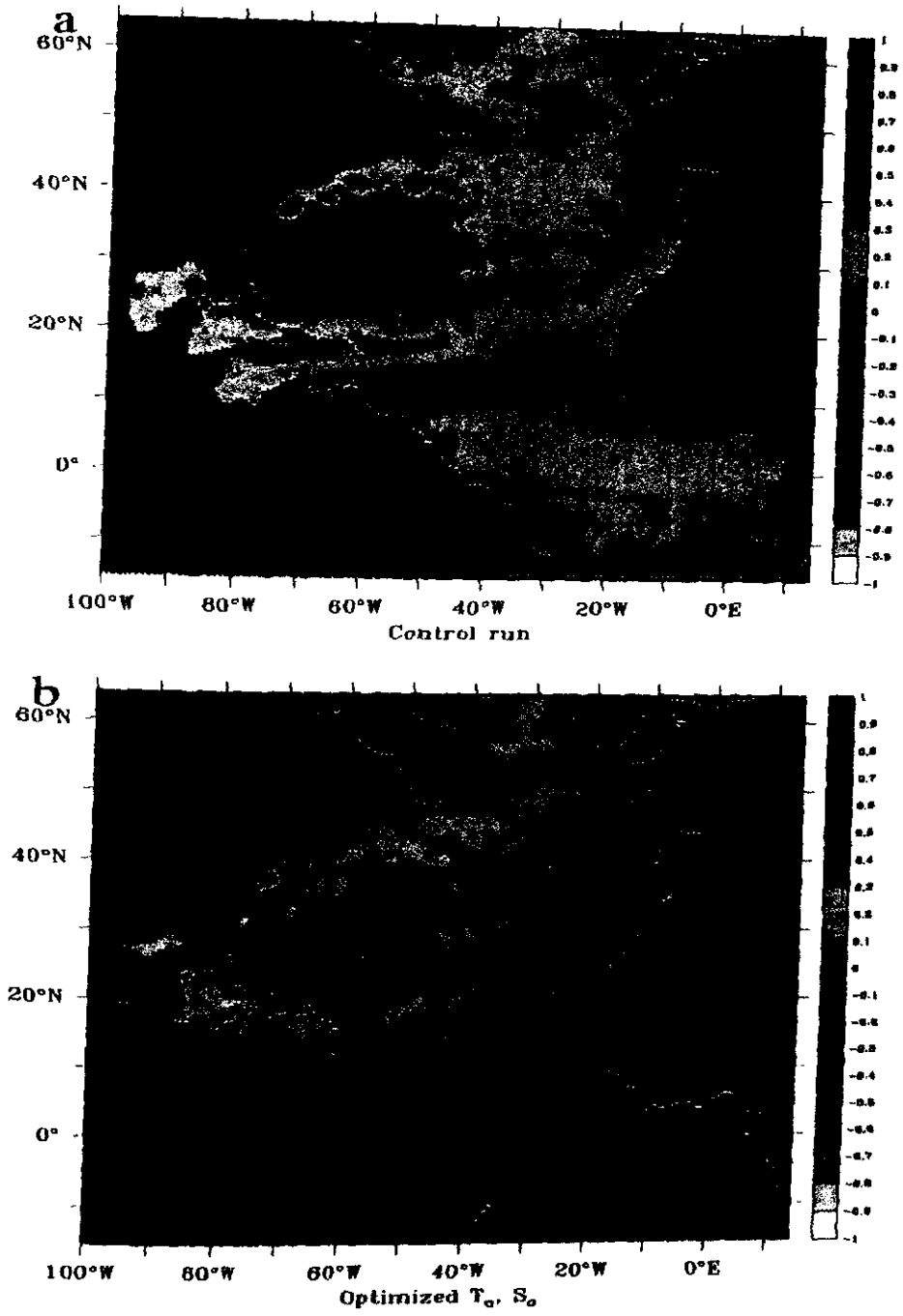


Figure 6.10: *Salinity difference in PSU in 580 m: Control run - WOA97 data (a) and assimilation run - WOA97 data (b).*

not claimed to be perfect as will be shown in the following, but since the remaining error is to a large extend caused by the irreducible part emerging from the eddy effects a further reduction is difficult to achieve.

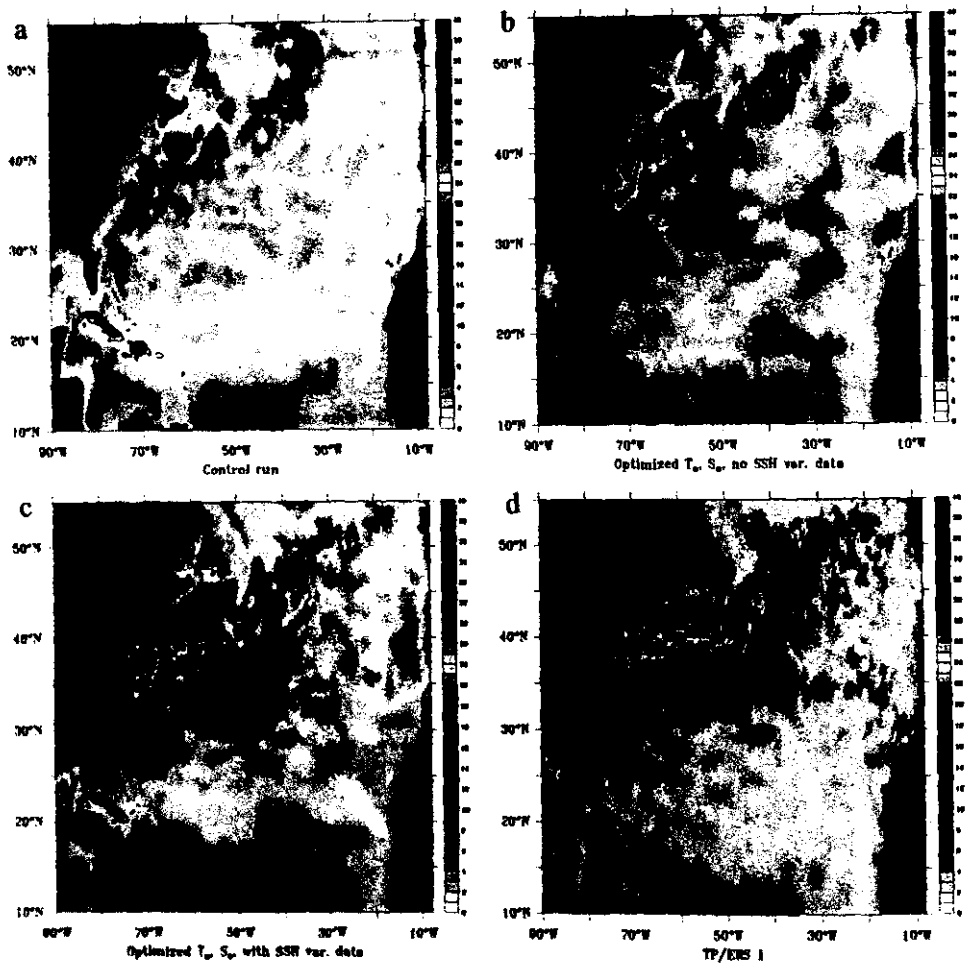


Figure 6.11: *RMS SSH variability in cm. (a) control run. (b) and (c) are from optimizing initial conditions for temperature and salinity. (b) includes only WOA97 and (c) both WOA97 and SSH variance data. (d) is from mapped TP/ERS1 data.*

Although SSH variance is assimilated, it is not clear whether a realistic representation of the observations could be found by only estimating initial conditions for temperature and salinity. The rms SSH variance error is reduced from 7.2 cm to 4.8 cm which means a reduction to 45% in terms

of cost function values. The sequential method of Oschlies and Willebrand (1996) in a slightly different model configuration reduces the rms error from 6.5 cm to 3.5 cm which corresponds to a reduction of a cost function value to 28% (pers. com.). Figure 6.11 shows the SSH variance from the control and the assimilation run together with the observational data. The maximum of the variability of the control is displaced northward and overall too low with a maximum around 47°N. By assimilation of SSH variance and WOA97 data, the erroneous maximum vanished and the mean position and amplitude of maximal variability is fairly well matched, though it extends further to the north around 55°W and the northward extension at 42°W is not captured. However, the turnoff of the Azores Current is visible by increased variability. The eddy scales are clearly too large which may be explained by still too low resolution. The level of SSH-variability in the Gulf Stream region is also enhanced to a realistic magnitude by assimilating only the WOA97 data, but its position and the maximum around 47°N is not changed.

