MET report
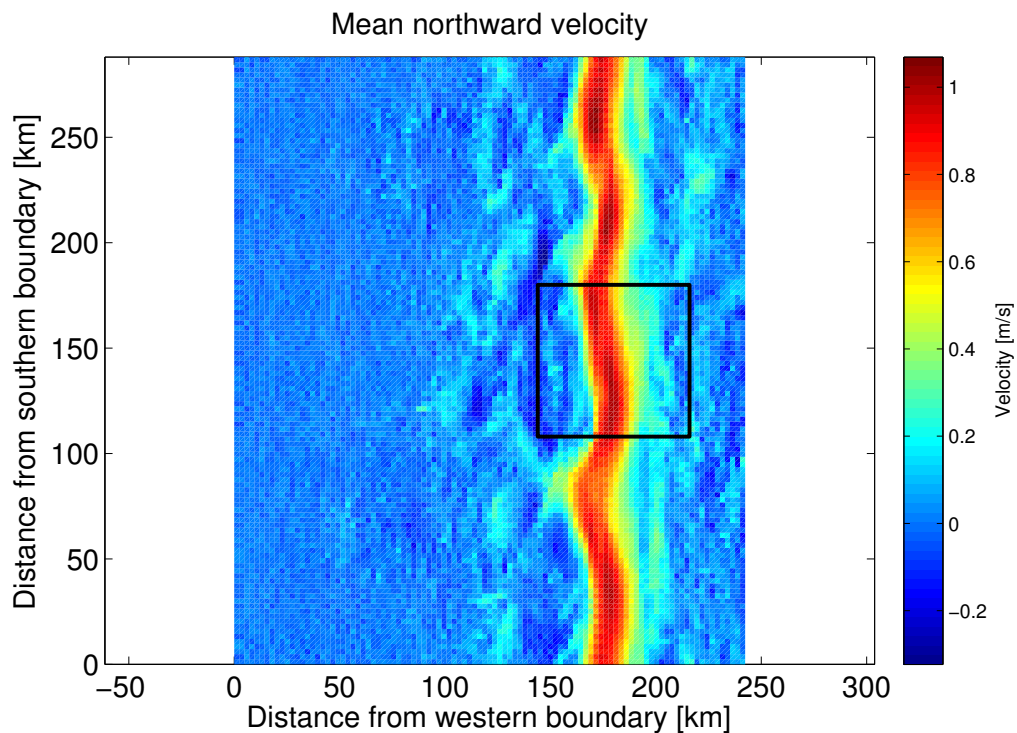
Assimilation of HF radar total current vectors in an idealised version of ROMS-4DVAR

Kai H. Christensen

Mean northward velocity

Norwegian Meteorological Institute

MET report

Title		Date
Assimilation of HF radar total current vectors in an idealised version of ROMS-4DVAR		December 5, 2013
Section		**Report no.**
Oceanography		24
Author(s)		**Classification**
Kai H. Christensen		● Free ○ Restricted
Client(s)		**Client's reference**
Norwegian Clean Seas Association for Operating Companies (NOFO)		Petter Reed

Abstract

This report summarises the results of task 1, activity 1 in the ENI/NOFO project on mobile HF radar systems. The work involves experiments with synthetic HF radar and hydrographic observations using ROMS-4DVAR in an idealised model setup. Tests are made for a case of strongly nonlinear, unstable baroclinic flow along a steep slope. The tests are set up such that the model system is essentially without any systematic error (bias) and the evaluation of the test cases is primarily based on the standard deviation of the analysis error. The main motivation for this work is to test the system so that any technical issues, e.g. with regard to model resolution, vertical mixing schemes, or observation types and errors, can be addressed before work on a realistic model is started. The main findings are:

- Both HF radar and hydrographic observations yield improvements,
- horizontal error correlation scales should be small to avoid spreading information across fronts,
- outer loops should be used to re-linearise adjoint and tangent linear model components,
- short assimilation windows should be used to overcome nonlinearities.

Keywords

Assimilation, ROMS, idealised, synthetic, currents, HF radar

Disciplinary signature

Responsible signature

Contents

1 General model configuration

The basis for the idealised model simulations is the NorKyst800 version[Asplin, 2011] of the Regional Ocean Modeling System (ROMS), which is run operationally by the Norwegian Meteorological Institute. NorKyst800 covers the coast of mainland Norway and has a horizontal resolution of 800 meters. Ocean data assimilation at such high resolution is a formidable task, requiring extensive supercomputing resources, careful analysis to estimate representation errors (which reflects the model's ability to represent the observed dynamics), and a long history of observation data to estimate observation error decorrelation scales. The latter is potentially a serious problem with HF-radar data since observation errors are likely to be correlated in time and/or space[1]. In this study we have therefore chosen to use a coarser horizontal resolution, with horizontal grid cells that are 2.4x2.4 km^2.

Otherwise, the idealised model presented here closely resembles NorKyst800, for example with regard to vertical coordinates and choice of vertical mixing scheme. It should be noted that the forcing of the model is done by specifying all the necessary momentum, radiative, and freshwater fluxes, which means that no bulk flux schemes are used. Also, salinity is kept constant and all changes in water density are caused by changes in temperature. A linear equation of state is used.

There are seven dependent variables in ROMS: $(u, v, T, S, \zeta, \bar{u}, \bar{v})$, representing horizontal velocity in easterly direction, horizontal velocity in northerly direction, potential temperature, salinity, vertically averaged velocity in easterly direction, vertically averaged velocity in northerly direction, and sea surface height, respectively. In the ROMS code and ROMS output these variables are referred to as (`u`, `v`, `temp`, `salt`, `zeta`, `ubar`, `vbar`), respectively.

1.1 Model grid

The model domain is configured as a channel with periodic north-south boundary conditions and solid walls along the eastern and western boundaries, see figure 1. The grid is Cartesian with a horizontal resolution of 2.4 kilometers. We use the f-plane approximation with constant Coriolis parameter equivalent to 65° north. The domain size is 100x120 interior grid points (easterly/northerly directions), and we use 35 vertical levels.
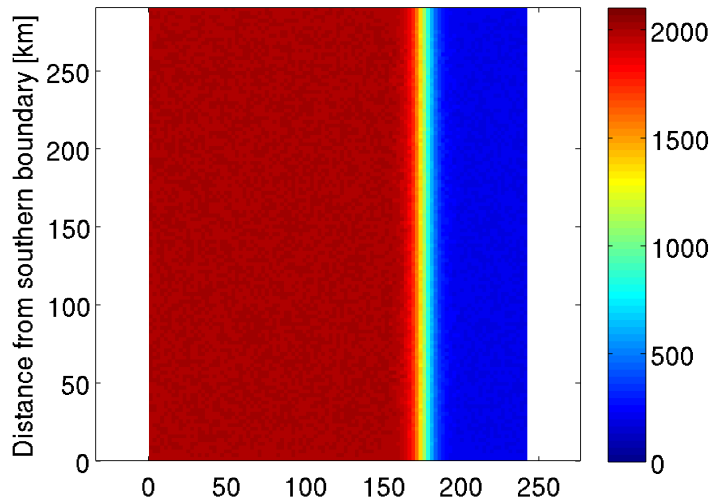
---

[1]W. G. Zhang, pers. comm.

