Nansen Environmental and Remote Sensing Center

A non-profit environmental research center affiliated with the University of Bergen

Thormølensg. 47
N-5006 Bergen, Norway
Tel: +47 55 20 58 00
Fax: +47 55 20 58 01

NERSC technical report 325

Forecasting non-linear systems with the Ensemble Kalman Filter and related data assimilation methods (eVITA-EnKF)

by

Bertino L. [1], G. Evensen [1,4], P. Sakov[1], E. Simon [1], F. Counillon [1], J. Bojarova [2], I. Lie [3], D. Kalise [3], A. Seiler [1,4], S. Kvamsdal [5], H. Wackernagel [1,6]

1. Mohn-Sverdrup Center (MSC) / NERSC, Bergen.
2. Norwegian institute for weather prediction (met.no), Oslo.
3. Storm Weather AS, Bergen/Oslo.
4. Statoil AS, Bergen.
5. Norwegian school of business and economic management (NHH), Bergen.
6. Center for Geostatistics, Ecole des Mines de Paris, Fontainebleau, France.

This work is a deliverable to the RCN under contract 178894 of the eVITA programme

September 30, 2011

Nansen Environmental and Remote Sensing Center (NERSC)
Thormølensgate 47
N-5006 Bergen
Norway

phone +47 55205800
fax +47 55205801
email admin@nersc.no

TITLE	**REPORT IDENTIFICATION**
eVITA EnKF Forecasting non-linear systems with the EnKF and related data assimilation methods Sept. 2010 to Sept. 2011	325 4th year report
CLIENT	**CONTRACT NO. 178894**
The Research Council of Norway (eVITA programme)	
CLIENT REFERENCE	**AVAILABILITY**
Gudmund Høst, RCN	Public
AUTHORS	**AUTHORIZATION**
Laurent Bertino and the eVITA EnKF consortium	Bergen, September 30, 2011 Prof. Stein Sandven Director NERSC

Contents

Forecasting non-linear systems with the Ensemble Kalman Filter and related data assimilation methods (eVITA-EnKF)

Bertino L. [1], G. Evensen [1,4], P. Sakov[1], E. Simon [1], F. Counillon [1], J. Bojarova [2], I. Lie [3], D. Kalise [3], A. Seiler [1,4], S. Kvamsdal [5], H. Wackernagel [1,6]

1. **Mohn-Sverdrup Center (MSC) / NERSC, Bergen.**
2. **Norwegian institute for weather prediction (met.no), Oslo.**
3. **Storm Weather AS, Bergen/Oslo.**
4. **Statoil AS, Bergen.**
5. **Norwegian school of business and economic management (NHH), Bergen.**
6. **Center for Geostatistics, Ecole des Mines de Paris, Fontainebleau, France.**

Introduction

The present document reports the work carried out under the fourth and final year of the project "Forecasting non-linear systems with the EnKF and related data assimilation methods (eVITA-EnKF)" in the period October 2010 - September 2011. The first section describes transversal methodological developments and each following section describes an application from each of the partners. Perspectives of continuation are summarized in the end.

1 Methodological developments

This section contains a brief list of the methodological results and activities in the frame of the eVITA project which considered the iterative solutions of the EnKF (Iterated EnKF) and geostatistical perspectives. The work has been performed on analytical solutions and idealized models and lead to the following communications.

Publications

- P. Sakov, D. Oliver and L. Bertino, *An iterative EnKF for strongly nonlinear systems*, Monthly Weather Review, *Submitted*, 2011.

  This study considers an iterative formulation of the ensemble Kalman filter (EnKF) for strongly nonlinear systems in the perfect-model framework. In the first part a scheme is introduced that is similar to the ensemble randomised maximal likelihood (EnRML) filter by Gu and Oliver (2007). There are two new elements in the scheme. Firstly, it uses the ensemble square root filter for better optimality instead of the traditional (perturbed observations) EnKF. Secondly, it uses rescaling of the ensemble anomalies with the ensemble transform matrix from the previous iteration, instead of estimating the sensitivities by linear regression between the ensemble observations and ensemble anomalies at the start of the assimilation cycle. A simple modification turns the scheme into an ensemble formulation of the iterative extended Kalman filter (EKF). We refer to the two versions of the algorithm as the iterative EnKF (IEnKF), and the iterative EKF (IEKF). In the second part we test the performance of the IEnKF and IEKF in five numerical experiments; two with the 3-element Lorenz models and three with the 40-element Lorenz model. Both the IEnKF and IEKF show a considerable advantage over the EnKF in strongly nonlinear systems when the quality or density of observations are sufficient to constrain the model to the regime of mainly linear propagation of the ensemble anomalies as well as constrain the fast growing modes, and a much smaller advantage otherwise. Like the EnKF, the

IEnKF and IEKF can be used with large-scale models, and can be viewed as an alternative to the particle filter (PF) and hybrid PF-EnKF schemes in strongly nonlinear systems.

- H. Wackernagel and L. Bertino, *Localization, or how to speed up the Kalman filter update without spoiling the covariance structure (a geostatistical perspective).*

  Hans Wackernagel and Laurent Bertino have reviewed the geostatistical literature about kriging with large multivariate datasets and are examining several possibilities of implementing geostatistical findings in the context of sequential data assimilation.

  There is a growing literature in geostatistics about the problem of kriging with large datasets, which bears similarities with the discussion about localisation in sequential data assimilation. In both cases, the aim is to reduce the problem dimension of the prediction/forecasting by simplifying the kriging/update step, while preserving the properties of the statistical/physical model describing the phenomenon.

  In geostatistics, as compared to data assimilation, the discussion is more easy to lead as a statistical model of the coregionalisation is explicitly formulated and, in particular, the covariance structure of the multivariate space-time data is explicitly modeled with direct and cross-covariance functions, yielding a coregionalisation model. Studying the structure of the coregionalisation may then suggest simplifications in the cokriging equations and also provide guidelines for specifying an appropriate moving neighborhood. In data assimilation, the various receipts proposed for speeding up the update step in the Kalman filter are more difficult to evaluate and to compare in terms of knowing whether the properties of the physical system at hand are preserved, because empirical, non-stationary covariances are used and, furthermore, these covariances are error covariances.

  First results have been presented recently in an oral presentation at the 6th International EnKF Workshop workshop and a corresponding journal paper is in preparation.

Conferences, Outreach

- P. Sakov has co-arranged jointly with CIPR and IRIS, the *6th international workshop on ensemble Kalman filtering 20-22 May, 2011. Brakanes Hotel, Ulvik, Norway.*

- M. Gray from the Laboratoire d'Astrophysique de Marseille (LAM), has visited the MSC in Bergen from 31st January to 4th February for initial implementation and testing of an Ensemble Transform Kalman Filter in the Optimal control of the adaptive optics of the European Extremely Large Telescope.

- P. Sakov has given 5 lectures (8 hours) at the NORDITA school on Data Assimilation, Stockholm, 25th to 29th April 2011.

- L. Bertino has given 3 lectures on operational ocean forecasting and data assimilation on a workshop 16th to 18th May 2011, at INCOIS, Hyderabad, India.

- P. Sakov, D. Oliver and L. Bertino, *Iterative EnKF for strongly nonlinear systems* talk at the 6th International EnKF Workshop, Ulvik.