The mean SSH indicating the position of the mean near surface current is illustrated in Figure 6.12a. It supports the findings from regarding SSH variance. The position and amplitude around 60°W matches the data of Singh and Kelly (1997) with a slightly weaker front. However the front at 42°W is not present, although it may be seen within the estimated initial conditions which indicates a dynamical deficit of the model. The amplitude of the stationary anticylone east of Cape Hatteras is reduced together with the disappearance of the front which resembles a wrong turnoff position of the Azores Current. The mean SSH in Fig. 6.12b represents a somewhat intermediate state between the control state in Fig. 6.7a and data in 6.12a. The Gulf Stream front is stronger but the position is nearly unchanged in comparison to 6.7a. The pattern in the Azores region resembles the one of Fig. 6.12b but the strength of the front is too weak.

Within the following two years of integration the state returns to the pattern of the control run. The mean front starts continuously to split into

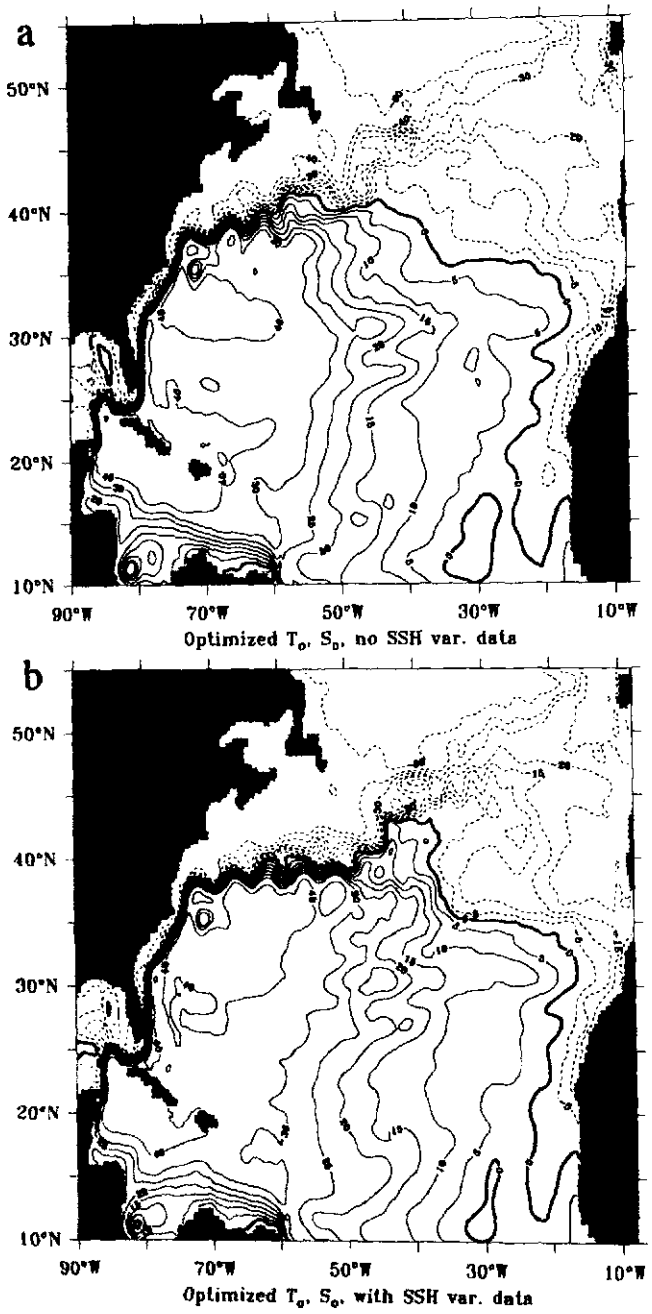


Figure 6.12: *Mean sea surface height in cm, with optimized initial conditions for temperature and salinity including informations from WOA97 data (a) and including additionally SSH variance data (b).*

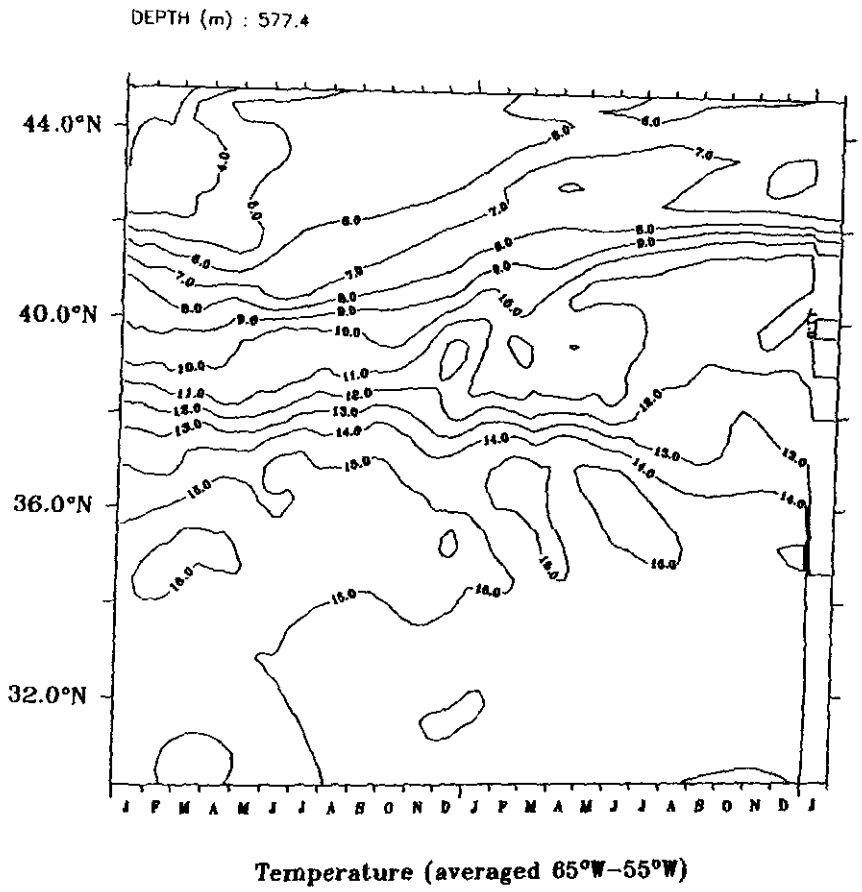


Figure 6.13: *Time series starting after insertion of optimized initial conditions for temperature and salinity. The temperature in 550 m depth is averaged from 65° W to 55° W. Mean values of the control run are added in January after the second year to mark the difference to the control run.*

a northward displaced and the southern front visible in Figure 6.7a. The return is documented in Figure 6.13 which shows a meridional temperature section in the Gulf Stream region. Cold water inserted north of the Gulf Stream is rapidly removed, whereas the mean frontal position only slowly returns to the position of the control run. In some iterations the stationary anticylone northeast of Cape Hatteras completely disappears and the Gulf Stream separation is almost realistic. This state is then found to be stable for two years.