Figure 1: Model domain and topography. The variable `hraw` is shown, i.e. before the internal smoothing that is applied in ROMS. Depth is in meters.

The main topographic features are (i) a shelf along the eastern boundary, with average depth of 200 meters, and width of approximately 70 kilometers, (ii) a sharp shelf break with a hyperbolic tangent profile (maximum slope is approximately 0.1), and (iii) a deep ocean floor width average depth of approximately 1800 meters towards the western boundary. In order to trigger instabilities, random noise is added to the depth with maximum amplitudes of 50 meters.

1.2   Initialisation and forcing of the model

The initial conditions are uniform with $T = 10$, $S = 35$, $\zeta = 0$ and zero velocities. The model is run applying constant heat fluxes over a period of 150 days. The aim is to produce unstable baroclinic currents that will be guided by topography. To accomodate this, we apply a net bottom heat flux into the ocean on the shelf, and a net surface heat flux out of the ocean over the deep ocean Isachsen [2011]. These fluxes are balanced such that the net heat flux into the ocean is zero. Since the deep ocean is wider than the shelf, the surface fluxes are smaller than the bottom fluxes (approximately 160 Wm^{-2} and 440 Wm^{-2}, respectively). The momentum and net freshwater fluxes are both set to zero.

The water heated over the shelf bottom is rapidly mixed upwards through the entire water column resulting in a sharp temperature front near the shelf break (see figs. 2 and 3). Due to geostrophic adjustment a northward flowing, topographically controlled slope current develops. This current is baroclinically unstable
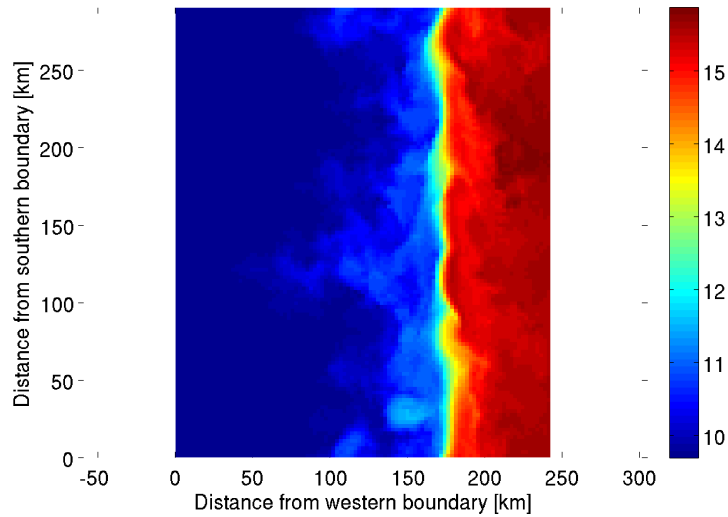
2

Figure 2: Sea surface temperature at the end of the spinup period. The values are given in degrees Celsius.

and heat exchange with the deep water region is facilitated by macro turbulence (ocean eddies).
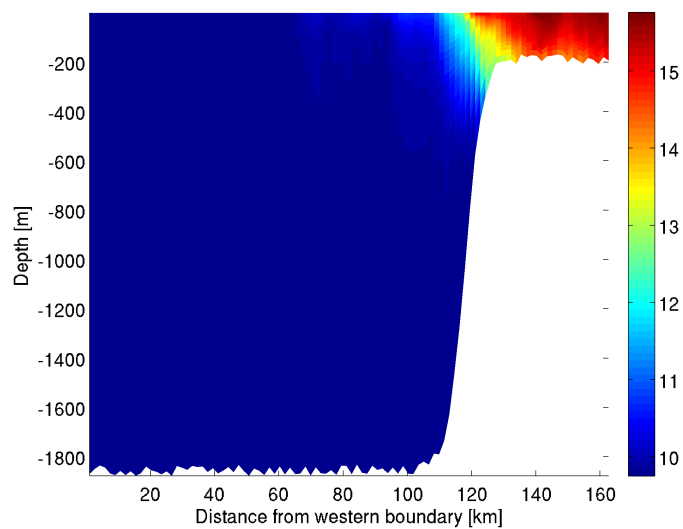

Figure 3: A vertical east-west section in the middle of the basin showing the temperature distribution at the end of the spinup period. Note the sharp front over the shelf break. Values are given in degrees Celcius.

1.3 Sensitivity tests

The domain is periodic so that the water masses flowing out of the domain at the northern boundary reappear at the southern boundary. It is of interest for the assimilation experiments to investigate how quickly the various dynamical features propagate through the domain. As shown in fig. 4, the mean northward surface velocity has a maximum of about 1 m/s over the slope, which means that a drifting object can potentially be advected through the domain in about three days. The flow is also characterised by baroclinic and barotropic waves and eddies, however, with propagation speeds that are generally very different from the mean flow speeds. Figure 4 also shows a region of special interest: it is here that observations are made, both hydrographic and HF-radar currents. The statistical analysis used to assess the performance of the data assimilation system is restricted to this region (see Sec. 3).

To investigate how upstream conditions influence the dynamics in this region we perform an adjoint sensitivity studyMoore et al. [2009], Zhang et al. [2009]. Such studies can reveal what variables influence a specific quantity of interest in a specific region and at a specific time. We focus on the advection of water masses and investigate the sensitivity of the surface velocity variance to temperature, choosing a short section over the slope in the middle of the domain. Figures 5 and 6 show how the influence reaches increasingly further upstream as we move backwards in time, but that the main upstream influence is restricted to the southern part of the domain. Based on these results we conclude that an upper limit of three days for the assimilation experiments is adequate.
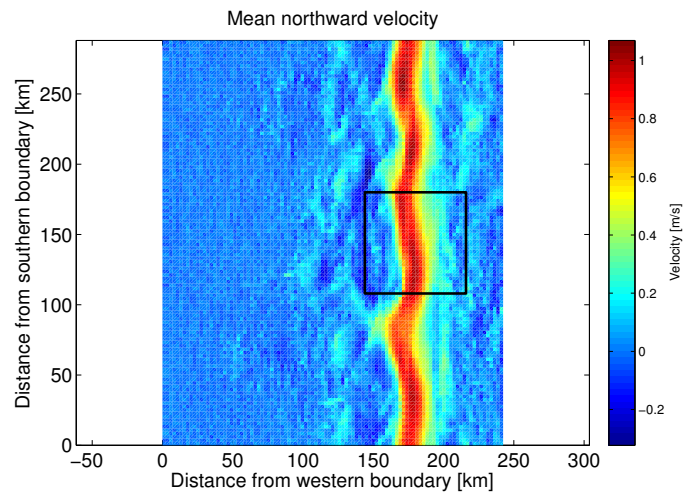
Figure 4: The mean northward surface velocity during a three days simulation. The region enclosed by the black line is used for assessing the performance of the data assimilation system, as further discussed in Sec. 3.
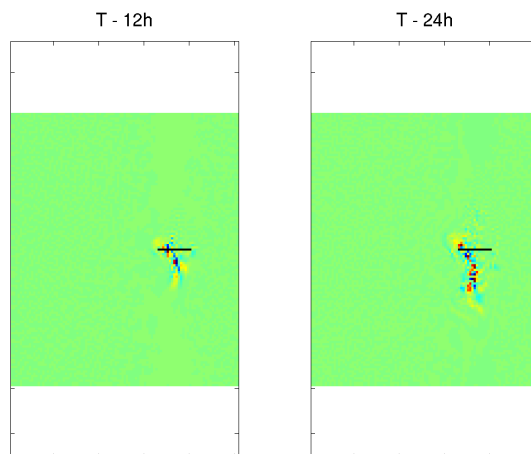