- H. Wackernagel and L. Bertino, *Localization, or how to speed up the Kalman Filter update without spoiling the covariance structure*, talk at the 6th International EnKF Workshop, Ulvik.

- The eVITA-EnKF meeting on 16th-17th May was receiving Dr. Svetlana Losa (Alfred Wegener Institute, De) in Bergen as guest lecturer.

- F. Counillon has given 3 tutorials on data assimilation in ocean models at the Nansen-Tutu Winter School, Cape Town, South Africa, June 2011.

- P. Sakov has given 4 tutorials on the EnKF at the 2nd Summer School on Data Assimilation, Iasi, Romania, 25th to 29th July 2011.

- L. Bertino has given 3 lectures (4 hours) on applications of the EnKF in oceanography at the 2nd Summer School on Data Assimilation, Iasi, Romania, 1st-4th August 2011.

- Continued development and distribution of the EnKF-Matlab package, available from the EnKF home page:
  `http://enkf.nersc.no/Code/EnKF-Matlab`
- The eVITA-EnKF news and presentations are posted on the project home page:
  `http://msc.nersc.no/?q=evita-enkf`

2 Physical oceanography - MSC

2.1 Context: The Arctic Ocean reanalysis

The Pilot Reanalysis (2003-2008) has been completed in Year 3 and served as part of the MyOcean V1 product portfolio `http://www.myocean.eu.org` and a scientific publication is in the review process for the AOMIP (Arctic Ocean Model Intercomparison Project) special edition of the Journal of Geophysical Research. Meanwhile, the full 20-years reanalysis is being integrated on the NOTUR facilities at Parallab, Bergen, and a preliminary set of data for the years 1991-1998 has been provided on the MyOcean public server. The full Reanalysis has taken up all the methodological developments from last year.

Following last year's findings on the TOPAZ Pilot Reanalysis, F. Counillon has started a more systematic evaluation of the bias estimation procedure with an idealized model being afflicted of SSH bias.

2.2 Estimation of a bias in sea surface heights

The Ensemble Kalman Filter assumes that the error made in the model is Gaussian and non biased. The latter assumption is often not satisfied in realistic ocean models. Ignoring model biases can have negative impact in combination with data asimilation for two reasons: first, as the optimal estimate is a combination of model and observations, the final estimate will still contain a part of bias. Second and probably more problematic, the models are often attracted to a biased solution and this also affects the ensemble estimates of error covariances. Thus, through successive assimilation increments the erroneous updates (extrapolation in space or updates of non-observed variables) will pile up in unpredictable ways. To circumvent this problem, a common approach in data assimilation is to remove an estimate of the bias prior to assimilation, with a bias-corrected forecast. Thus, the average estimate of the analysis and forecast have the correct mean, and the observation is not used for correcting the bias but rather for correcting the model anomaly. The EnKF allows for estimating this bias online in the same way as for parameter estimation (3): through successive iteration the bias estimate will converge to the values that minimise the error, as any other parameter. This technique was adopted in TOPAZ system to handle for bias of Sea Surface Height and Sea Surface Temperature. It proved relatively successful, as the resulting model bias decreased, and the estimated bias could be interpreted against the known deficiencies of the model. However, in places where no observations are not assimilated (or where the model bias is related linearly to observations), spurious features appear, similar to random red noise with mesoscale decorrelation length scale (Figure 1). An idealised study was set up to understand the reason for the development of these features and improve the bias estimation procedure.

The model used is the quasi-geostrophic model available within the EnKF-matlab tool box that can mimic the behaviour of mesoscale circulation. Two different approaches were tested: the classical model error estimation suggested in (4) (referred to as Method 1 in the following) and the bias inflation method used in TOPAZ system (referred to as Method 2 in the following). These two methods differ by the way the ensemble is initialised and the ensemble spread is maintained. The finite ensemble size and the associated sampling errors tend to a systematic reduction of the ensemble spread that must be maintained high enough for the bias parameters to converge. The first method relies on additive red noise increments for the bias (correlated in space and time), while the second uses inflation of bias anomalies. In addition, the ensemble in the first method is initialised with random red noise while in the second method it is simply uses uniform values. Note that the length scales of the red noise is an arbitrary parameter in the first method. The two methods are tested for estimation of a constant bias (red noise with known or unknow length scales) assimilating an observational network that cover fully or partially the model domain, in the intention to mimic different situations occurring with the TOPAZ system. With a uniform observation coverage, it appears that the two methods succeed. Method 2 provides slightly better estimates than Method 1, and the ensemble initialisation has almost no impact because the final scale of the model bias estimate depends

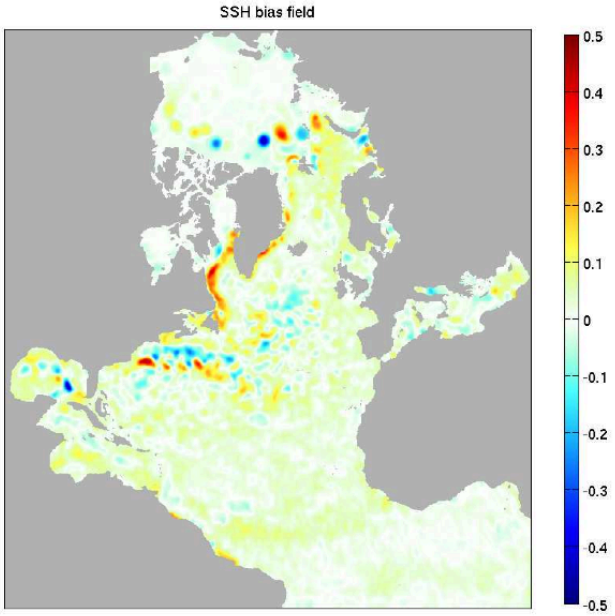

Figure 1: Example of SSH bias estimated from the TOPAZ Pilot Reanalysis

more on the localisation radius. It is also found that with Method 1, the length scale of red noise should rather be set too large than too short when it is poorly known. With an observation network that covers only part of the model domain, Method 1 is still succesfull while Method 2 causes spurious features of large amplitude in the non-observed part of the domain, causing the model to crash. Small changes were proposed to Method 2 based on the adaptive inflation method proposed by (1). Using this small correction, Method 2 provides a slightly better accuracy than Method 1 and without spurious features. This work is intended to be submitted in a scientific journal.

2.3 Extension to P. R. China

This subsection relates to the follow-up of the BILAT extension obtained during Year 2, which has lead to a publication by Xie, Counillon, Bertino and Zhu is now accepted in Ocean Science (8), demonstrating the possibility of assimilation of altimeter data into a 3D oecan model resolving both mesoscale currents and tides.

Dr. Jiping Xie has visited NERSC for 3 months in Fall 2010, this time funded by the SeaFINE Joint Industry Project. This visit has produced a 1-year test run of the 5 km resolution HYCOM model of the South China Sea with assimilation of altimeter data, using the Ensemble Optimal Interpolation (EnOI) in similar settings to the results published in Ocean Science, but at a higher horizontal resolution (See Figure 2). This upgraded model resolves better the detailed ocean mesoscale processes and the bathymetric features of interest to the oil and gas industry. In April 2011, the SeaFINE JIP has approved the 18-years reanalysis of the South China Sea during the London OGP meeting. The SeaFINE JIP represents a group of 15 major oil and gas operators, capturing most of the main actors in this region.