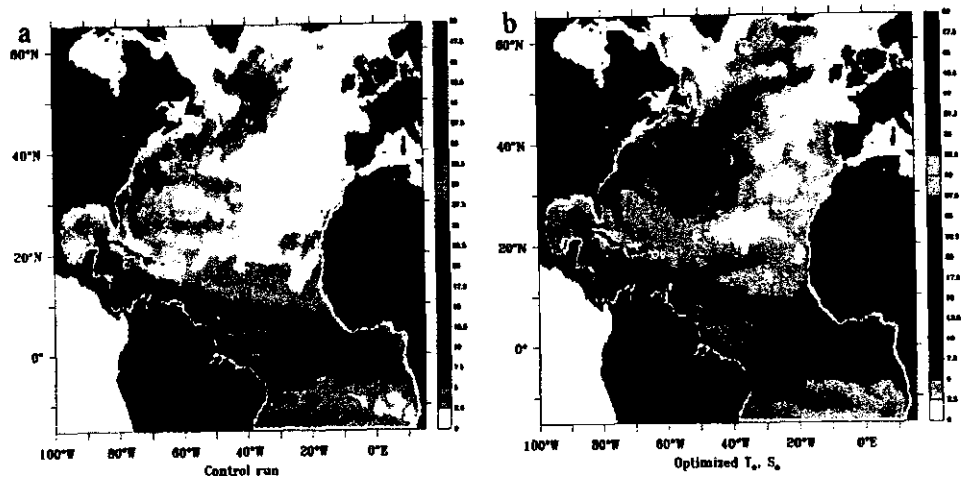


Figure 6.14: *Square root of near surface eddy kinetic energy, (a) control run, (b) with optimized initial conditions for temperature and salinity from assimilating SSH variance and WOA97 data.*

The structure of the eddy kinetic energy (EKE) field shown in Figure 6.14 reflects the pattern from the SSH variance, demonstrating that the former is in fact the reason for the improved SSH variance. The EKE level is increased in almost all regions, particularly also in areas of beforehand low EKE as it is associated with the Azores current. In the control run the current is only represented as a markly southward shifted band branching east of Cape Hatteras from the Gulf Stream as visible from the near surface velocities in Fig. 6.15a. Assimilation shifts the band to the observed position at 34°N (Gould, 1985) and the current originates, as suggested by Sy (1988), from the separation of the Gulf Stream into the North Atlantic Current (NAC) and the Azores Current (Fig. 6.15b). The erroneous flow of the NAC towards the Irminger Sea visible in Figure 6.15a is corrected and the transport follows after assimilation the realistic route towards the Norwegian Sea, although it deflects east of the Rockall Plateau to the northwest.

The position of Azores front is difficult to identify in the control from the potential density section averaged between $35°W - 30°W$ as illustrated in

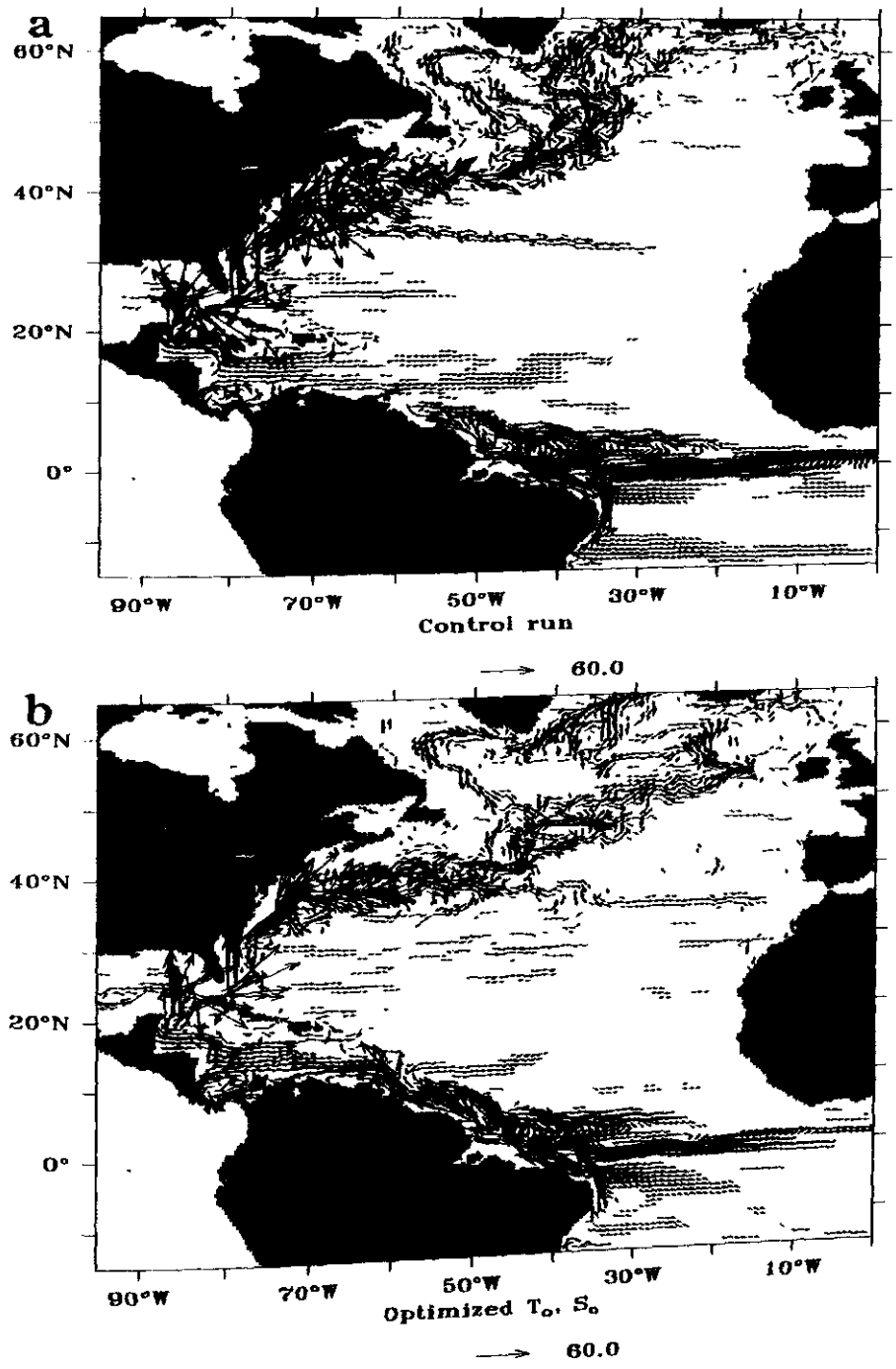


Figure 6.15: *Annual mean velocity at 90 m in cm/s. Control (a) and assimilation run (b).*

Figure 6.16, whereas in the assimilation run the frontal structure is even more pronounced than in the WOA97 data. Likewise, the position is shifted slightly to the north, indicating that the current and the front is mainly established by informations from the altimeter data.

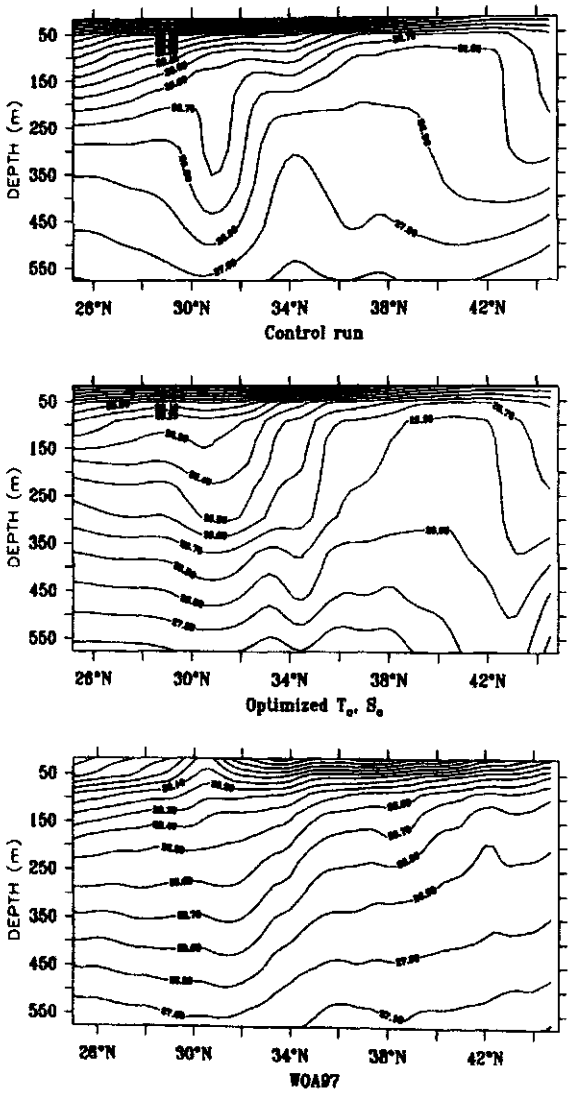


Figure 6.16: *Sections of potential density averaged from 35° W-30° W (ci=0.1). Control run (upper), assimilation run (middle) and data from Boyer and Levitus (1997) (lower).*