Figure 5: The panels show the sensitivity of surface velocity variance in the short section denoted by the black line. The sensitivity to temperature at intermediate depths (level 18) is shown. Put more simply, the panels indicate where the water masses come from that will cause changes in the surface velocities (at the section) at time $T = 0$, in this case showing where these water masses are 12 and 24 hours ahead of $T = 0$, respectively.
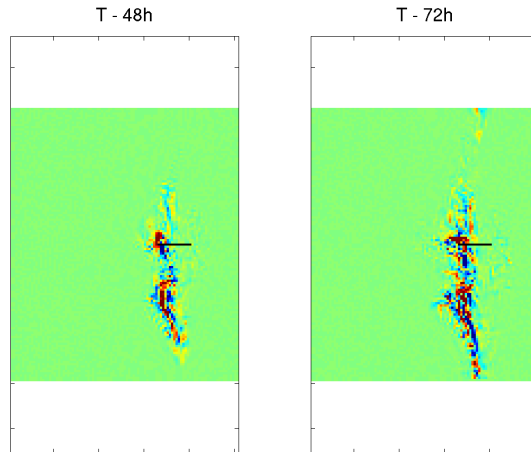
Figure 6: Same as fig. 5, but for $T = -48$ h and $T = -72$ h, respectively.

2 Configuration of the data assimilation system

ROMS contains a four dimensional variational data assimilation system (ROMS-4DVAR) that accepts observations in any of the dependent variables Moore et al. [2011]. As yet the 4DVAR system does not allow assimilation of radial currents from HF radars. In our experiments, data assimilation is only considered for interior model points and no adjustment of the surface/bottom fluxes or boundary conditions are made. We use incremental, strong-constraint 4DVAR. ROMS-4DVAR also has options for multivariate background error correlations (e.g. temperature observations can directly influence sea surface height), but since the underlying theories are dubious for high latitudes and eddy resolving models, we do not make use of any such options here. The estimates of background (model) errors are taken from day 120 to 150 of the spinup period, using the standard deviation of each variable in each grid point. The vertical error correlation lengths for all variables is set to 30 m.

The tuning of ROMS-4DVAR focuses on three components of the assimilation system:

- **Horizontal background error correlation scales:** These control the influence radius of a single observation, i.e. a small scale implies that an observation will only have direct influence in a small area centered around the observation. The actual influence of the observation is weighted by the estimates of background (model) and observation errors.

- **Inner/outer loops:** Central to 4DVAR is the forward and backward time integration using linearised model dynamics. A pair of forward (tangent lin-

ear) and backward (adjoint) model integrations is called an *inner loop*. If the problem is strongly nonlinear, a temporary solution can be used to re-linearise the tangent linear and adjoint models, which is referred to as an *outer loop*.

- **Assimilation window length:** The 4DVAR system relies on the linearised models being reasonably accurate during the assimilation period. If the problem is strongly nonlinear, shorter assimilation windows may have to be used.

2.1 Synthetic observations

Synthetic observations are taken from a separate model simulation that has been forced with time varying momentum and heat fluxes, see fig. 7. An example of the difference between the simulation used for synthetic observations and the simulation that forms the basis for 4DVAR is shown in fig. 8.

An observation error is assigned to each synthetic observation, but no random or systematic errors are added to the observation values. The observation errors are used in the assimilation system to give the observations proper weight compared to the background (model) values.

The observation locations are shown in fig. 9. Hydrographic observations, e.g. such as taken from a research vessel with a CTD, are simulated by taking a single vertical profile of temperature every hour, zigzagging southwards making four sections across the slope. A total of 64 temperature profiles are processed and each individual observation is assigned a constant error of 0.05 K. Two simulated HF-radars are used to provide hourly total current vectors in 61 locations. More details about these synthetic observations are presented in the following section.
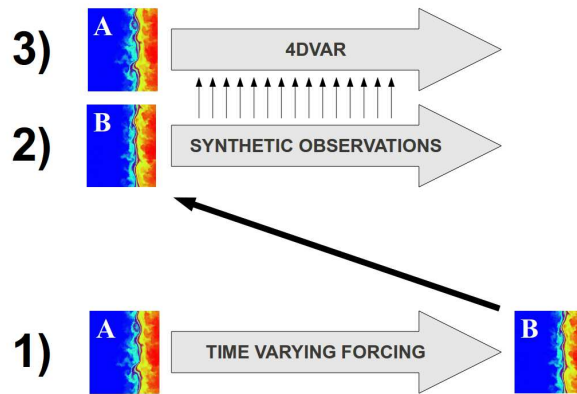
Figure 7: After the spin-up period we have a restart field "A". A separate simulation (1) is then forced with time varying heat and momentum fluxes to produce a different restart field "B". The restart field "B" is then used for a simulation (2) from which synthetic observations are taken, and which represent the "true" state of the ocean. The original restart field "A" is used as a starting point for simulation (3) with ROMS-4DVAR to assimilate the synthetic observations. Ideally, the entire simulation (3) should become closer to simulation (2) as a result of assimilating the synthetic observations.
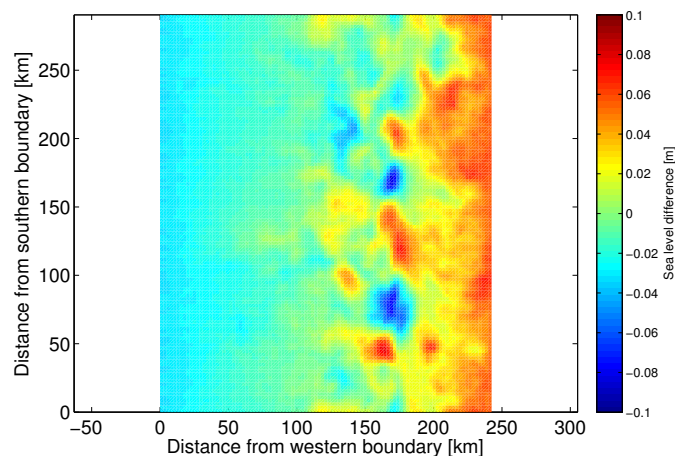