The reanalysis period of 18 years will use the whole altimeter record since the launch of the first altimeter Topex/Poseidon in 1992. The duration of 18 years is a full tidal epoch, necessary for design criteria in the tidal-dominated South China Sea. The database will upgrade a previous hindcast carried out by NERSC in 2007, which was using a coarser 12 km nested HYCOM model and ran without data assimilation. The SeaFINE JIP (similarly to the work initiated in Year 3) contracted Terra Orbit AS, a private company owned by the Nansen Center with offices in Bergen, Norway.

The simulation will produce 18-years full-depth time series in almost twenty thousand locations and will be validated against in situ measurements from moorings all around the South China Sea during the same period. Building on the success of the South China Sea simulations, similar hindcast products are envisioned for different oceans of the world. The full hindcast simulation is expected for the OGP meeting in March 2012.

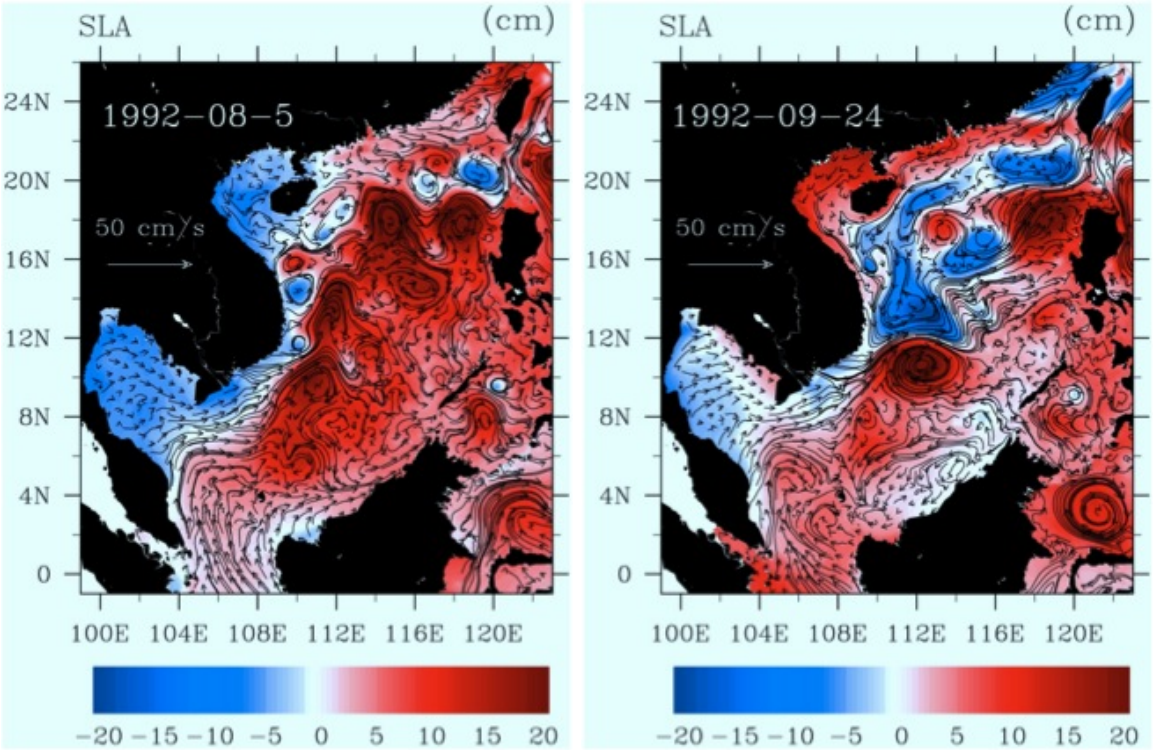

Figure 2: Surface currents and sea surface heights on two typical days in 1992 simulated with the HYCOM 5 km model, showing clear mesoscale dynamic features. Courtesy of Dr. Jiping Xie, IAP, Chinese Academy of Science, Beijing while visiting the MSC.

Publications

- P. Sakov, F. Counillon, L. Bertino, K.A. Lisæter, P. Oke, A. Korablev, *TOPAZ4: an ocean-sea ice data assimilation system for the North Atlantic and Arctic*, J. Geophys. Res., Submitted.
- J. Xie, F. Counillon, L. Bertino and J. Zhu *An eddy resolving tidal-driven model of the South China Sea assimilating along-track SLA data using the EnOI*, Ocean Sci. Discuss., 8, 873-916, doi:10.5194/osd-8-873-2011. Accepted for publication.
- Srinivasan, A, Chassignet, EP, Bertino, L, Brankart, JM, Brasseur, P, Chin, TM, Counillon, F, Cummings, J, Mariano, AJ, Smedstad, OM and Thacker, WC, *A comparison of sequential assimilation schemes for ocean prediction with the HYbrid Coordinate Ocean Model (HYCOM): Twin experiments with static forecast error covariances*, Oc. Modeling, 37(3-4), pp. 85-111, 2011.
- L. Bertino, *Data assimilation in oceanography, capabilities and challenges*, Proceedings of the joint Nansen-Tutu Scientific opening Symposium and OceansAfrica meeting. Nansen Tutu Center for Marine Environmental Research. Cape Town, South Africa. July 2011.

Conferences

- L. Bertino, *Capabilities and challenges of data assimilation*, presentation at the Nansen-Tutu opening symposium, 8th December 2010, Cape Town, South Africa.
- P. Sakov, L. Bertino and F. Counillon, *TOPAZ4: an ocean-sea ice forecasting system for the North Atlantic and Arctic* Seminar 25th Jan 2011 at CMAR, Hobart, Australia.
- L. Bertino, P. Sakov, F. Counillon, A. Samuelsen, B. Hackett, A. Burud, H. Wehde and V. Volkov, *The MyOcean Arctic Forecasting Center: a new pan-European service*, Arctic Frontiers, 28th January 2011, Tromsø.

- P. Sakov, L. Bertino and F. Counillon *The TOPAZ4 Pilot Reanalysis (2003-2008)*, talk at the EGU, 4th-9th April 2011, Vienna, Austria.
- L. Bertino, P. Sakov and F. Counillon *The TOPAZ4 Pilot Reanalysis (2003-2008)*, talk at the 6th EnKF workshop, Ulvik, Norway. 20th June 2011.
- P. Sakov, L. Bertino, F. Counillon and P. Oke, *Metrics for quantifying observation impact in data assimilation: application to the TOPAZ pilot reanalysis*, Presentation at the GODAE/OceanView GSOP/CLIVAR workshop, Santa Cruz, USA, 15th June 2011.
- P. Sakov, F. Counillon and L. Bertino *Status of the main TOPAZ4 reanalysis (1991-2010)*, presentation at the MyOcean WP05 meeting, IMR, Bergen.

Other

- The TOPAZ4 Pilot Reanalysis is part of the MyOcean V1 service launched on the 15th December 2010 and presented to the user community during the User Workshop in Stockholm, 7th and 8th April 2011.