An amplification of the EKE level shown in Figure 6.17 from values less

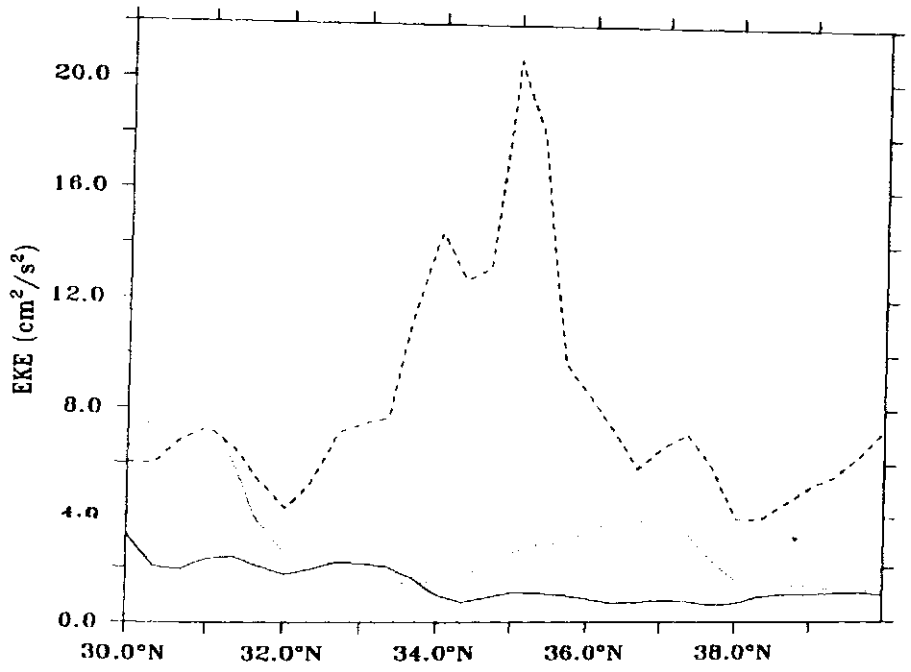


Figure 6.17: *Near surface eddy kinetic energy averaged from 35° W to 25° W for the control run together with values from the assimilation with (dashed) and without SSH-variance data (dotted).*

than 4 $cm^2 s^{-2}$ to values up to 20 $cm^2 s^{-2}$ in average accompanies the intensification, but is still much too low compared to values up to 200 $cm^2 s^{-1}$ Brügge (1995) calculated from drifter data. Some iterations show values up to 100 $cm^2 s^{-1}$. However, the associated SSH variability is then found to be higher than the values from TP/ERS1. Values over 100 $cm^2 s^{-1}$ are thus not consistent with the assimilated SSH variability from mapped TP/ERS1 data. The interpolation of track data generally underestimates the amplitude of the the real signal which causes a systematic error. It would be, particularly in association with an adjoint method, more appropriate to use along track variability data.

There is nearly no improvement through the assimilation of only climatological data and a slight enhancement is visible in Figure 6.17 only north and south of the current position for this experiment.

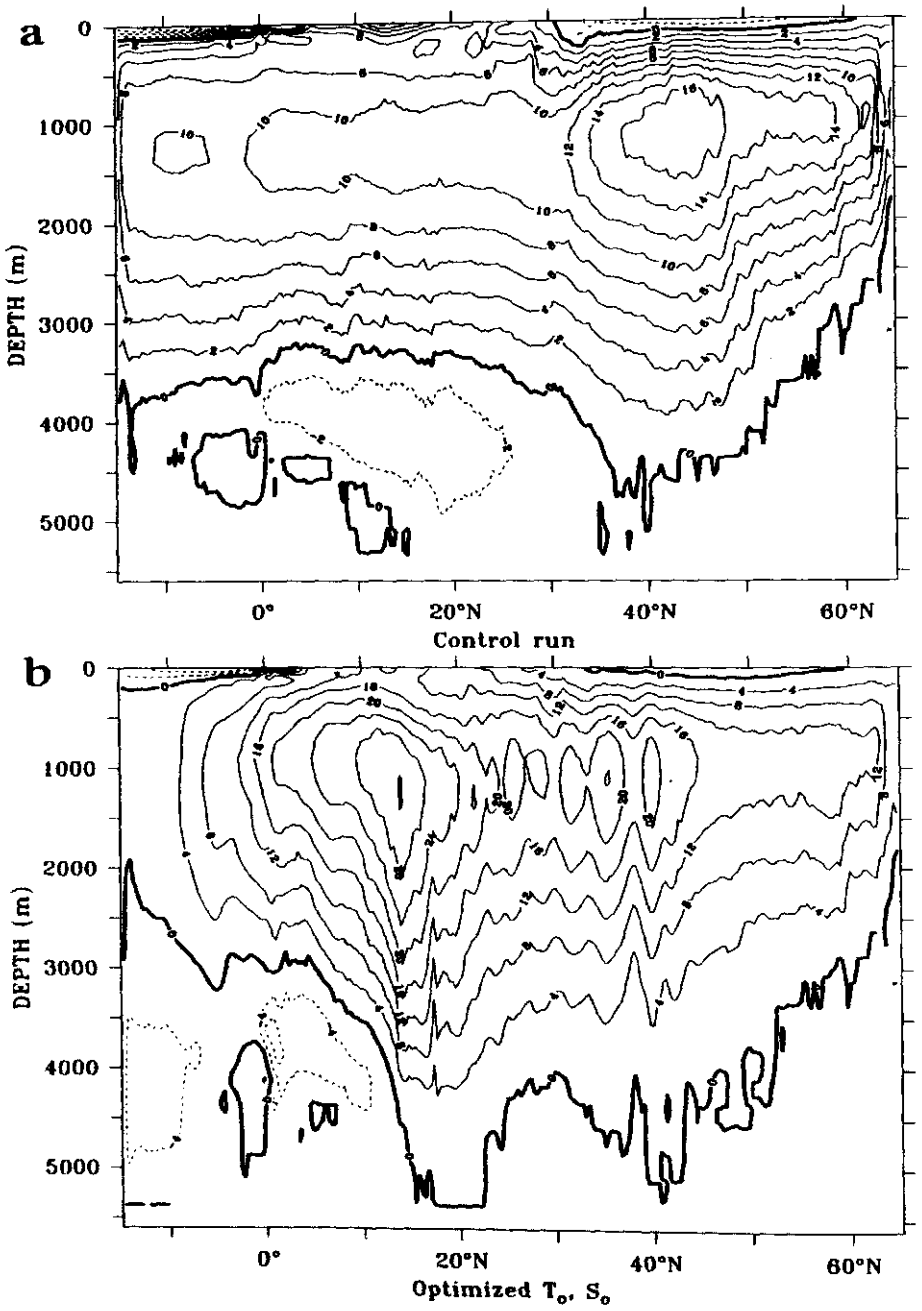


Figure 6.18: *Meridional overturning stream function; (a) control run (ci= 2Sv), (b) with optimized initial conditions for temperature and salinity (ci= 4Sv).*