Figure 8: Example of the difference between the restart fields "A" and "B", here showing sea surface height.
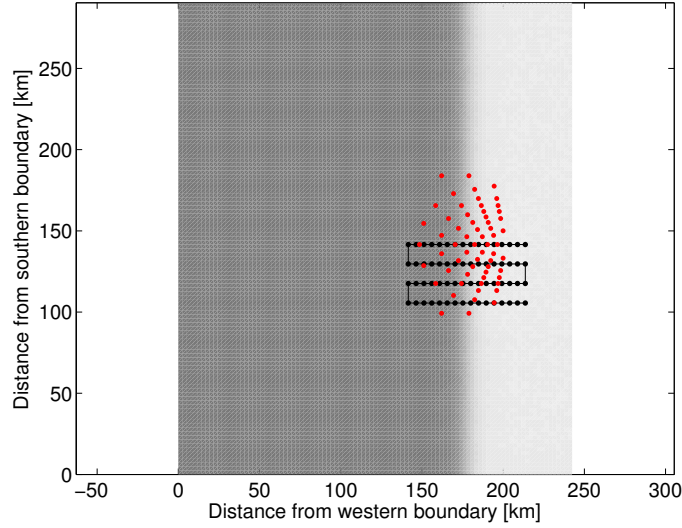
Figure 9: The black dots show where temperature profiles are taken. One full vertical profile is taken every hour, and four sections are made, starting from the north. The red dots are the intersection points of the two simulated HF-radars. Hourly observations of total current vectors are produced for the assimilation. The water depth is indicated in gray.

2.1.1 Simulated HF-radar observations

We simulate two HF-radar stations positioned a distance of y_a = 118 km and y_b = 166 km from the southern boundary. We assume that the radars retrieve radial currents at an effective depth D_e = 2 m, with an azimuthal resolution of $\Delta\theta$ = 11.25 degrees. Furthermore we assume that the maximum range of the radars is R = 80 km, and that the observation error σ_R associated with the radial currents is a linear function of radial distance r. The azimuthal resolution determines the number of beam directions, and, combined with the maximum range R, also the number of intersecting beams from which we can estimate total current vectors (see Fig. 10).

To obtain the synthetic HF-radar observation errors we have first used standard vector algebra to determine the positions were the beams intersect and then calculated the errors in the easterly and northerly directions $(\sigma_{GDOP}^{(E)}, \sigma_{GDOP}^{(N)})$ due geometric dilution of precisionChapman et al. [1997], see fig. 11.

The total errors are then assumed given by

$$(\sigma_{tot}^{(E)}, \sigma_{tot}^{(N)}) = \left(\sigma_{GDOP}^{(E)}\sigma_R|v^{(E)}|, \sigma_{GDOP}^{(N)}\sigma_R|v^{(N)}|\right) \tag{1}$$

where v is the observed speed. The relative observation errors, i.e. the product $\sigma_{GDOP}\sigma_R$, are shown in fig. 12. The errors in a real HF-radar system are more
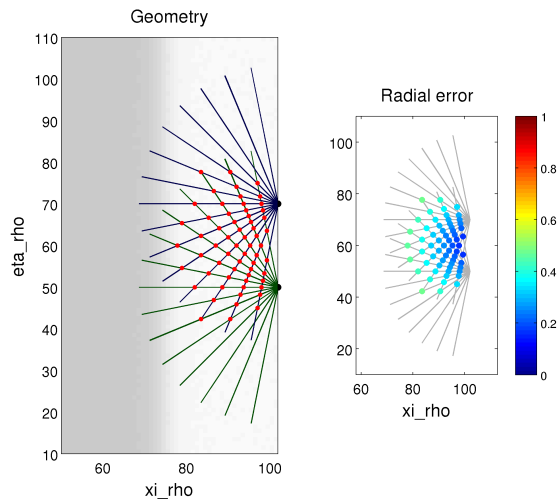
Figure 10: The left panel shows the radar beams and the intersection points where total current vectors are obtained. The right panel shows the observation error σ_R that is attributed to the distance from the HF-radar stations.
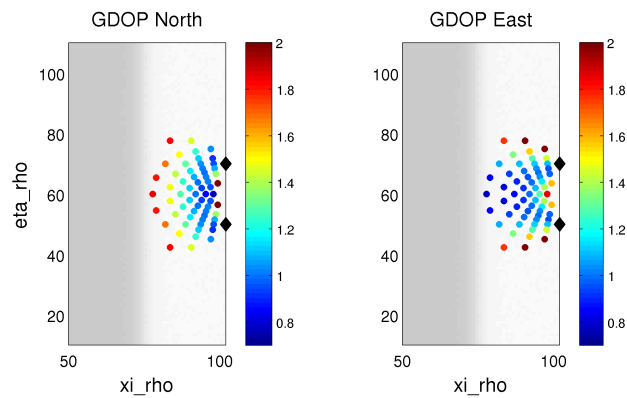

Figure 11: The left panel shows the GDOP error in the northerly component of the total current vector, while the right panel shows the GDOP error in the easterly component.
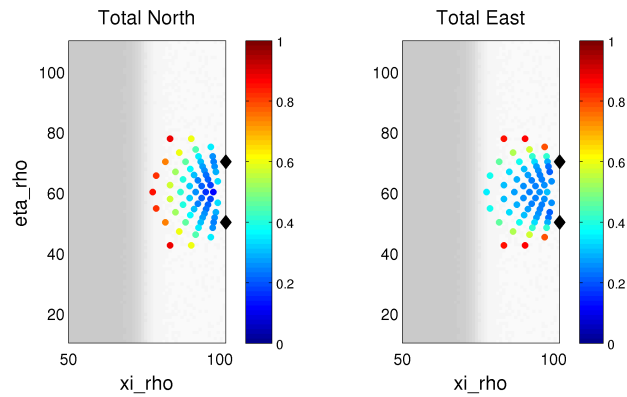
Figure 12: The panels show the total relative error ($\sigma_{GDOP}\sigma_R$) in the northerly and easterly directions, respectively.

complex than those given by (1), but a comparison with data from the CODAR system deployed near Fedje demonstrate that the values obtained from (1) are reasonable (see figs. 13 and 14). The data from the Fedje system are from the period Feb. 1 to Apr. 30, 2009. HF-radar observations from the Finnmark system collected late autumn 2012 possess similar error statistics (not shown here).