3 Biological oceanography - MSC

This document is a brief description of the work realized on the eVITA EnKF project by E. Simon in period October 2010 - December 2011.

1. Reanalysis of the biogeochemical component of the North Atlantic and Arctic oceans

   In the prospect of a reanalysis of the biogeochemistry component of the Arctic ocean for the period 2008-2010, an experiment of combined state-parameter estimation is conducted in a North Atlantic and Arctic configuration of the HYCOM-NORWECOM coupled model. The main task consisted in the development of an assimilation system based on the core of the TOPAZ4 forecasting system. Beside the update of the two components of the coupled model, a deterministic ensemble Kalman filter (DEnKF) and its log-transformed extension have been implemented to assimilate both physical and biological observations. The system assimilates observations in two steps every 8 days. In the first one, only surface satellite physical data (SST, TSLA and ice concentration) are assimilated in the physical component of the coupled model (HYCOM). For each member, the biogeochemical variables are then interpolated on the analyzed vertical grid in order to maintain the conservation of tracer quantities. In a second step, satellite surface chlorophyll-a data are assimilated in the biogeochemical component of the coupled model (NORWECOM). For that step, variables, parameters to calibrate and observations are log-transformed in order to handle issues that must raise due to the positiveness of the variables and parameters. The system is almost finalized and the reanalysis is expected before the end of 2011. The reanalysis will contribute to the V2 public service of the European MyOcean project.

2. Application of model reduced 4D variational data assimilation to ecosystem models

   This work has been realized during the 3-month stay of J. Pelc (doctoral student from TU Delft, Netherlands) at MSC. The aim of this study was to investigate the application of model reduced 4D-Var algorithms to ocean ecosystem models. Twin experiments have been realized in a simple 1D ocean ecosystem model in the prospect of parameters calibration. Two types of control have been considered: the control of the parameters to estimate and the combined control of the initial condition and parameters. The study highlighted the effectiveness of the model reduced 4D-Var to calibrate biased parameters for both types of control.

3. Estimation of sum-to-one constrained parameters with the EnKF: application to the zooplankton grazing preferences

The aim of this study is to demonstrate that ensemble-based Kalman filters can be applied to estimate sum-to-one constrained parameters. More specifically, we focus on ecosystem models and the estimation of the grazing preferences parameters of zooplankton species. These parameters are introduced to model the diet of zooplankton species among phytoplankton species and detritus. They are positive values and their sum is equal to one. Because the sum-to-one constraint cannot be handled by ensemble-based Kalman filters, a reformulation of the parameterization is proposed. We investigated two types of changes of variables for the estimation of sum-to-one constrained parameters that result in the estimation of normal or bounded distributed parameters. This led to the development of an ensemble-based Kalman filtering system, including the Gaussian anamorphosis, for a 1D configuration of the Global Ocean Turbulence Model (GOTM) coupled with the ecosystem model NORWECOM. This task has been done in collaboration with A. Samuelsen from MSC and D. Dumont from ISMER (Canada). Twin experiments that are currently realized with this system tend to demonstrate the effectiveness of these approaches to estimate sum-to-one constrained parameters like zooplankton grazing preferences.

Publications

- E. Simon and L. Bertino, Gaussian anamorphosis extension of the DEnKF for combined state and parameter estimation: application to a 1D ocean ecosystem model. *Journal of Marine Systems*, 89, 1-18, 2011.
- J. S. Pelc, E. Simon, L. Bertino, G. El Serafy and A.W. Heemink, Application of a model reduced 4D-Var to a 1D ecosystem model. In preparation.
- E. Simon, A. Samuelsen, L. Bertino and D. Dumont, Estimation of sum-to-one constrained zooplankton grazing preferences with the DEnKF: twin experiments. In preparation.
- Milutinovic, S, Bertino, L, *Assessment and propagation of uncertainties in input terms through an ocean-color-based model of primary productivity.* Remote Sensing of Environment . Vol.: 115, pp. 1906-1917, 2011.

Communications

- L. Bertino, E. Simon and A. Samuelsen, *Development of data assimilation in the TOPAZ-ECO system*, presentation at the Kick-Off meeting of the ESA Climate Change Initiative Ocean Colour project. Plymouth, UK, 31st Aug 2010.
- E. Simon and L. Bertino, *MyOcean Science Days*, Non-Gaussianity and biased parameter estimation of an ocean biological system with the EnKF, 1-2 December 2010, MyOcean science days, Toulouse, France.
- E. Simon and L. Bertino, *Seminar at Geofysisk institutt*, Observation-based parameter estimation in a 3D physical-biogeochemical model, 21 February 2011, Bergen, Norway.
- J. S. Pelc, E. Simon, L. Bertino, G. El Serafy and A.W. Heemink, *The Ninth International Workshop on Adjoint Model Application in Dynamic Meteorology*, Model-reduced 4D-Var data assimilation in application to 1D ecosystem model, 10-14 October 2011, Cefalu, Italy.
- E. Simon, L. Bertino and A. Samuelsen, *Estimation of sum-to-one constrained parameters with non-Gaussian extensions of ensemble-based Kalman filters : application to a 1D ocean biogeochemical model*, presentation at the MyOcean Arctic MFC meeting, IMR, Bergen, 9th September 2011.
- E. Simon, L. Bertino and A. Samuelsen, *AGU Fall meeting 2011*, Estimation of sum-to-one constrained parameters with non-Gaussian extensions of ensemble-based Kalman filters : application to a 1D ocean biogeochemical model, abstract accepted for Poster presentation at the AGU, 5 - 9 December 2011, San Francisco, USA.

4 Weather forecasting - met.no

Brief description of the work conducted under the eVITA-EnKF project at met.no (October 2010-May 2011). The following tasks where addressed

1. As it was decided earlier, the ETKF rescaling scheme was extended to allow the ATOVS radiances. The idea behind this extension was that the introduced ATOVS radiances will make the observation framework more uniform for different assimilation times which could lead to a positive impact on the performance of the variance inflation. However, the desirable effect was not obtained. To understand the reason we have investigated the model response versus data assimilation system response (conventional observations + ATOVS radiances) for HIRLAM forecasting system (the normalized ensemble spread versus normalized squared innovation for different types of observations and different assimilation cycles). The variability of the observation network for different assimilation cycles (00, 06, 12 and 18 UTC) as well as the relative size of each observation network were investigated as well. It turned out that the vertical variability of the HIRLAM forecasting system seems to be very small. This feature reduces impact of the ATOVS radiances on the performance of the ETKF rescaling scheme. At the same time the HIRLAM forecasting system seems to be very sensitive to the surface pressure (over land) disturbances. We have noticed as well that the assumed observation error variance is not always consistent with the "observed " observation error variance. At the same time the performance of the ETKF rescaling scheme is very sensitive to the assumptions on the observation error variance, because the ETKF scheme works in the observation space normalized with the assumed observation error variance. The main conclusion of this investigation was that the assumed observation error variance for AIREP, DRIBU, SHIP, and in particular ATOVS is unrealistically large. Some experiments were performed with the decreased observation error variance. According to the standard HIRLAM verification scores the reduction of observation error variance improved forecast performance, in particular for larger forecast length time. However, reducing observation error variance the efficient size of the observation network has been increased, and this lead to further increase of the inflation factor, amplifying the problem of rank deficiency.