An effect on the deep circulation is demonstrated in Figure 6.18. The meridional overturning stream function remains mostly unchanged north of 40°N where the overturning is mainly controlled by the northern boundary condition (Döscher et al., 1994). It increases to very unrealistic values of about 30 Sv north of the equator. The strength of the overturning continuously increases with the number of iteration. The increased sinking between 15°N and 20°N is limited to the region of the Caribbean Sea and north east of the Antilles where also large differences to the WOA97 data occurs. In spite of this difference, continuously denser water is introduced between 600 m and 1000 m depth in this region during the optimization. It is not quite clear if the adjoint variables in this region are affected by remote model-data differences or if approximations made within this method are responsible for a wrong estimation of the gradient in this area. The heat transport is due to the close relation to the overturning strength (Böning et al., 1996) also nearly doubled in the area around 15°N. The quite small difference to the WOA97 data in terms of the contribution to the cost function (differences due to eddy signatures at the same level in the Gulf Stream region or due to insufficient smoothing of the WOA97 data are much larger) has important dynamical consequences. The overturning and heat transport have returned in the following year back to the values of the control. Although the unrealistic transports values could be reduced to deviation to WOA97 data, the results demonstrate a poor precision for determining these quantities from assimilation experiments when quasi-stationarity is not guaranteed.

### Energy transfers

The invention of a closure scheme was handled by a simple parameterization approach. However, it is not necessarily guaranteed, that the action of this closure, although physically motivated, will result into a physically sensible representation of the parameterized observations. SSH variance for example is equally possible to generate in a non equilibrium situation by the steric effect through a density flux. By analyzing the energy balances

it is demonstrated, that the generation mechanism for eddy kinetic energy (EKE) follows the processes, on which the closure is based. The Gulf Stream region (60°W-40°W and 38°N-44°N) is chosen to demonstrate the effect of the insertion of the optimized initial conditions for temperature and salinity in the experiment in Section 6.7.2, where SSH variance is assimilated.

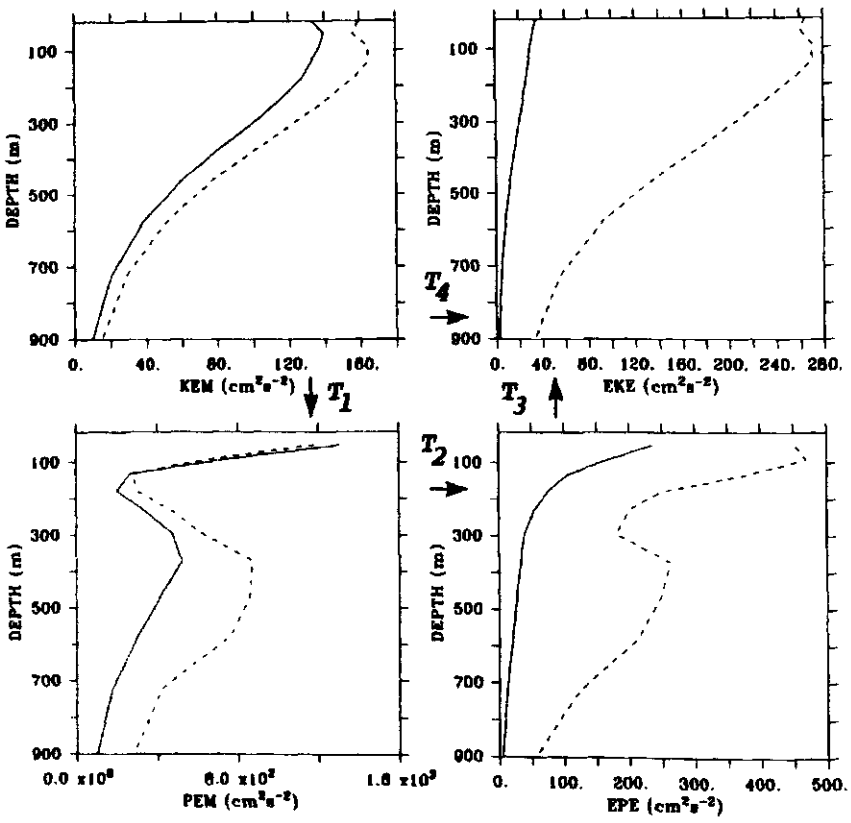


Figure 6.19: *Vertical profiles of the energy components averaged over the region 60°W-40°W and 38°N-44°N for control (solid) and assimilation run (dashed). Definitions for the energy components and the energy transfers are given in the text and in the appendix.*

The vertical profiles of the four energy components depicted in Figure 6.19 indicate an enhancement of all components in the assimilation run. The

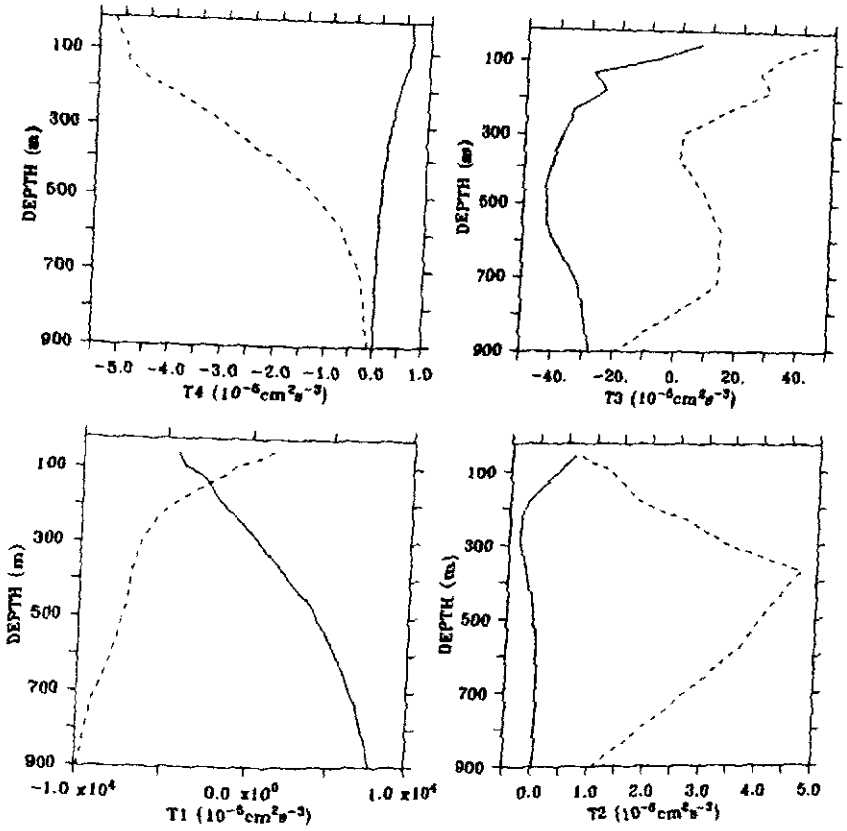


Figure 6.20: *Vertical profiles of the energy transfer coefficients averaged over the region 60°W-40°W and 38°N-44°N for control (solid) and assimilation run (dashed). Definitions are given in the appendix and indicated in Figure 6.19.*

effect on the mean potential energy (PEM) is maximal at about 500 m depth corresponding to the position of the maximum of the profile derived from the inverse Richardson number used in the parametrization (6.1). The effect of the parameterization was explained in Section 6.6 to be based on improving the conditions for baroclinic instability by steepening the front, which simultaneously enhances the amount of mean available potential energy (PEM). Following the energy route of the baroclinic process, PEM is

subject to instability and converted into eddy potential (EPE) and finally into eddy kinetic energy (EKE).

The picture of an enhancement of PEM as the primary source for energizing the flow is supported by regarding the energy transfer terms displayed in Figure 6.20. The baroclinic conversion term $T_2$ and term $T_3$, which describes the conversion from EPE to EKE, characterize the path, how the enhancement of mean potential energy cascades into EKE. These terms are markly enlarged although $T_3$ is strongly sensitive to the choice of the subregion and is therefore not a reliable quantity. The barotrophic conversion term T4, showing that barotropic instability is weak in the control run, reverses its sign and indicates a transfer of energy from the eddy component to the mean flow.