Finally, in order to cover the region with energetic currents near the slope, the entire simulated HF-radar system is translated approximately 40 km westwards from the eastern boundary (compare figs. 9 and 10).
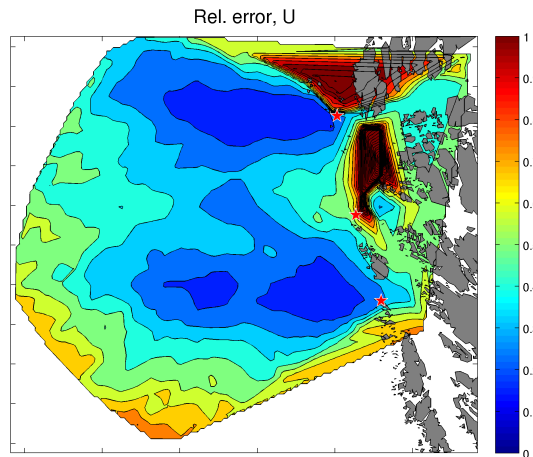
Figure 13: The figure shows the average relative error, i.e. error divided by velocity value, for the easterly component of the total current vector. The data are taken from the Fedje site, using all observations made in the period Feb. 1 to Apr. 30, 2009. The stations are marked by stars.
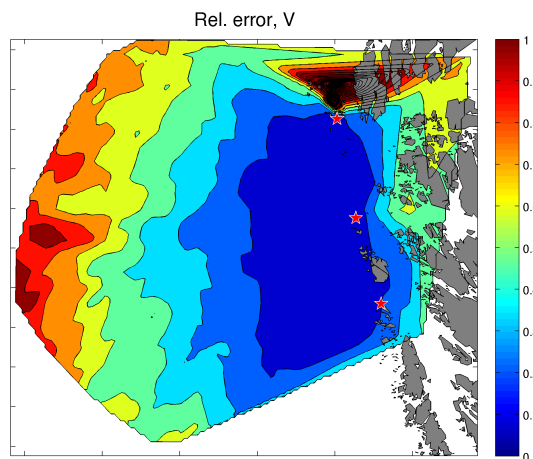

Figure 14: The figure shows the average relative error, i.e. error divided by velocity value, for the northerly component of the total current vector. The data are taken from the Fedje site, using all observations made in the period Feb. 1 to Apr. 30, 2009. The stations are marked by stars.

12

3 Results

We only consider the velocities (u, v) and the temperature T in the evaluation of the data assimilation system. Furthermore, we only consider a limited region similar to the HF-radar coverage area (see fig. 4), and also restrict the analysis to the two uppermost vertical levels of the model.

The procedure is straightforward: each numerical experiment results in an *analysis*, i.e. a solution of ROMS-4DVAR. This solution is compared with the "truth" as represented by the simulation that provides the synthetic observations. Average bias and standard deviation for each of the variables (u, v, T) are calculated based on all grid points and all output times in the region of interest.

Due to the way the experiments are designed, the (u, v, T) variables have little bias compared to what we would expect from a realistic model, and the standard deviation is used when deciding what options that yield the best results.

3.1 Numerical experiments

A total of 15 different numerical experiments are made as outlined in Table 1.

3.2 Error correlation scales

Figures 15-17 show the results from the experiments with different horizontal error correlation scales. It is clear that the smaller scale (5 km) yields the best results. One likely explanation is that observation values are erroneously distributed across sharp gradients when the larger correlation scales are used.

Exp. no.	Corr. scale	Obs. types	BG. error	Inner/outer	Assim. window
0	N/A	N/A	N/A	N/A	N/A
1	5 km	All	Normal	5/4	72 h
2	10 km	All	Normal	5/4	72 h
3	20 km	All	Normal	5/4	72 h
4	Best	All	Normal	20/1	72 h
5	Best	All	Normal	10/2	72 h
6	Best	All	Normal	4/5	72 h
7	Best	All	Normal	6/5	72 h
8	Best	HF	Normal	Best	72 h
9	Best	Temperature	Normal	Best	72 h
10	Best	All	Inflated	Best	72 h
11	Best	HF	Inflated	Best	72 h
12	Best	Temperature	Inflated	Best	72 h
13	Best	All	Normal	6/5	24 h
14	Best	All	Normal	10/2	6 h
15	Best	All	Normal	6/5	6 h

Table 1: The numerical experiments used for evaluation of the data assimilation system. From left to right the columns denote horizontal error correlation scales; observation types; background (model) error estimates; number of inner and outer loops; length of assimilation window. Exp. 0 is a control run without data assimilation.
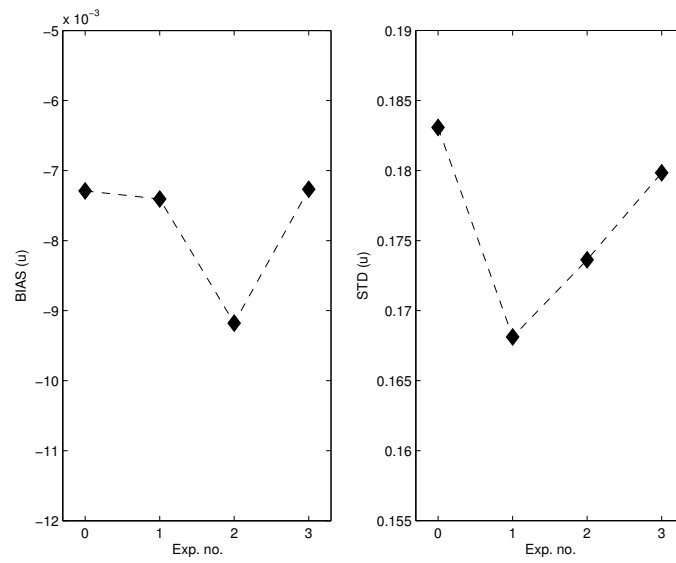
Figure 15: Bias and standard deviation of u for the experiments with different horizontal error correlation scales. Exp 0: no assimilation; Exp 1: 5 km; Exp 2: 10 km; Exp 3: 20 km.
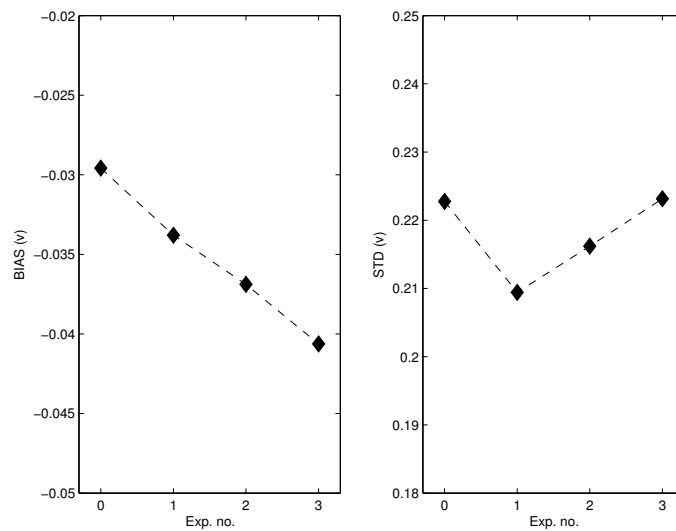