2. The EDA (Ensemble of Data Assimilation runs) system for generation of the upper air perturbations was developed as an alternative to the ETKF rescaling scheme. The scheme is implemented into the HIRLAM forecasting system and is adopted to the hybrid settings. The ensemble of perturbations is obtained via assimilating the perturbed observations. The EDA scheme is consistent with the EnsDA schemes used at ECMWF and one used at Meteo-France. The DFS ("degree of freedom for signal"), a standard tool for investigating response of the observing system, can be efficiently computed using EDA. The estimation of the DFS from the ETKF rescaling system can be more questionable, but is important for comparison. The EDA system can be a useful device investigating the sensitivity of the data assimilation and the response of the forecasting system to perturbations of the particular signal.

3. A case study (for winter case: 17 January 2008- 3 March 2008) was conducted in order to illustrate performance of the hybrid ensemble variational data assimilations scheme in HIRLAM/hirlam-7.3 (using three different ensemble generation scheme: downscaling of EuroTEPS; ETKF; EDA). The hybrid ensemble variational data assimilation scheme gives improvement over the pure 3D-variational data assimilation scheme both for the winter and for the summer case using each of three upper air perturbation methods (downscalingof EuroTEPS, ETKF and EDA). The hybrid scheme using the ETKF rescaling scheme gives the most improvement, however it could depend on the fact that the ETKF rescaling scheme was the most elaborated and the best tuned among the others. The improvement in the RMSE scores is largest over Scandinavia, and in the most significant for the vertical profiles of wind-speed and humidity, for forecasting the surface pressure and of the cloud cover. The impact of the hybrid approach on the RMSE score of the vertical temperature profiles and for the precipitation seem to be neutral. The improvement originates from a more proper treatment of observations in the vicinity of fronts. Using hybrid scheme the analysis increments are anisotropic and dynamically consistent and follow clearly the front structures. Besides that under the hybrid ensemble variational data assimilation scheme (using anyone of the three ensemble generation schemes) the analysis increments show baroclinic structures. Baroclinic structures are very essential for a proper modelling of quickly developing low pressure systems (polar lows systems in this case study) but they are not captured at all in the background error covariance used in variational data assimilation.

Existence of baroclinic structures for analysis increment in the hybrid mode indicate clearly that the hybrid algorithm is able to extract useful flow-dependent information even from a heavily rank-deficient ensemble. However, we have noticed that the control run for all ensemble scheme, with or without the hybrid data assimilation, is significantly degraded compared to the deterministic 3D-Var (4D-Var) runs. There could be several reasons for degraded performance of the control forecast in the ensemble mode using EuroTEPS boundaries. First all these deficiencies are to be discovered and improved, only after this the final comparison of the hybrid ensemble variational data assimilation versus the 3d/4d-Var schemes will be documented.

4. A various system script improvements were performed in order to assure the efficient pre-operational flow of the hybrid ensemble variational data assimilation scheme for the HIRLAM forecasting system. The main implementation was the introduction of the high horizontal resolution (HR) operational lateral boundary conditions (LBC) when low resolution global (LR) system upper air ensemble of perturbations was used to generate Limited Area Model high horizontal resolution perturbations. The following scheme was implemented for generation of the lateral boundary conditions:

$$\mathrm{HRLBC}_i = \mathrm{HRLBC}_0 + (\mathrm{LRLBC}_i - \mathrm{LRLBC}_0)$$

5 Weather forecasting - Storm Weather AS

5.1 Introduction

This part of the project investigates the application of the EnKF on high-resolution ($\mathcal{O}(10^2)$ meters) atmospheric models.
The project is divided in roughly two parts,

1. Development of high order numerical schemes for full and simplified Euler equations.
2. Construction of, and numerical experiments with the EnKF in very high resolution atmospheric models.

In the modelling part, the focus has been on development and analysis of equations systems that enables us to experiment with EnKF with a reasonable amount of computational resources.

In the assimilation part, the emphasis is on model configurations with high resolution, a very limited number of, but frequent observations. This is important in applications as wind-power, but also in general in order to investigate the applicability of EnKF and other data assimilation methods to high resolution models.

Dante Kalise will submit his Ph.D. thesis in October 2011.

5.2 Modelling

The third year activities in modelling, which is work done by Ph.D student Dante Kalise, have focused on development of rather untraditional, higher-order numerical methods for atmospheric models.

The activities have been:

- Detailed study of the so-called 2.5D layered model, consisting in a layered coupled set of Euler equations where the vertical velocity appears as a source term providing coupling between both layers. Another approach to layering has also been investigated, namely the Galerkin semi-discretization by Temam et al.
- Using the higher-order Godunov type ADER method finite volume method with WENO reconstruction for the linearized Euler equations.
- Using a Runge-Kutta Totally Variation Diminishing method with WENO reconstruction on the full Euler equations.

For all these activities, there are at the time of writing 4 papers that will be included in Dante's Ph.D thesis. In addition to this, work on using higher-order methods with unstructured grids will be reported in the Ph.D thesis.

Results with the ADER method was presented at the "Workshop on Higher Order Nonlinear Numerical Methods for Evolutionary PDEs" in Trento in April 2011.

5.3 Data assimilation

In 2011 Astrid Holstad and Ivar Lie have been working construction of an EnKF system and on setting up assimilation experiments related to wind-power applications.

We have been using an extensive set of observations and runs from fine scale atmospheric models that were results from a BIP research project "Vindressurskartlegging i komplekst terreng" performed in the years 2006-2008 and sponsored by the Research Council of Norway and Hydro Oil & Energy. See (2) for a detailed description. It is to be emphasized that the observations and model runs form the BIP-project has been very important for this work.

We chose to develop an EnKF system based on the stochastic update algorithm ourselves since it is relatively straightforward when the observations are few and are of the same category. Moreover, it gives us a possibility to investigate the EnKF system in detail, which is important since data assimilation with EnKF with very high resolution models is quite new.

The atmospheric model we used is the SAM (Storm Airflow Model) model developed by Astrid Holstad. We ran the model with approximately 150 m resolution over an area covering Gurskøy in northwestern Norway, the site where the observation campaign.

The results of the data assimilation using only one observation site with wind observations, were rather surprising and in fact very good. EnKF was able to correct the fields in the whole model domain in a very physical way with only 4 members. This is, to our knowledge, the first time that EnKF has been used for such high resolution models, and the results are very promising for example in contexts of:

- "Distributing" model point corrections spatially. This is a very important topic in weather forecasting, and EnKF seems to be the only solution based on atmosphere dynamics (which is in fact how distributed corrections should be performed).
- Distributed validations. Another important topic in the evaluation of model results and forecasting quality.
- Forecasting of wind fields in a wind farm in complex terrain.

In addition to the numerical experiments without EnKF system, we have developed a method to compute the "size" of the error covariance matrices so that the performance of the EnKF can be inspected. The EnKF system and extensive numerical experiments are described in Holstad and Lie (2011).