The reversal of the conversion of mean kinetic energy (KEM) to mean potential energy (PEM) is represented by the term $T_1$. The reversal of sign results in the assimilation run from the enhancement of sinking southeast of the boundary current, which was discussed in association with the overturning. The picture of a constructed baroclinic front, that is after insertion eroded by energy conversion due to baroclinic instability, is supported when regarding the following years of the assimilation run and it appears that the reduction of EKE precedes the return of the mean flow and not vice versa.

The interpretation of the results should only be regarded as a qualitative support for the arguments given in Section 6.6 since the flow is not stationary and returns to the reference state for longer integrations. However, the situation is generic for the action of the parametrization and also holds in the Azores Current region with shifted maxima towards the surface layers. An exception concerns the $T_1$ term, which remains almost invariant in the Azores region.

# Chapter 7

# Conclusion

The experiments presented here demonstrate that it is possible to improve the state of a dynamical system by assimilating statistical moments. An extension of the original adjoint formulation is derived which offers a way for improving model parameters and may be used to correct systematic errors in the climatological state of ocean models.

Cost functions which depend on statistical moments are shown to describe invariant characteristics of the dynamical system in a sensible way. Secondary minima emerge when the period of assimilation is increased beyond the forecast range and hinder the convergence of the standard adjoint method. For long integration periods of high resolution models, the statistical quantities are regarded as single realizations of Gaussian distributed moments, which are shown to change smoothly for variations of parameters.

The new approach invents the adjoint to a prognostic model that describes the dynamics of the moments. The formalism then calculates cost function gradients that ignore the fine structure of secondary minima. The application of coarse resolution models as approximations for the moment models makes it necessary to add special closure schemes for the assimilation of higher order moments. Very simple parametrization approaches are tested for this purpose and shown to be able to predict reasonable gradients. A

limited number of iteration steps are found to be sufficient to minimize below the noise level. The stochastic nature of moments calculated from high resolution models enables the estimation of error covariances from ensembles of simulations.

By regarding statistical moments the method provides as a by-product a Gaussian distribution for the quantities used to define the cost function. This condition enables the validity of the minimum variance argument used in variational methods which is generally found to be violated in highly nonlinear models (Miller et al., 1999). However, the restoration of this condition should not be overrated since its validity is limited by a time scale separation argument which probably will not hold for ocean models in same distinct way as in the Lorenz case.

The precision up to which parameters could be recovered is quite low for the QG model. Both, from the optimization and from a simple theoretical consideration, error values higher than 20% are found, which at the first sight is not very encouraging. However, this is not a deficiency of the method, but results from very high variability of the averaged quantities, when applying the method to the most nonlinear areal of western boundary currents.

The original intention was to provide a framework for the improvement of climatological states of high resolution ocean models by data assimilation. With respect to this aim the experiments with the CME model only provide a first step towards a true model improvement. The feasibility of the method for improving the climatological state of eddy resolving models is however indicated by assimilating SSH variance and climatological data, although it was not possible to find a new stable equilibrium state by estimating the initial conditions for temperature and salinity. In order to answer the question whether the climatological state of the model is consistent with the observations, a time scale has to be included. On annual periods a mean state almost consistent with the assimilated data could be constructed. However, the state is unstable and still subject to deficits.

The method turned out to provide gradients that, although they are only approximative, perform efficiently in minimizing cost functions. From an identical twin experiment with the CME model it seems that the precision of the estimated parameters is not limited that seriously as in the experiments with the QG model. A greater precision for parameter estimations could be achieved due to a lower noise level in the basin wide application.

The assimilation of SSH variability with a simple diagnostic closure scheme derived from baroclinic instability theory is shown to be able to introduce information that is complementary as well as consistent with the climatology. An important aspect of this new technique is to extract characteristics of frontal structures from variability, since this type of information is not well represented in climatological data since they are usually subject to extensive smoothing. In this way the method allows to use informations from the variable signal to determine the mean state. In order to assess the success of the assimilation it is not sufficient to regard only assimilated quantities. Since the adjoint method enforces consistency with dynamics of the used model, the quality of the mean state and further quantities that were derived via the dynamical equations are strongly limited by dynamical deficits or unestimated parameters of the model. This was one obstacle in the presented experiments.

It appears from some iterations that a marginally stable solution showing a correct Gulf Stream separation with a small basin of attraction is possible for more than one year. The solution may then be repelled by eddy induced variability or by a slow change that comes along with increased heat transport. From this it seems to be plausible that only minor changes in the model parametrizations may be sufficient to enable a correction for at least this point. It remains at the end open whether estimable parameters are responsible for systematic errors that a great mayority of this type of models share.

# Appendix A

# Energy Transfer Coefficients

A brief definition for the various components concerning the energy cycle is given. The terminology of Böning and Budich (1992) is followed but potential density $\rho$ was applied instead of a linear relation to potential temperature. If $\tilde{\rho}$ denotes the horizontal average of $\rho$ and $\rho$ is divided into a time mean $\bar{\rho}$ and a transient eddy part $\rho'$, the available potential energy per volume $V$ can be separated into a mean (PEM)

$$\frac{1}{2}g\int\frac{(\bar{\rho}-\tilde{\rho})^2}{d\tilde{\rho}/dz}dV \tag{A.1}$$

and transient part (EPE)

$$\frac{1}{2}g\int\frac{\overline{\rho'^2}}{d\tilde{\rho}/dz}dV. \tag{A.2}$$

The same separation applied for the horizontal component of the kinetic energy yields to a term for the mean kinetic energy (KEM)

$$\frac{1}{2}\int(\bar{u}^2+\bar{v}^2)dV \tag{A.3}$$

and the eddy kinetic energy (EKE)

$$\frac{1}{2}\int\overline{(u'^2+v'^2)}dV. \tag{A.4}$$

Regarding the equations governing the rates of change of potential and kinetic energy, the division into mean and fluctuation components yields

four terms which describe the redistribution of energy among the above defined four parts. The work of mean buoyancy forces converts mean kinetic in mean potential energy

$$T_1 = -g \int \bar{w}\bar{\rho} dV. \tag{A.5}$$

Positive values of

$$T_2 = -g \int \frac{\overline{u'\rho'}\partial\bar{\rho}/\partial x + \overline{v'\rho'}\partial\bar{\rho}/\partial y}{d\bar{\rho}/dz} dV \tag{A.6}$$

indicate baroclinic conversion of mean to eddy potential energy, which may then convert to eddy kinetic energy connected with a positive

$$T_3 = g \int \overline{w'\rho'} dV \tag{A.7}$$

term. Barotropic instability is characterized by the conversion of mean to eddy kinetic energy governed by

$$T_4 = -\int \left[ \overline{u'^2}\frac{\partial\bar{u}}{\partial x} + \overline{u'v'}\left(\frac{\partial\bar{v}}{\partial x} + \frac{\partial\bar{u}}{\partial y}\right) + \overline{v'^2}\frac{\partial\bar{v}}{\partial y} \right] dV. \tag{A.8}$$

This set of eight terms gives no a complete description of the generation and destruction mechanisms of energy. External forces and internal dissipation and diffusion processes have to be considered, supplemented by advection terms if only subregions with open boundaries are analyzed.

# Bibliography


Boffetta, G., P. Giuliani, G. Paladin, and A. Vulpiani, 1998: An extension of the Lyapunov analysis for the predictability problem, *J. Atmos. Sci.*, **23**, 3409–3416.

Böning, C. W., F. O. Bryan, W. R. Holland, and R. Döscher, 1996: Deep-water formation and meridional overturning in a high-resolution model of the North Atlantic., *J. Phys. Oceanogr.*, **26**, 1142–1164.

Böning, C. W., and R. Budich, 1992: Eddy dynamics in a primitive equation model: Sensitivity to horizontal resolution and friction, *J. Phys. Oceanogr.*, **22**, 361–381.