Figure 16: Bias and standard deviation of v for the experiments with different horizontal error correlation scales. Exp 0: no assimilation; Exp 1: 5 km; Exp 2: 10 km; Exp 3: 20 km.
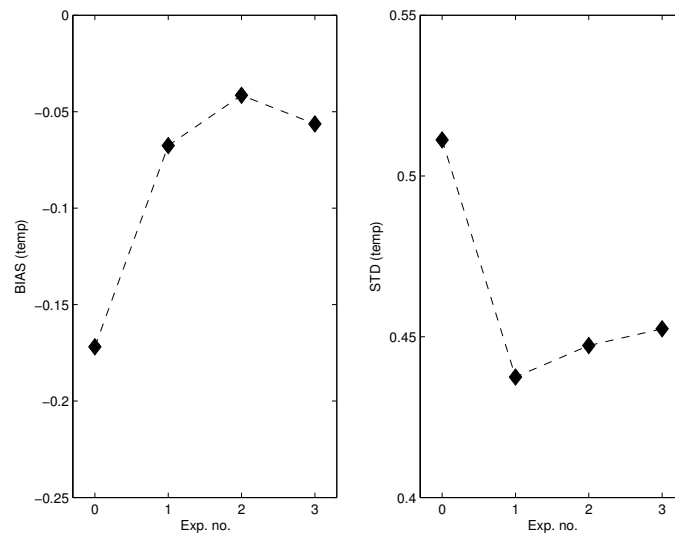
Figure 17: Bias and standard deviation of T for the experiments with different horizontal error correlation scales. Exp 0: no assimilation; Exp 1: 5 km; Exp 2: 10 km; Exp 3: 20 km.
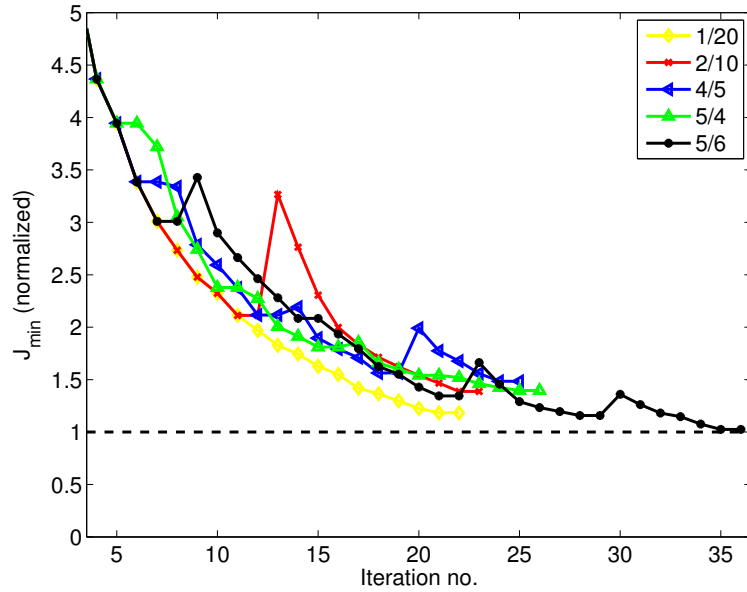
Figure 18: The convergence of the normalised cost function minimum is shown for experiments 1 and 4-7.

3.3 Inner and outer loops

The next set of experiments aim to determine the number of inner and outer loops that should be used. In 4DVAR the analysis is obtained by minimising a cost function, and the convergence of the cost function minimum, J_{min}, is shown in fig. 18. It appears that using outer loops helps reducing the value of J_{min}, which indicates that a more optimal solution is found. In the last (and best) of these experiments, with 6 inner and 5 outer loops, the value of the normalised cost function minimum comes very close to 1, which is an indication that our error statistics are properly specifiedWeaver et al. [2002]. It should be noted that increasing the number of outer loops at the expense of reducing the number of inner loops is not necessarily the best option, even though the final analysis will be slightly better. The representation of e.g. the gain matrix in ROMS-4DVAR is based on the so-called Lanczos vectorsMoore et al. [2011] and each inner loop produces one such vector. The Lanczos vectors are e.g. useful for *a posteriori* analysis of system performance.

3.4 Impact of different observation sets

The next set of experiments is made to evaluate the impact of the different observation types and the results are shown in figs. 19-21. It is interesting to note that the errors in both u and v are smallest when only HF-radar observations are
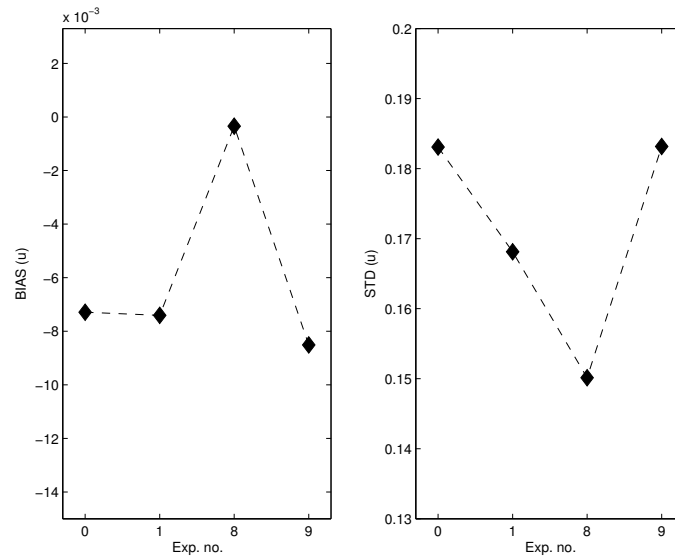
Figure 19: Bias and standard deviation of u for the experiments with different observation sets. Exp 0: no assimilation; Exp 1: both temp./HF; Exp 8: HF only; Exp 9: temp. only.

assimilated. The error in T is also smaller when only HF-radar observations are assimilated compared to when only temperature profiles are assimilated. One likely explanation is that the HF-radar observations improves the positioning of the eddies, and as these carry water masses the temperature fields also improve.
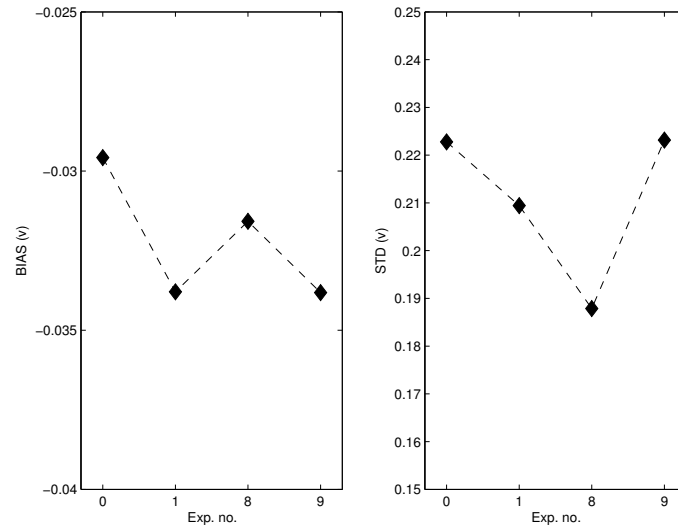
Figure 20: Bias and standard deviation of v for the experiments with different observation sets. Exp 0: no assimilation; Exp 1: both temp./HF; Exp 8: HF only; Exp 9: temp. only.
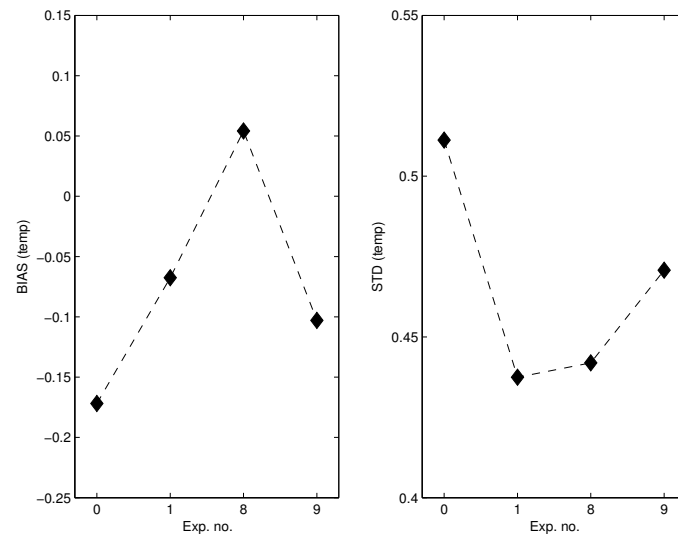

Figure 21: Bias and standard deviation of T for the experiments with different observation sets. Exp 0: no assimilation; Exp 1: both temp./HF; Exp 8: HF only; Exp 9: temp. only.