As an example of the performance of the EnKF, see Figure 3, which shows wind profiles in the range $[0, 200]$ m above the terrain. The upper panel shows the results from the first assimilation cycle and the lower panel the results from the second cycle. The black profile is the initial conditions for the model runs, in the upper panel it is the initial condition from the outer model (with 1 km resolution), and in the lower panel it is the analysis from the first cycle. The ensemble mean of the model runs is the red profile and the analysis is the gray profile. In the case shown, the model initial condition give a wind speed that is much higher than the observation, so we see in the upper panel that the first cycle has a large increment. The analysis id in fact very close to the observation. The lower panel shows that the increment is much smaller (gray - red), but the ensemble still has a reasonable spread.

Publications

- A. Holstad and I. Lie, *An Ensemble Kalman filter data assimilation system for high resolution atmospheric models - implementation and numerical experiments*, StormGeo Technical Report, Bergen, June 2011.

6 Oil reservoir forecasting - MSC/Statoil

The PhD study of Alexandra Seiler was defended successfully on the 26th November 2010 against the opponents Prof. Peter Jan van Leeuwen (Uni. Reading, UK) and Dr. Remus Hanea (TNO, TU Delft, Nl) after only two years and a half of PhD. Another highlight of the 4th year of the eVITA project is the study by Skjervheim et al., that revives the Ensemble Smoother, a Monte Carlo method originally introduced by (7), but that now finds an appropriate field of application with history matching of oil reservoirs.

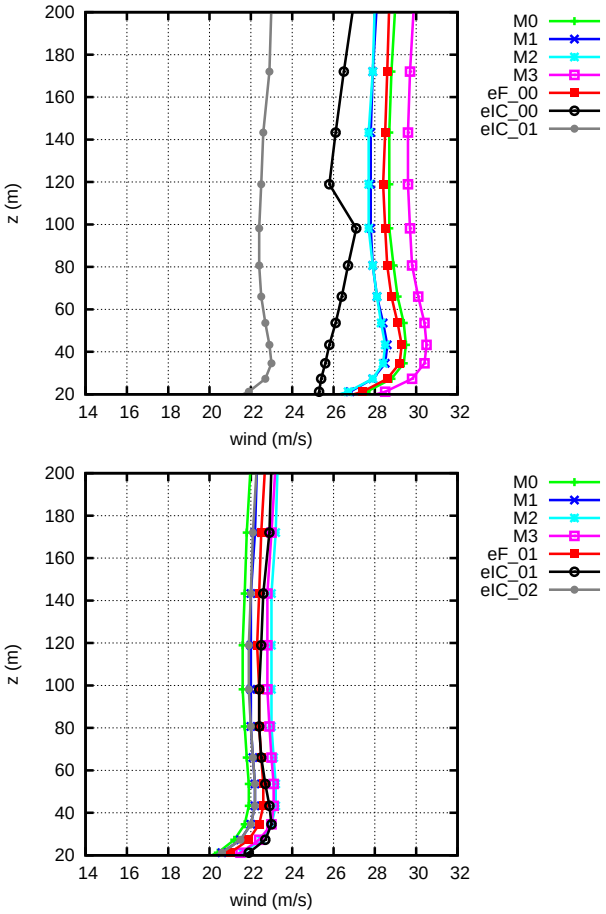

Figure 3: *Vertical profiles for two cycles of EnKF. For explanation of the profiles, see the text*

Publications

- Alexandra Seiler, *Reservoir structural model updating using the Ensemble Kalman Filter*, PhD, Mathematical Institute, Uni. of Bergen. 26th November 2010.

  In reservoir characterization, a large emphasis is placed on risk management and uncertainty assessment, and the dangers of basing decisions on a single base-case reservoir model are widely recognized. In the last years, statistical methods for assisted history matching have gained popularity for providing integrated models with quantified uncertainty, conditioned on all available data. Structural modelling is the first step in a reservoir modelling workflow and consists in defining the geometrical framework of the reservoir, based on the information from seismic surveys and well data. Large uncertainties are typically associated with the processing and interpretation of seismic data. However, the structural model is often fixed to a single interpretation in history-matching workflows due to the complexity of updating the structural model and related reservoir grid.

  This thesis present a method that allows to account for the uncertainties in the structural model and continuously update the model and related uncertainties by assimilation of production data using the Ensemble Kalman Filter (EnKF). We consider uncertainties in the depth of the reservoir horizons and in the fault geometry, and assimilate production data, such as oil production rate, gas-oil ratio and water-cut.

  In the EnKF model-updating workflow, an ensemble of reservoir models, expressing explicitly the model uncertainty, is created. We present a parameterization that allows to generate different realizations of the structural model to account for the uncertainties in faults and horizons and that maintains the consistency throughout the reservoir characterization project, from the structural model to the prediction of production profiles. The uncertainty in the depth of the horizons is parameterized as simulated depth surfaces, the fault position as a displacement vector and the fault throw as a throw-

scaling factor. In the EnKF, the model parameters and state variables are updated sequentially in time, as new measurements become available. Updates in the structural model impact the reservoir grid, and when the grid architecture is modified, all the cell-referenced and grid-region parameters need to be updated as well. The entire reservoir modelling workflow, from structural modelling to flow simulation, needs to be rerun, and thus, automated. Furthermore, a major constraint is that the current EnKF implementation considers a fixed dimension of the state vector, which implies a constant number of active cells in the reservoir grid. The requirements of an automated workflow and a fixed grid architecture leads to the proposed method, where the geometry of the base-case grid, representing the most likely interpretation, is deformed to match the different realizations of the structural model. In this project, grid deformation algorithms for updating the geometry of a corner-point grid have been developed and integrated in the EnKF model-updating workflow.

The proposed method for updating the structural-model uncertainties has been applied to synthetic cases and is implemented on real field cases. The result is an updated ensemble of structural models, conditioned to all production data, and with reduced and quantified uncertainty. The updated ensemble of structures provides a more reliable characterization of the reservoir architecture, including the top and bottom horizons and the fault geometry, and a better estimate of the field oil-in-place.

- A. Seiler, O. Lia and K.M Okstad, *Grid Updating for Fault Uncertainty Modelling*, Submitted to Journal of Petroleum Science and Engineering, August 2010.

- J.-A. Skjervheim, G. Evensen, J. Hove and J. G. Vabø, *An Ensemble Smoother for assisted History Matching*, SPE 414929, 2011.

  This paper compares two ensemble-based data-assimilation methods when solving the history-matching problem in reservoir- simulation models. The methods are the Ensemble Kalman Filter (EnKF) and the Ensemble Smoother (ES). Several publications have discussed the use of EnKF in petroleum applications while ES is now used for the first time for history matching. ES differs from EnKF by computing a global update in the space-time domain, rather than using recursive updates in time as in EnKF. Thus, the sequential updating of the realizations with associated restarts is avoided. EnKF and ES provide identical solutions for state estimation with linear dynamical models. However, for nonlinear dynamical models, and in particular models with chaotic dynamics, EnKF is superior to ES, due to the fact that the recursive updates keep the model on track and close to the true solution. Thus, ES is not much used and EnKF has been the method of choice in most data assimilation studies where ensemble methods are used. On the other hand, reservoir simulation models are rather diffusive systems when compared to the chaotic dynamical models that were previously used to test ES. If we can assume that the model solution is stable with respect to small perturbations in the initial conditions and the history-matching parameters, then ES should give similar results to EnKF, and ES will be a more efficient and much simpler method to implement and apply. The technical advantages of using ES compared to EnKF are severe, especially when the methods are applied with complex real reservoir models. ES provides a significant reduction in simulation time. Furthermore, a more flexible parameterization is possible, which makes it easier to handle structural and geological model parameters in the history-matching process. In this paper we compare EnKF and ES and show that ES indeed provide for an efficient ensemble-based method for history matching.