Boyer, T. P., and S. Levitus, 1997: Objective analyses of temperature and salinity for the world ocean on a 1/4 degree grid. NOAA Atlas NESDIS 11, *U.S. Gov. Printing. Office, Washington, D.C.*.

Bryan, F. O., and W. R. Holland, 1989: A high-resolution simulation of wind- and thermohaline-driven circulation in the North Atlantic Ocean, *in Parametrizations of Small Scale Processes, Proceedings of the Aha Huliko a Hawaiian Winter Workshop*, 99–115, edited by P. Müller and D. Henderson, University of Hawaii, Honolulu.

Bryan, F. O., and R. D. Smith, 1998: From eddy-permitting to eddy-resolving, *Int. WOCE Newsletter*, **33**, 12–14.

Cooper, M., and K. Haines, 1996: Altimetric assimilation with water property conservation, *J. Geophys. Res.*, **101, C1**, 1059–1077.

Cooper, M., and K. Haines, 1996: Altimetric assimilation with water property conservation, *J. Geophys. Res.*, **101, C1**, 1059–1077.

Courtier, P., J. N. Thépaut, and A. Hollingsworth, 1994: A strategy for operational implementation of 4D-Var, using an incremental approach, *Q. J. R. Met. Soc.*, **120**, 1367–1387.

Dijkstra, H. A., and M. J. Molemaker, 1999: Imperfections of the North Atlantic wind-driven ocean circulation: Continental geometry and wind-stress shape, *J. Marine Sience*, **57**, 1–28.

Döscher, R. C., C. W. Böning, and P. Herrman, 1994: Resposne of circulation and heat transport in the North Atlantic to changes in forcing in northern latitudes: a model study, *J. Phys. Oceanogr.*, **24**, 2306–2320.

Eckmann, J.-P., and D. Ruelle, 1981: Roads to turbulence in dissipative dynamical systems, *Rev. Mod. Phys.*, **53**, 643–654.

Evensen, G., 1992: Using the extended Kalman filter with a multilayer quasi-geostrophic ocean model, *J. Geophys. Res.*, **97**, 17,905–17,924.

Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, *J. Geophys Res.*, **99 (C5)**, 10,143–10,162.

Evensen, G., D. Dee, and J. Schröter, 1998: Parameter estimation in dynamical models, *NATO ASI, Ocean Modeling and Parameterizations edited by E. P. Chassignet and J. Verron.*

Evensen, G., and P. J. van Leeuwen, 1996: Assimilation of Geosat altimeter data for the Agulhas Current using the ensemble Kalman filter with a quasigeostrophic model, *Month. Wea. Rev.*, **124**, 85–96.

Farge, M., 1992: Wavelet transforms and their applications to turbulence, *Ann. Rev. Fluid. Mech.*, **24**, 395–457.

Farrell, B. F., and P. J. Ioannou, 1996: Generalized stability theory. part I: Autonomous operators, *J. Atmos. Sci.*, **53**, 2025–2040.

Fox, A., K. Haines, B. de Cuevas, and D. Webb, 1998: Assimilation of TOPEX data into the *OCCAM model*, *Int. WOCE Newsletter*, **31**, 12–15.

Fu, L.-L., E. J. Christensen, C. A. Yamarone, M. Lefebvre, Y. Ménard, M. Dorrer, and P. Escudier, 1994: TOPEX/POSEIDON mission overview, *J. Geophys. Res.*, **99**, 24,369–24,381.

Gates, W. L., P. R. Rowntree, and Q.-C. Zeng, 1990: *Validataion of climate models*, Climate Change: The IPCC Scientific Assessment, Cambridge University Press, 93–130.

Gavart, M., and P. De Mey, 1997: Isopycnal EOFs in the Azores Current region : a statistical *tool* for dynamical analysis and data assimilation, *J. Phys. Oceanogr*, **27**, 2146–2157.

Giering, R., and T. Kaminski, 1998: Recipes for Adjoint Code Construction, in press ACM Trans. On Math. Software.

Gilbert, J. C., and C. Lemaréchal, 1989: Some numerical experiments with variable-storage Quasi-Newton algorithms, *Math. Program.*, **45**, 407–435.

Gould, W. J., 1985: Physical oceanography of the Azores front, *Progress in Oceanography*, **14**, 167–190.

Gouretski, V., and K. Jancke, 1998: A new world ocean climatology: objective analysis on neutral surfaces, *WHP-SAC Tech. Report No. 3, WOCE Report*, **256/17**.

Green, J. S. A., 1970: Transfer properties of the large-scale eddies and the general circulation of the atmosphere, *Q. J. R. Met. Soc.*, **96**, 157–417.

Griffies, S. M., and K. Bryan, 1997: Predictability of North Atlantic multi-decadal climate variability, *Science*, **275**, 181–184.

Han, Y.-J., 1984: A numerical world ocean general circulation model part II. a baroclinic experiment, *Dyn. Athmos. Oceanogr.*, **8**, 141–172.

Hasselmann, K., 1976: Stochastic climate models, *Tellus*, **5**, 473–485.

Hellerman, S., and M. Rosenstein, 1983: Normal monthly windstress over the world ocean with error estimates, *J. Phys. Oceanogr.*, **13**, 1093–1104.

Hersbach, H., 1997: The adjoint of the WAM model, *KNMI Scientific Report*.

Holland, W. R., 1978: The role of mesoscale eddies in the general circulation of the ocean – Numerical experiments using a wind-driven quasi-geostrophic model, *J. Phys. Oceanogr.*, **8**, 363–392.

Holloway, G., and M. C. Hendershott, 1977: Stochastic closure for nonlinear Rossby waves, *J. Fluid Mech.*, **82**, 747–765.

Jung, T., E. Ruprecht, and F. Wagner, 1998: Determination of cloud liquid water path over the oceans from Special Sensor Microwave/Imager (SSM/I) data using neural networks., *J. Applied Met.*, **37**, 832–844.

Kaminski, T., M. Heimann, and R. Giering, 1996: Sensitivity of the seasonal cycle of CO2 at remote monitoring stations with respect to seasonal surface exchange fluxes determined with the adjoint of an atmospheric transport model, *Physics and Chemistry of the Earth*, **21(5-6)**, 457–462.

Killworth, P., C. Dietrich, J.-M. Molines, A. Oschlies, and J. Willebrand, 1999: Assimilation of altimetric data into an eddy-permitting model of the North Atlanic, *Progr. Oceanogr. (submitted)*.

Krauss, W., and R. H. Käse, 1984: Mean circulation and eddy kinetic energy in the eastern North Atlantic, *J. Geophys. Res.*, **89**, 3407–3415.

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

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

Lea, D. J., M. R. Allen, and T. W. N. Haine, 1999: Sensitivity analysis of the climate of a chaotic system, *Tellus (submitted)*.

Levitus, S., 1982: Climatological atlas of the world ocean, *NOAA Tech. Paper*, 173 pp.

Li, Y., 1991: A note on the uniqueness problem of variational adjustment approach to four-dimensional data assimilation, *J. Met. Soc. Japan*, **69**, 581–585.

Liptser, R. S., and A. N. Shiryayev, 1974: *Statistics of Random Processes I & II*, Springer.

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

Lorenz, E. N., 1975: Climate predictability: The physical basis of climate modeling., *WMO, GARP Pub. Ser.*, **16**, 132–136.

Marotzke, J., and C. Wunsch, 1993: Finding the steady state of a general circulation model through data assimilation: Application to the North Atlantic Ocean, *J. Geophys. Res*, **98**, 20,149–20,167.

Miller, R. N., E. F. Carter, and S. T. Blue, 1999: Data assimilation into nonlinear stochastic models, *Tellus*, **51 A**, 167–194.

Miller, R. N., M. Ghil, and F. Gauthiez, 1994: Advanced data assimilation in strongly nonlinear dynamical systems, *J. Atmos. Sci.*, **51**, 1037–1056.