3.5 Impact of inflated background error

We also made several experiments to investigate how sensitive the solution is to the specification of the background error. In these experiments we used inflated background error variances so that the normal values were increased by a factor 3. Figures 22-24 show the difference between using inflated and normal background error statistics. The overall impact is small, in fact it is only when both observation sets are used that any improvement can be seen.
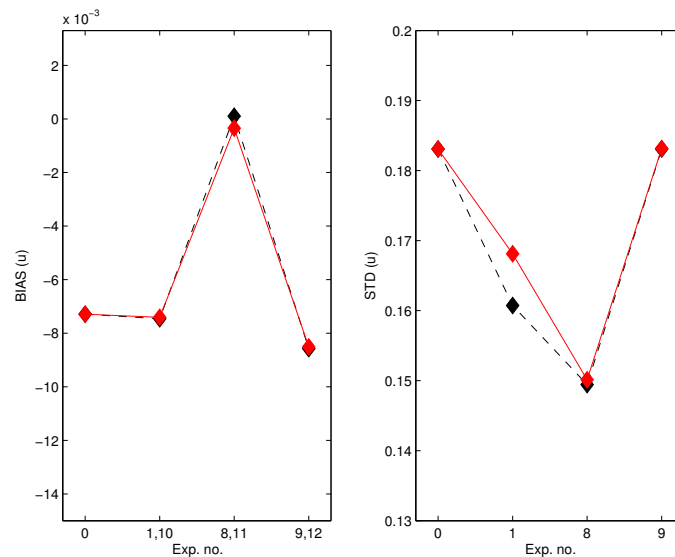

Figure 22: Bias and standard deviation of u for the experiments with inflated background error (red solid line: normal errors, black dashed line: inflated errors). Exp 0: no assimilation; Exp 1/10: both temp./HF; Exp 8/11: HF only; Exp 9/12: temp. only.
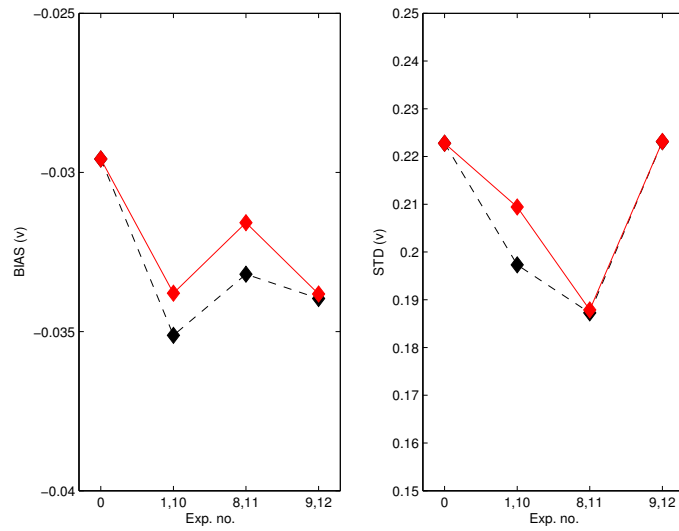
Figure 23: Bias and standard deviation of v for the experiments with inflated background error (red solid line: normal errors, black dashed line: inflated errors). Exp 0: no assimilation; Exp 1/10: both temp./HF; Exp 8/11: HF only; Exp 9/12: temp. only.
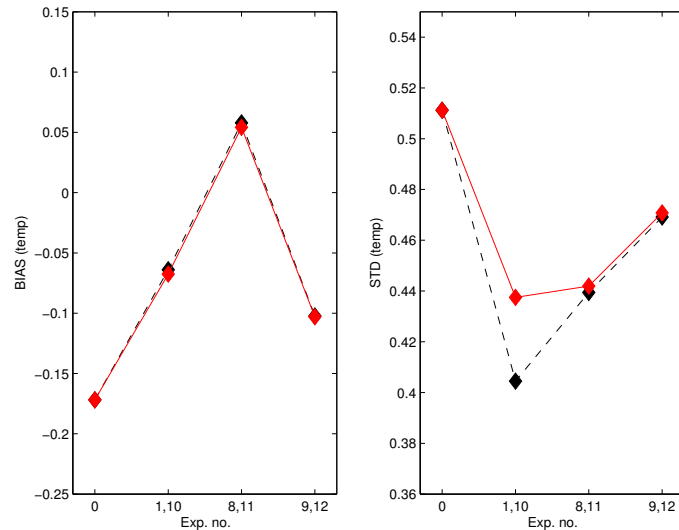

Figure 24: Bias and standard deviation of T for the experiments with inflated background error (red solid line: normal errors, black dashed line: inflated errors). Exp 0: no assimilation; Exp 1/10: both temp./HF; Exp 8/11: HF only; Exp 9/12: temp. only.

3.6 Length of assimilation window

The last experiments are made to investigate different assimilation window lengths. The results are shown in figs. 25-27. Reducing the window length from 72 h to 24 h yields a slight improvement, but a further reduction to 6 h is not beneficial.
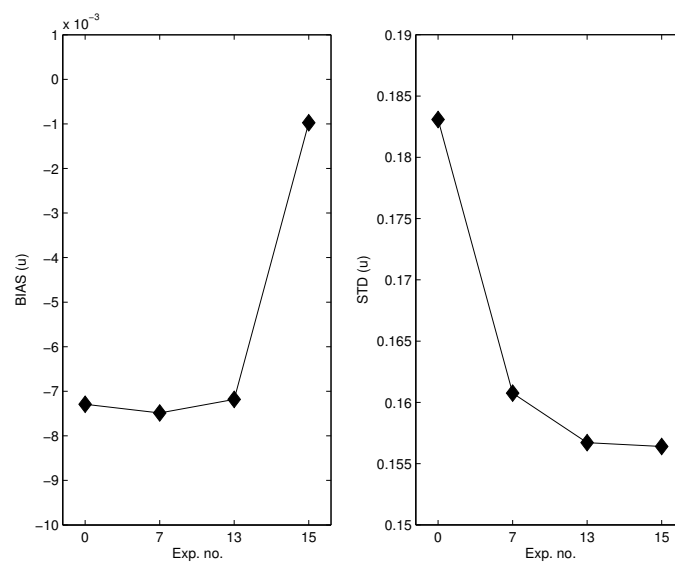

Figure 25: Bias and standard deviation of u for the experiments with different assimilation window lengths. Exp 0: no assimilation; Exp 7: 72 h; Exp 13: 24 h; Exp 15: 6 h.