Communications

- J.-A. Skjervheim and G. Evensen, *An Ensemble Smoother for assisted History Matching*, SPE Reservoir Simulation Symposium, 21-23 February 2011, The Woodlands, Texas, USA.

- G. Evensen, J.-A. Skjervheim J. Hove and J. G. Vabø, *An Ensemble Smoother for assisted History Matching*, 6th EnKF workshop, Ulvik, Norway.

7 Bioeconomic management - NHH

Sturla F. Kvamsdal, Norwegian School of Economics.

Earlier, the extended Kalman filter and the variational adjoint method were applied to data and a simple growth model of North East Arctic Cod (NEAC). Recently, also the ensemble Kalman filter (EnKF) has

been applied to the same data and model. The method yields relatively stable parameter estimates, and the model is able to capture the dynamics in the data very well. (Figure 4 shows the model solution with 95 percent confidence intervals along with stock and harvest observations; Figure 5 shows parameter estimates with 95 percent confidence intervals.) The EnKF-method has also been applied to a two-dimensional biological model of NEAC and Barents Sea Capelin. While results are preliminary, the model solution seems able to capture the high volatility present in the data. (Figure 6 shows preliminary results of the model solution.) Estimated parameters are already tested in optimal control models for the Barents Sea fisheries, and EnKF-methods plays an important role in the BMAME-project (Bioeconomic Multispecies Analysis of Marine Ecosystems) at the Norwegian School of Economics. Future work in this line of research includes improvement of the two-dimensional biological model and extensions to a three-dimensional model. The EnKF is currently applied to a number of problems, also outside bioeconomics, for example to estimate technical change in production functions for resource-based industries and to estimate a global carbon assimilation function. Potential applications considered are dynamic price equations, colored noise in one-dimensional models, and smoothing problems with nonlinear measurements.

Oral communications:

- "Extensions and Applications of the Ensemble Kalman Filter: Colored Noise, Global Carbon Assimilation, & Marine Ecosystem Management," Lunch Seminar, Norwegian School of Economics, Bergen, March 14, 2011.
- "Managing the Barents Sea Ecosystem," The 2011 North American Association of Fisheries Economists meeting (NAAFE Forum 2011), Honolulu, Hawaii, May 11 - 13, 2011.
- "The Ensemble Kalman Filter in Bioeconomics," eVITA-EnKF meeting, Nansen Center, Bergen, June 16 17, 2011.
- "Managing the Barents Sea Ecosystem," Fisheries Economics and Management Future Challenges (The Warming Conference), Copenhagen, Denmark, Sept 1 - 3, 2011.
- "Revisiting Technological Change and the Tragedy of the Commons: A Kalman Filter Approach," Management Science Seminar, Norwegian School of Economics, Bergen, Sept 23, 2011.

Working papers:

- "Does Species Interaction and Stochasticity Matter for the Optimal Management of Multispecies Fisheries?", Diwakar Poudel, Leif K. Sandal, Stein I. Steinshamn, Sturla F. Kvamsdal, 2011.
- "Revisiting Technological Change and the Tragedy of the Commons: A Kalman Filter Approach," Sturla F. Kvamsdal, Leif K. Sandal, 2011.

8 Petroleum Geostatistics - Ecole des Mines de Paris

In the field of EnKF development and application Hans Wackernagel has supervised the thesis (2007-2010) of Leila Heidari (5) at IFP Energies Nouvelles, which was then reviewed by Geir Evensen in January 2011.

Leila Heidari has proposed to combine the EnKF with two parameterization methods: the pilot point method and the gradual deformation method, which are both capable of preserving second order statistical properties (mean and covariance). Their integration with the EnKF was aimed at ensuring the preservation of the prior model properties during the assimilation process. Both developed algorithms were applied to a simple synthetic case study. For the pilot point method, the application was successful when defining an adequate number of properly positioned pilot points. While for gradual deformation, the method was found to be successful provided the background ensemble is large enough. However, this parameterization increases the inherent nonlinearity in the algorithm and may result in unfavorable effects. For both cases, some improvement scenarios were proposed and further applications to more complex scenarios are recommended.

One paper has been published as a book chapter (6) and a journal paper is close to submission.

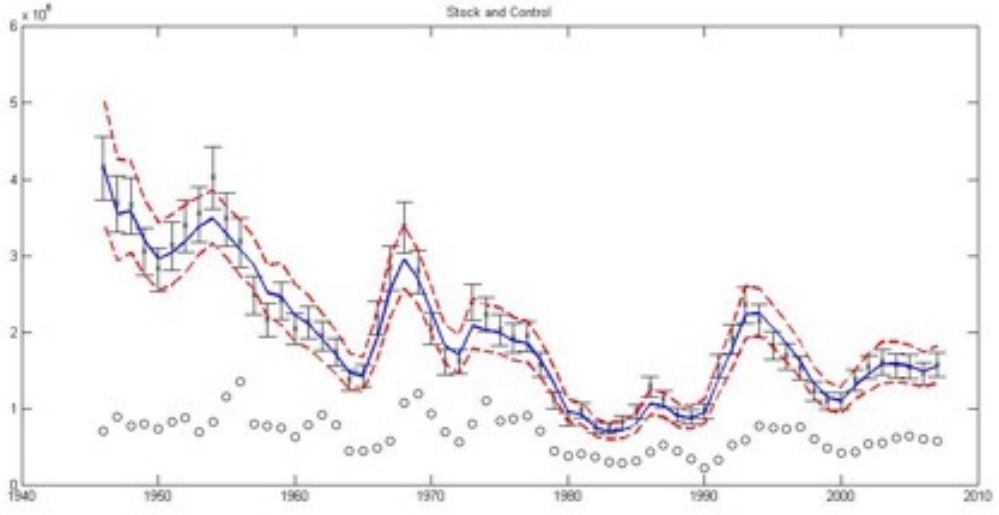

Figure 4: EnKF model solution (blue) with 95 percent confidence intervals (red), stock (x) and harvest (o) observations.