Moore, A. M., 1991: Data assimilation in a quasi-geostrophic open-ocean model of the Gulf Stream region using the adjoint method, *J. Phys. Oceanogr.*, **21**, 398–427.

Morris, C., and S. Gill, 1994: Evaluation of the TOPEX/POSEIDON altimeter system over great lakes, *J. Geophys. Res.*, **99**, 24,527–24,539.

Morrow, R., and P. De Mey, 1995: Adjoint assimilation of atlimetric, surface drifter, and hydrographic data in a quasigeostrophic model of the Azores Current, *J. Geophys. Res.*, **100**, 25,007–25,025.

Nese, J. M., J. A. Dutton, and R. Wells, 1987: Calculated attractor dimensions for low-order spectral models, *J. Atmos. Sci.*, **44**, 1950–1972.

Oschlies, A., and V. Garçon, 1998: Eddy induced enhancement of primary production in a model of the North Atlantic Ocean, *Nature*, **394**, 266–269.

Oschlies, A., and J. Willebrand, 1996: Assimilation of Geosat altimeter data into an eddy-resolving primitive equation model of the North Atlantic Ocean, *J. Geophys. Res.*, **101**, 14,175–14,190.

Oseledec, V. I., 1968: A multiplicative ergodic theorem. ljapunov characteristic numbers for dynamical systems., *Trans. Moscow Math. Soc.*, **19**, 197–231.

Pacanowski, R., K. Dixson, and A. Rosati, 1993: *The GFDL modular ocean model users guide, Tech. Rep. 2,*, Tech. rep., Geophy. Fluid Dyn. Lab., Ocean Group, Princeton Univ., Princeton, N. J.

Palmer, T. N., 1993: Extended-range atmospheric prediction and the Lorenz model, *Bul. Am. Met. Soc.*, **74**, 49–65.

Pires, C., R. Vautard, and O. Talagrand, 1996: On extending the limits of variational assimilation in nonlinear chaotic systems, *Tellus*, **48 A**, 96-121.

Podesta, G. P., S. Shenoi, J. W. Brown, and R. H. Evans, 1995: AVHRR Pathfinder Oceans Matchup Database 1985–1993 (Version 18), *Report from the University of Miami, Rosenstiel School of Marine and Atmospheric Science 33pp.*

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, 1993: *Numerical Recipes*, Cambridge University Press.

Rabier, F., J. N. Th'epaut, and P. Courtier, 1998: Extended assimilation and forecast experiments with a four-dimensional variational assimilation system, *Q. J. R. Met. Soc.*, **124**, 1861–1887.

Richardson, P. L., 1983: Eddy kinetic energy in the North Atlantic from surface drifters, *J. Phys. Oceanogr.*, **88**, 4355–4367.

Schiller, A., and J. Willebrand, 1995: The mean circulation of the Atlantic Ocean north of 30s determined with the adjoint method applied to an ocean general circulation model, *J. Marine Res.*, **53**, 453–497.

Schröter, J., U. Seiler, and M. Wenzel, 1993: Variational assimilation of Geosat data into an eddy resolving model of the Gulf Stream extension area, *J. Phys. Oceanogr.*, **23**, 925–953.

Singh, S., and K. A. Kelly, 1997: Monthly maps of sea surface hight in the North Atlantic and zonal indices for the Gulf Stream using TOPEX/Poseidon altimeter data., *Woods Hole Oceanographic Institution Technical Report, WHOI-97-06, 50pp.*

Sirkes, Z., and E. Tziperman, 1997: Finite difference of adjoint or adjoint of finite difference?, *Mon. Wea. Rev.*, **125**, 3373–3378.

Smith, E., J. Vazquez, A. Tran, and R. Sumagaysay, 1996: Satellite-derived sea surface temperature data available form the NOAA/NASA Pathfinder program, *http://www.agu.org/eos_elec/95274e.html*.

Stammer, D., 1997: Global characteristics of ocean variability estimated from regional TOPEX/POSEIDON altimeter measurements, *J. Phys. Oceanogr.*, **27**, 1743–1768.

Stammer, D., C. Wunsch, R. Giering, Q. Zhang, J. Marotzke, J. Marshall, and C. Hill, 1997: The global ocean circulation estimated from TOPEX/POSEIDON altimetry and the MIT general circulation model, *Technical Report Nr.49, MIT, Center for Global Change Science.*

Stensrud, D. J., and J. W. Bao, 1992: Behaviors of variational and nudging techniques with a chaotic low-order model, *Mon. Wea. Rev.*, **120**, 3016–3028.

Stone, P. H., 1972: A simplified radiative-dynamical model for the static stability of rotating atmospheres, *J. Atmos. Sci.*, **29**, 405–418.

Swanson, K., R. Vautard, and C. Pires, 1998: Four-dimensional variational assimilation and predictability in a quasi-geostrophic model, *Tellus*, **50 A**, 369–390.

Sy, A., 1988: Investigation of large scale circulation patterns in the central North Atlantic: the North Atlantic Current, the Azores Current, and the Mediterranean Water plume in the area of the Mid-Atlantic Ridge., *Deep-Sea Res.*, **35**, 383–413.

Tanguay, M., P. Bartello, and P. Gauthier, 1995: Four-dimensional assimilation with a wide range of scales, *Tellus*, **47A**, 947–967.

Thacker, W. C., 1986: *Relationships between statistical and deterministic methods of data assimilation*, Variational Methods in Geosiences; Ed. Y. K. Sasaki, Elsevier, Amsterdam.

Thacker, W. C., 1988: Fitting models to inadequate data by enforcing spatial and temporal smoothness, *J. Geophys. Res.*, **93**, 10,655–10,665.

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

The DYNAMO Group, 1997: *DYNAMO Dynamics of North Atlantic Models: Simulation and assimilation with high resolution models*, Berichte aus dem IFM Kiel Nr. 294.

Traon, P. L., F. Nadal, and N. Ducet, 1998: An improved mapping method of multi-satellite altimeter data, *J. Atm. Ocean. Techn*, **25**, 522–534.

Treguier, A. M., I. Held, and V. Larichev, 1997: Parametrization of quasigeostrophic eddies in primitive equation ocean models, *J. Phys. Oceanogr.*, **27**, 567–580.

Tziperman, E., and W. Thacker, 1989: An optimal-control/adjoint equations approach to studying the oceanic general circulation, *J. Phys. Oceanogr.*, **19**, 1471–1485.

Tziperman, E., W. Thacker, R. Long, S.-M. Hwang, and S. Rintoul, 1992: A North Atlantic inverse model using an oceanic general circulation model, *J. Phys. Oceanogr.*, **22**, 1458–1485.

Visbeck, M., J. Marshall, T. Haine, and M. Spall, 1997: Specification of eddy transfer coefficients in coarse resolution ocean circulation models, *J. Phys. Oceanogr.*, **27**, 381–402.

Vogeler, A., and J. Schröter, 1995: Assimilation of satellite altimeter data into an open ocean model, *J. Geophys. Res.*, **100**, 15,951–15,963.

Yu, L., and P. Malanotte-Rizzoli, 1998: Inverse modeling of seasonal variations in the North Atlantic Ocean, *J. Phys. Oceanogr.*, **28**, 902–922.

# Acknowledgments

I would like to thank my supervisor, Prof. Dr. Jürgen Willebrand, who inspired this work for his guidance and support as well as the independence he confidentially allowed me. Thanks to Dr. J. Kielmann and the assistants of the computer center at the Institut für Meereskunde and at Christian-Albrechts Universität Kiel for service and expert advice. J. Schöter and A. Vogeler kindly provided the QG model and its adjoint and Dr. A. Oschlies provided his version of the CME model and the maps of satellite altimeter data. I especially would like to thank him and my office colleague Dr. N. Rix for many helpful advises. Finally I wish to thank all colleagues at the department of Theoretical Oceanography for interesting discussions and colleagues at the department of Regional Oceanography who allowed me to participate at an Indian Ocean cruise.