Figure 26: Bias and standard deviation of v for the experiments with different assimilation window lengths. Exp 0: no assimilation; Exp 7: 72 h; Exp 13: 24 h; Exp 15: 6 h.
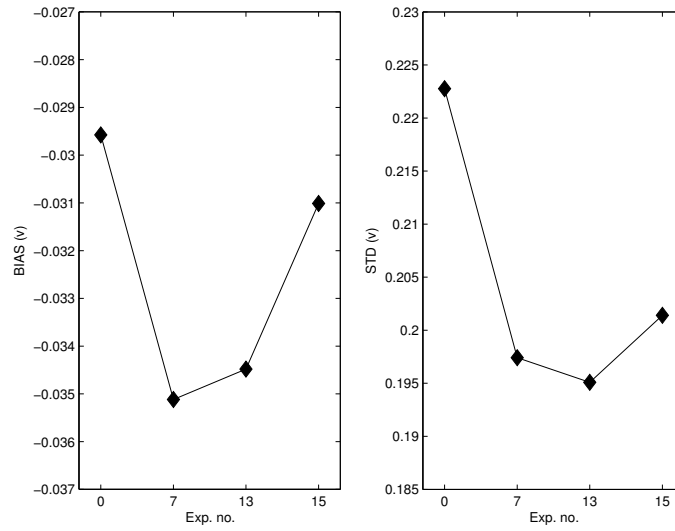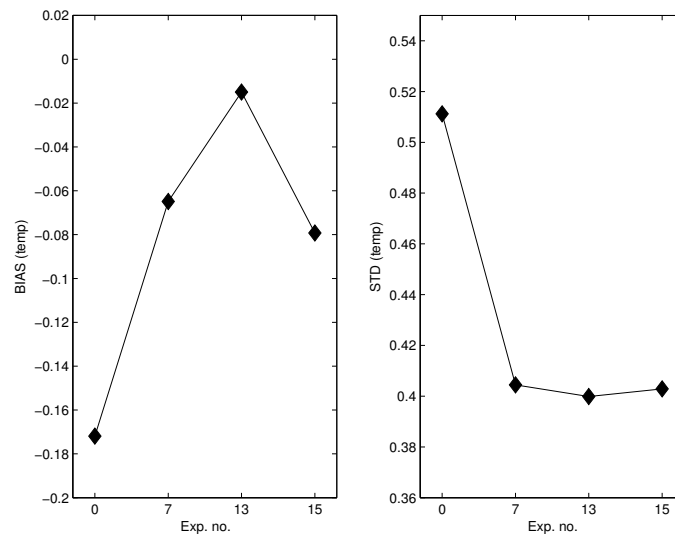

Figure 27: Bias and standard deviation of T for the experiments with different assimilation window lengths. Exp 0: no assimilation; Exp 7: 72 h; Exp 13: 24 h; Exp 15: 6 h.

4   Concluding remarks

A total of 15 different numerical experiments with ROMS-4DVAR has been made. The main aim of these experiments has been to test the use of HF-radar observations in a model setup that is similar to the operational ocean model systems at the Norwegian Meteorological Institute, and to explore the various options of the ROMS-4DVAR system to aid the implementation of a realistic model with data assimilation.

Both synthetic HF-radar currents and synthetic hydrographic profiles have been assimilated. Both data sets yield an improvement, and the data assimilation system works well in combination with the general configuration of the ocean model (i.e., vertical coordinates, mixing schemes, etc.).

For a realistic model it is recommended that (i) small horizontal error correlation scales should be used, (ii) outer loops should be used, and (iii) short assimilation windows should be used. It is also recommended that the nonlinear model should produce at least hourly output for tangent linear and adjoint models to ensure that linearised solutions do not deviate significantly from the nonlinear solutions.

The test case considered in this study focuses on a specific mode of ocean circulation and the length of the experiment is short. The quantitative results obtained here are therefore not necessarily representative of the results that will be obtained when using realistic ocean model setups. Real flows are more complex and e.g. storm surge and tides are not considered here. Nevertheless, the test case represents a specific type of ocean dynamics that is very challenging to model, and the fact that the data assimilation system yields an improvement is encouraging.

Acknowledgements

Bibliography

J. Albretsen A.K. Sperrevik A. Staalstrøm A. D. Sandvik F. Vikebø L. Asplin. Norkyst-800 report no. 1 - user manual and technical descriptions. Technical report, Fisken og Havet 2/2011, Institute of Marine Research, 2011.

R. D. Chapman, L. K. Shay, H. C. Graber, J. B. Edson, A. Karachintsev, C. L. Trump, and D. B. Ross. On the accuracy of hf radar surface current measurements: Intercomparisons with ship-based sensors. *J. Geophys. Res.*, 102(C8): 18737–18748, January 1997. ISSN 2156-2202. URL `http://dx.doi.org/10.1029/97JC00049`.

P.E. Isachsen. Baroclinic instability and eddy tracer transport across sloping bottom topography: How well does a modified eady model do in primitive equation simulations? *Ocean Modelling*, 39(1-2):183–199, 2011. ISSN 1463-5003. URL `http://www.sciencedirect.com/science/article/pii/S1463500310001460`.

Andrew M. Moore, Hernan G. Arango, Emanuele Di Lorenzo, Arthur J. Miller, and Bruce D. Cornuelle. An adjoint sensitivity analysis of the southern california current circulation and ecosystem. *J. Phys. Oceanogr.*, 39(3):702–720, March 2009. ISSN 0022-3670. URL `http://dx.doi.org/10.1175/2008JPO3740.1`.

Andrew M. Moore, Hernan G. Arango, Gregoire Broquet, Brian S. Powell, Anthony T. Weaver, and Javier Zavala-Garay. The regional ocean modeling system (roms) 4-dimensional variational data assimilation systems: Part i - system overview and formulation. *Progress in Oceanography*, 91(1):34–49, 2011. ISSN 0079-6611. URL `http://www.sciencedirect.com/science/article/pii/S0079661111000516`.

A.T. Weaver, J. Vialard, D.L.T. Anderson, and P. Delecluse. Three- and four-dimensional variational assimilation with a general circulation model of the tropical pacific ocean. Technical Report 365, European Centre for Medium Range Weather Forecasts (ECMWF), March 2002.

Weifeng G. Zhang, John L. Wilkin, Julia C. Levin, and Hernan G. Arango. An adjoint sensitivity study of buoyancy- and wind-driven circulation on the new jersey inner shelf. *J. Phys. Oceanogr.*, 39(7):1652–1668, July 2009. ISSN 0022-3670. URL `http://dx.doi.org/10.1175/2009JPO4050.1`.