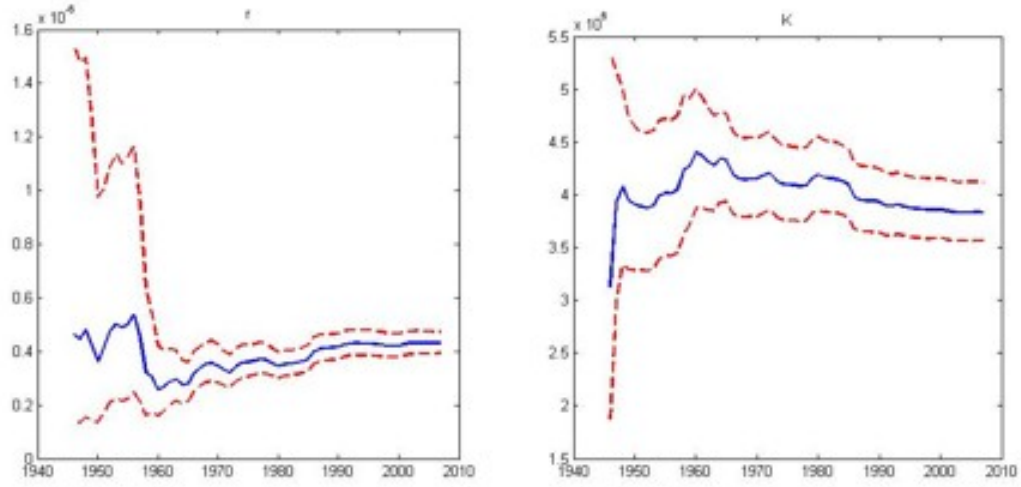

Figure 5: Parameter estimates of the NEAC model.

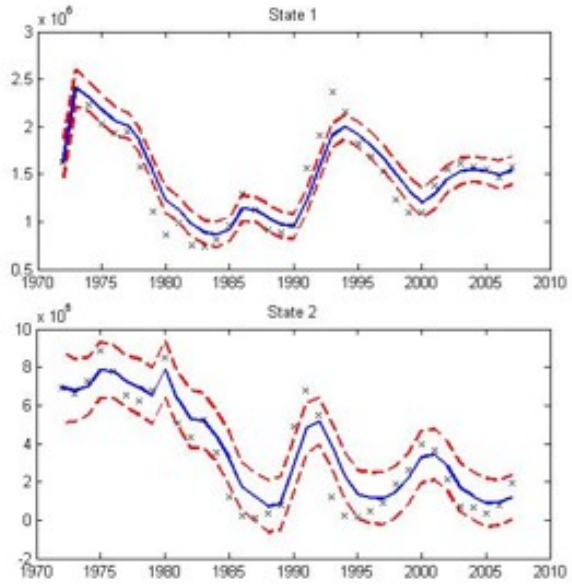

Figure 6: EnKF model solution of a two-dimensional biological model.

9 Spin-off from the eVITA-EnKF project

The eVITA-EnKF project runs in parallel with complementary projects (EU MyOcean, HIRLAM, etc.) until 2012, some parts of the project are already approved for continuation:

- An internal project on oil reservoir history matching at Statoil. Following the eVITA project and additional internal developments within Statoil, Statoil has initiated a commercialization and implementation process of EnKF and ES for history matching, together with external vendors.
- Since July 2011, Dr. J. Bojarova has started as a project leader for Data Assimilation and Use of Observations for HIRLAM-B at met.no, with activities overlapping with eVITA-EnKF.
- The new Center for Climate Dynamics (SKD) at the Bjeknes Center for Climate Research, has approved the PRACTICE project, where the DEnKF is applied to the physical ocean and sea ice components of the Norwegian Earth System Model in view of possible climate "predictions", and not only "scenarios" as presently practiced by climate research centers. L. Bertino is co-manager of the PRACTICE project and F. Counillon is heavily involved, in collaboration with the SKD partners GFI, Uni Research and IMR. The PRACTICE project has started in 2011 and should last for 5 years. The MSC is also a partner in the project BioGeoChemical cycles from the SKD with activities related to the biological oceanography module of eVITA-EnKF.
- A European project lead by University of Bergen (C. Heinze) entitled *CarboChange* has been endorsed and should cover the period 2011-2013. The MSC will implement the EnKF into the coupled ocean-carbon model MICOM-HAMOCC to assimilate observations of ocean carbon and estimate ocean uptake parameters. This work is related to the task "biological oceanography" and will contribute to understanding the carbon cycle in a changing climate. An extension of this project to land and atmopheric Carbon has more recently been approved with the GeoCarbon project, where the MSC is also a partner.
- A European project led by NERSC (J. A. Johannessen) entitled *GREENSEAS* has been endorsed for constructing a database of ocean plankton. The MSC will contribute with data assimilative estimates of plankton build by assimilation of ocean colour data into the couped HYCOM-NORWECOM model. This project is a direct continuation of the module "biological oceanography".
- The continuation of the MyOcean project (MyOcean-II) has been approved for the period 2012-2014, where the MSC keeps the lead of the Arctic Marine Forecasting Center (met.no is also involved for the real-time operations of the TOPAZ and TOPAZ-ECO systems). Norway would probably not have maintained its position in the MyOcean Integrated Project without the considerable contribution from the eVITA-EnKF project. After the completion of MyOcean-II it is expected that a new legal entity, the European Center for Ocean Monitoring and Forecasting (ECOMF) will continue the public marine service, with a Norwegian node for the Arctic.
- A new European GMES project Stochastic Assimilation for the Next Generation Ocean Model Applications (SANGOMA) has been approved for the period 2012-2014, where the MSC will be responsible for the work package on impact studies of new satellite and in situ measurements. The work initiated on parameter estimation in physical ocean models will be continued there.
- A new European Infrastructure project Gliders for Research, Ocean Observation and Management (GROOM) has also been approved for the period 2012-2014, in which the MSC will evaluate the potential impact of a network of ocean gliders for monitoring the North Atlantic. The methodology will heavily rely on the metrics developed during the eVITA-EnKF project (Degrees of Freedom of Signal and Spread Reduction Factor).

Acknowledgments

The project acknowledges the Norwegian supercomputing project NOTUR2 for a grant of computer time on the national High Performance Computing facilities (project `nn2993k`).

References

[1] Jeffrey L. Anderson. Spatially and temporally varying adaptive covariance inflation for ensemble filters. *Tellus A*, 61(1):72–83, 2009.

[2] E. Berge, I. Lie, O. Undheim, A. Holstad, P. Rothmund, and L. Tallhaug. Gurskøy fou, autumn 2006. results from the measuring and modelling campaign, report kvt/eb/2007/001. Technical report, Kjeller Vindteknikk, Kjeller, Norway, 2007.

[3] Geir Evensen. The ensemble Kalman filter for combined state and parameter estimation. *IEEE Trans. Autom. Control*, 29(3):83–104, 2009.

[4] Geir Evensen. *Data assimilation: the Ensemble Kalman Filter*. Springer, 2009 (2nd Ed.).

[5] L Heidari. *History-matching of petroleum reservoir models by the Ensemble Kalman filter and parameterization methods*. PhD thesis, MINES ParisTech, Paris, France, 2011.

[6] L. Heidari, V. Gervais, M. Le Ravalec, and H. Wackernagel. History matching of reservoir models by ensemble Kalman filtering: the state of the art and a sensitivity study. In Y.Z. Ma and P.R. La Pointe, editors, *Uncertainty Analysis and Reservoir Modeling*, volume 96 of *American Association of Petroleum Geologists' Memoirs*, pages 249–264, 2011.

[7] Peter Jan van Leeuwen and Geir Evensen. Data assimilation and inverse methods in terms of a probabilistic formulation. *Mon. Weather Rev.*, 124:2898–2913, 1996.

[8] J. Xie, F. Counillon, J. Zhu, and L. Bertino. An eddy resolving tidal-driven model of the south china sea assimilating along-track sla data using the enoi. *Ocean Science Discussions*, 8(2):873–916, 2011.